



**HAL**  
open science

## Conférence Nationale d'Intelligence Artificielle Année 2019

Caroline Chanel, Tiago de Lima, Sylvie Doutre, Amal Elfallah-Seghrouchni,  
Frederick Garcia, Nathalie Jane Hernandez, Jérôme Lang, Maxime Lefrançois,  
Juliette Mattioli, Emmanuel Rachelson, et al.

► **To cite this version:**

Caroline Chanel, Tiago de Lima, Sylvie Doutre, Amal Elfallah-Seghrouchni, Frederick Garcia, et al..  
Conférence Nationale d'Intelligence Artificielle Année 2019 : Activité AFIA 1er août 2017 – 31 juillet  
2019. Association Française d'Intelligence Artificielle, 2019. hal-02458049

**HAL Id: hal-02458049**

**<https://hal.science/hal-02458049>**

Submitted on 28 Jan 2020

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# Conférence Nationale d'Intelligence Artificielle Année 2019

## Sections Spéciales

Caroline Chanel  
Tiago de Lima  
Sylvie Doutre  
Amal Elfallah-Seghrouchni  
Frédéric Garcia  
Nathalie Hernandez  
Jérôme Lang  
Maxime Lefrançois  
Juliette Mattioli  
Emmanuel Rachelson  
Olivier Simonin  
Pierre Zweigenbaum

**Coordination AfIA**  
Yves Demazeau  
Dominique Longin



**Conférence Nationale  
d'Intelligence Artificielle  
Année 2019**





# **Actes CNIA 2019**

Conférence Nationale d'Intelligence Artificielle

Éditeurs : Président(e)s des CP de CNIA 2019



# Table des matières

Mot des Présidents.

**Toulouse, au cœur de l'Intelligence Artificielle de demain** ..... 5

Comités.

**Comités 2019 de programmation, de programme, et d'organisation** ..... 7

Partenaires.

**Partenaires PFIA 2019** ..... 15

A. Aubret, L. Matignon, S. Hassas (JFPDA).

**Étude de la motivation intrinsèque en apprentissage par renforcement** ..... 17

Q. Baert, A.-C. Caron, M. Morge, J.-C. Routier, K. Stathis (JFSMA).

**Stratégie situationnelle pour l'équilibrage de charge** ..... 27

S. Belabbes, S. Benferhat, J. Chomicki (JIAF).

**Elect : An Inconsistency Handling Approach for Partially Preordered Lightweight Ontologies** ..... 35

Q. Cohen-Solal (CNIA).

**Apprendre à jouer aux jeux à deux joueurs à information parfaite sans connaissance** ..... 43

Y. Dimopoulos, J.-G. Maily, P. Moraitis (JIAF).

**Argumentation-based negotiation with incomplete opponent profiles** ..... 51

N. Gauville, F. Charpillat (JFSMA).

**Exploration et couverture par stigmergie d'un environnement inconnu avec une flotte de robots autonomes réactifs** ..... 59

H. Hajri, H. Zaatiti, G. Hebrail, P. Akinin (CAp).

**Apprentissage automatique sur des données de type graphe utilisant le plongement de Poincaré et les algorithmes stochastiques riemanniens** ..... 69

G. Kassel (IC).

**Trois conceptions du processus : les raisons d'un choix** ..... 77

T. Lorieul, A. Joly (CAp).

**Vers un désenchevêtrement de l'ambiguïté de la tâche et de l'incertitude du modèle pour la classification avec option de rejet à l'aide de réseaux neuronaux** ..... 87

J. Mattioli, S. Lamoudi, P.-O. Robic (APIA).

**La gestion d'actifs augmentée par l'intelligence artificielle** ..... 97

S. Montariol, A. Allauzen (TALN-RECITAL).

**Apprentissage de plongements de mots dynamiques avec régularisation de la dérive** ..... 105

S. Pageaud, V. Deslandres, V. Lehoux, S. Hassas (APIA).

**Application du *clustered deep Q-network* aux politiques tarifaires** ..... 113

R. Pierrard, J.-P. Poli, C. Hudelot (RJCIA).

**Apprentissage de relations floues pour l'annotation sémantique expliquée avec peu de données** ..... 121

W. Ragheb, J. Azé, S. Bringay, M. Servajean (IC).

**Pourquoi dois-je croire ta prédiction ? Comment expliquer les résultats d'une classification automatique de sentiments à partir de textes** ..... 129

M. Scholivet (TALN-RECITAL).

<b>Méthodes de représentation de la langue pour l'analyse syntaxique multilingue</b> .....	137
N. Seydoux, M. Lefrançois, L. Médini (IC).	
<b>Positionnement sur le Web Sémantique des Objets</b> .....	145
B. Taillé, V. Guigue, P. Gallinari (CAp).	
<b>Une étude empirique de la capacité de généralisation des plongements de mots contextuels en extraction d'entités</b> .....	157
T. Van de Cruys (TALN-RECITAL).	
<b>La génération automatique de poésie en français</b> .....	163
L. Vial, B. Lecouteux, D. Schwab (TALN-RECITAL).	
<b>Compression de vocabulaire de sens grâce aux relations sémantiques pour la désambiguïsation lexicale</b> .....	171
W. Younes, F. Adreit, S. Trouilhet, J.-P. Arcangeli (CAp).	
<b>Apprentissage en ligne par renforcement centré utilisateur pour la composition émergente d'applications ambiantes</b> .....	179

# Toulouse, au cœur de l'Intelligence Artificielle de demain

**Sylvie Doutre, Dominique Longin, Frédéric Maris,**

Maître de Conférence U. Toulouse 1, Chargé de Recherche CNRS, Maître de Conférence U. Toulouse 3  
Co-Présidents du Comité d'Organisation de la Plate-Forme IA 2019

**Sandra Bringay, Yves Demazeau, Catherine Faron-Zucker**

Professeur U. Montpellier 3, Directeur de Recherche au CNRS, Maître de Conférence U. Nice Sophia Antipolis  
Co-Présidents du Comité de Programmation de la Plate-Forme IA 2019

L'AFIA et l'IRIT organisent à Toulouse du 01 au 05 juillet 2019 la douzième 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), Caen (2017) et Nancy (2018).

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 en IA pour aborder des problématiques communes.

Pour son édition 2019, la plate-forme IA héberge les conférences suivantes, ainsi que leurs ateliers associés :

<b>APIA</b>	5 <sup>ème</sup> Conférence Nationale sur les Applications Pratiques de l'IA
<b>CAP</b>	21 <sup>ème</sup> Conférence sur l'Apprentissage Automatique
<b>CNIA</b>	22 <sup>ème</sup> Conférence Nationale d'Intelligence Artificielle
<b>IC</b>	30 <sup>èmes</sup> Journées Francophones d'Ingénierie des Connaissances
<b>JFPDA</b>	14 <sup>ème</sup> Journées Francophone Planification, Décision et Apprentissage
<b>JFSMA</b>	27 <sup>èmes</sup> Journées Francophones sur les Système Multi-Agents
<b>JIAF</b>	13 <sup>èmes</sup> Journées d'Intelligence Artificielle Fondamentale
<b>RJCIA</b>	17 <sup>ème</sup> Rencontres des Jeunes Chercheurs en Intelligence Artificielle
<b>TALN</b>	26 <sup>ème</sup> Conférence sur le Traitement Automatique des Langues Naturelles

Mais aussi les autres événements :

<b>France @ Intern.</b>	Présentation par leurs auteurs français d'articles acceptés à AAAI, IJCAI, PRICAI 2019
<b>DriveToGaether</b>	Compétition en intelligence artificielle et robotique mobile

Dix collègues de renommée ont également été invités pour des conférences en début de chaque demi-journée

<b>Isabelle Bloch</b>	Télécom Paris Tech	France
<b>Julien Bourgeois</b>	U. Franche-Comté	France
<b>Walter Daelemans</b>	U. Antwerp	Belgique
<b>Takayuki Ito</b>	Nagoya Institute of Technology	Japon
<b>Sylvie Thiebaut</b>	Australian National U.	Australie
<b>Jochen Triesch</b>	U. Frankfurt	Allemagne
<b>Ruben Verborgh</b>	U. Gent	Belgique
<b>Anaëlle Wilczynski</b>	Prix de Thèse IA 2018	France
<b>Michael Wooldridge</b>	U. Oxford	Royaume-Uni
<b>Valentina Zantedeschi</b>	Prix de Thèse IA 2018	France

L'ensemble des informations sur ces événements est disponible à partir du site de la Plate-Forme IA 2019 (<https://www.irit.fr/pfia2019/>) dont le contenu sera maintenu jusqu'à l'été 2023.

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.



# Comité de programmation

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

## Présidents

- Yves DEMAZEAU, LIG, CNRS

## Membres

- Sandra BRINGAY, LIRMM, Université de Montpellier
- Catherine FARON-ZUCKER, I3S, Université de Nice Sophia-Antipolis

# Comités de programme

Conférence Nationale sur les Applications Pratiques de l'Intelligence Artificielle (APIA)

## Présidents

- Amal EL FALLAH-SEGHROUCHNI, Sorbonne University
- Juliette MATTIOLI, Thales

## Membres

- Carole ADAM, LIG CNRS UMR 5217 - Université Grenoble-Alpes
- Florence AMARDEILH, Ticket for Change
- Ghislain ATEMEZING, Mondeca
- Jérôme AZÉ, LIRMM-UM-CNRS
- Alain BERGER, Ardans
- Moez BOUCHOUICHA, Université du Sud Toulon-Var
- Bertrand BRAUNSCHWEIG, INRIA
- Stéphane BRUNESSAUX, Airbus
- Nathalie CHAIGNAUD, LITIS - INSA Rouen Normandie
- Caroline CHOPINAUD, craft ai
- Laurent COSSERAT, Renault
- Jean-Marc DAVID, Renault
- Philippe DAVID, SNCF
- Etienne de SEVIN, SANPSY - University of Bordeaux
- Yves DEMAZEAU, CNRS - LIG
- Sylvie DESPRES, LIMICS - Université Paris 13
- Sébastien DESTERCCKE, CNRS, UMR Heudiasyc
- Hamza DIDARALY, IA pour Tous
- Christophe GUETTIER - SAFRAN ELECTRONICS & DEFENSE
- Béatrice FUCHS, LIRIS, IAE - Université Lyon 3
- Céline HUDELOT, Ecole Centrale Paris
- Christophe LABREUCHE, Thales
- Arnaud LALLOUET, Huawei
- Christine LARGOUËT, Irisa /Agrocampus Ouest
- Dominique LENNE, Heudiasyc - Université de Technologie de Compiègne
- Domitile LOURDEAUX, Heudiasyc - Université de Technologie de Compiègne
- Sylvain MAHE, EDF R&D
- Philippe MATHIEU, University of Lille 1
- Nada MATTA, University of Technology of Troyes
- Christophe MENICHETTI, IBM
- Eunika MERCIER-LAURENT, CRESTIC- Université Reims Champagne Ardennes
- Christophe MEYER, Thales
- Philippe MORIGNOT, SAFRAN
- Selmin NURCAN, Université Paris 1 Panthéon - Sorbonne
- Jean-Marc OGIER, L3i - University of La Rochelle
- Jean ROHMER, De Vinci Research Center (DVRC)

- Frédérique SEGOND, Bertin
- 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 - Sorbonne University

## Conférence sur l'Apprentissage automatique (CAp)

### Présidents

- Frédéric Garcia (INRA, MIAT, Toulouse)

### Membres

- Massih-Reza Amini (Université Grenoble Alpes, LIG, Grenoble)
- Dominique Bontemps (Université Paul Sabatier, IMT, Toulouse)
- Stéphane Canu (Insa de Rouen)
- Marie-Josée Cros (INRA, MIAT, Toulouse)
- Nicolas Dobigeon (INP-ENSEEIH, IRIT, Toulouse)
- Emmanuel Faure (CNRS, IRIT, Toulouse)
- Cédric Févotte (CNRS, IRIT, Toulouse)
- Sébastien Gerchinovitz (Université Paul Sabatier, IMT, Toulouse)
- Yves Grandvalet (CNRS, Heudiasyc, Compiègne)
- Timothée Masquelier (CNRS, CerCo, Toulouse)
- Thomas Oberlin (INP-ENSEEIH, IRIT, Toulouse)
- Nathalie Peyrard (INRA, MIAT, Toulouse)
- Emmanuel Rachelson (ISAE, Toulouse)
- Régis Sabbadin (INRA, MIAT, Toulouse)
- Michèle Sebag (CNRS, LRI, Gif sur Yvette)
- Rufin Van Rullen (CNRS, CerCo, Toulouse)

### Relecteurs

- Thierry Artières (Université d'Aix-Marseille, LIF)
- Stéphane Ayache (Université d'Aix-Marseille, LIF)
- Nicolas Baskiotis (Sorbonne Université, LIP6)
- Aurélien Bellet (Inria)
- Younès Bennani (Université Paris13, LIPN)
- Marianne Clausel (Lorraine University)
- Stéphane Cléménçon (Telecom ParisTech)
- Antoine Cornuéjols (AgroParisTech)
- Nicolas Courty (Université de Bretagne Sud, IRISA)
- Bruno Cremilleux (Université de Caen Normandie)
- Colin De La Higuera (University of Nantes)
- Gaël Dias (Normandie University)
- Remi Eyraud (Aix-Marseille University)
- Aurélien Garivier (Institut de Mathématiques de Toulouse)
- Gilles Gasso (INSA de Rouen, LITIS)
- Romaric Gaudel (Université de Rennes, Ensai, CREST)
- Pascal Germain (Inria)
- Remi Gilleron (Université de Lille, Inria)
- Alexandre Gramfort (Telecom ParisTech, CNRS LTCI)
- Yann Guermeur (CNRS)
- Vincent Guigue (Sorbonne Université, LIP6)
- Amaury Habrard (University of Saint-Etienne)
- Alexis Joly (Inria)
- Mikaela Keller (Université Lille 3, Inria)
- Christine Keribin (Université Paris Sud, Laboratoire de Mathématique)
- Engelbert Mephu Nguifo (University Clermont Auvergne, CNRS, LIMOS)
- Mohamed Nadif (University Paris Descartes)
- Liva Ralaivola (Université d'Aix-Marseille, LIF)



- Ievgen Redko (University Jean Monnet of Saint-Etienne, LHC)
- Marc Tommasi (University of Lille)

## Conférence Nationale en Intelligence Artificielle (CNIA)

### Présidents

- Jérôme Lang

### Membres

- Nathalie Aussenac
- Meghyn Bienvenu
- Tristan Cazenave
- Séverine Dubuisson
- Hélène Fargier
- Béatrice Fuchs
- Andreas Herzig
- Frédéric Koriche
- Pierre Marquis
- Nicolas Maudet
- Marie-Laure Mugnier
- Philippe Muller
- Christine Solnon
- Serena Villata
- Bruno Zanuttini

## Journées francophones d'Ingénierie des Connaissances (IC)

### Présidents

- Nathalie Hernandez IRIT Université Toulouse 2 Jean-Jaurès

### Membres

- Marie-Hélène Abel UTC
- Xavier Aimé Cogsonomy / LIMICS UMRS 1142 Inserm
- Yamine Ait Ameer IRIT/INPT-ENSEEIH
- Bruno Bachimont Sorbonne Université
- Jean-Paul Barthès UTC
- Aurélien Bénel Université de technologie de Troyes
- Nacéra Bennacer Seghouani LRI CentraleSupélec
- Bertrand Braunschweig INRIA
- Nathalie Bricon-Souf IRIT Université Paul Sabatier Toulouse
- Sandra Bringay LIRMM
- Patrice Buche INRA
- Davide Buscaldi LIPN, Université Paris 13, Sorbonne Paris Cité
- Sylvie Calabretto LIRIS
- Gaoussou Camara Université Alioune Diop de Bambey - Sénégal
- Pierre-Antoine Champin LIRIS, Université Claude Bernard Lyon1
- Jean Charlet AP-HP & INSERM UMRS 1142
- Olivier Corby INRIA
- Mathieu D'Aquin Insight Centre for Data Analytics, National University of Ireland Galway
- Sylvie Despres Laboratoire d'Informatique Médicale et de BIOinformatique (LIM&BIO)
- Jean-Pierre Evain EBU
- Gilles Falquet University of Geneva
- Catherine Faron Zucker Université Nice Sophia Antipolis
- Cécile Favre ERIC - Université Lyon 2
- Béatrice Fuchs LIRIS, IAE - université Lyon 3
- Frédéric Fürst MIS - Université de Picardie - Jules Verne
- Alban Gaignard CNRS
- Jean-Gabriel Ganascia Pierre and Marie Curie University - LIP6

- Serge Garlatti IMT Atlantique
- Alain Giboin INRIA
- Nathalie Guin LIRIS - Université de Lyon
- Mounira Harzallah LS2N
- Liliana Ibanescu AgroParisTech, INRA,
- Sébastien Iksal LIUM - Le Mans Université, France
- Antoine Isaac Europeana & VU University Amsterdam
- Clement Jonquet University of Montpellier - LIRMM
- Mouna Kamel IRIT - Université Paul Sabatier - Toulouse
- Gilles Kassel University of Picardie Jules Verne
- Pascale Kuntz Laboratoire d'Informatique de Nantes Atlantique
- Michel Leclère LIRMM (CNRS - UM2)
- Maxime Lefrançois MINES Saint-Etienne
- Alain Leger France Telecom R&D - Research - Orange Labs
- Dominique Lenne Heudiasyc, Université de Technologie de Compiègne
- Cédric Lopez Emvista
- Pascal Molli University of Nantes - LS2N
- Alexandre Monnin Origens Medialab
- Isabelle Mougenot Université Montpellier EspaceDev
- Fleur Mougin ERIAS, INSERM U1219 - Université de Bordeaux
- Amedeo Napoli LORIA Nancy (CNRS - Inria - Université de Lorraine) France
- Jérôme Nobécourt LIMICS
- Nathalie Pernelle LRI-Universit Paris SUD
- Camille Pradel Synapse Développement
- Yannick Prié LINA - University of Nantes
- Cédric Pruski Luxembourg Institute of Science and Technology
- Sylvie Ranwez LGI2P / Ecole des mines d'Alès
- Chantal Reynaud LRI, Univ. Paris-Sud, CNRS, Université Paris-Saclay
- Catherine Roussey Irstea Clermont-Ferrand Center
- Fatiha Saïs LRI ( Paris Sud University &CNRS8623), Paris Saclay University
- Pascal Salembier Université de Troyes
- Karim Sehaba LIRIS - Université Lumière Lyon 2
- Konstantin Todorov LIRMM / University of Montpellier
- Raphaël Troncy EURECOM
- Haifa Zargayouna University Paris 13

## **Journées Francophones sur la Planification, la Décision et l'Apprentissage pour la conduite de systèmes (JFPDA)**

### **Présidents**

- Emmanuel Rachelson (ISAE-SUPAERO)
- Caroline Chanel (ISAE-SUPAERO)

### **Membres**

- Olivier Buffet (LORIA)
- Nicolas Drougard (ISAE-SUPAERO)
- Alain Dutech (LORIA)
- Humbert Fiorino (Université de Grenoble - LIG)
- Raphaël Fonteneau (Université de Liège)
- Frédérick Garcia (INRA)
- Matthieu Geist (Google AI)
- Malik Ghallab (LAAS-CNRS)
- Alain Haït (ISAE-SUPAERO)
- Guillaume Infantes (Jolibrain)
- Félix Ingrand (LAAS-CNRS)
- Erwan Lecarpentier (ISAE-SUPAERO)
- Charles Lesire (ONERA)
- Frédéric Maris (Université Toulouse 3 - IRIT)

- Laetitia Matignon (LIRIS CNRS)
- Alexandre Niveau (Université de Caen - GREYC)
- Damien Pellier (Université de Grenoble - LIG)
- Nicolas Perrin (CNRS - ISIR)
- Mathieu Picard (IRT Saint Exupéry - Airbus DS)
- Cédric Pralet (ONERA)
- Philippe Preux (Université de Lille - INRIA)
- Serge Rainjonneau (IRT Saint Exupéry - Thalès Alenia Space)
- Régis Sabbadin (INRA)
- Nicolas Schneider (Airbus CRT)
- Olivier Sigaud (Université Paris Sorbonne - ISIR)
- Jonathan Sprauel (Thalès Alenia Space)
- Florent Teichtel (Airbus CRT)
- Vincent Thomas (LORIA)
- Paul Weng (UM-SJTU Joint Institute)
- Bruno Zanuttini (Université de Caen - GREYC)

## Journées Francophones sur les Systèmes Multi-Agents (JFSMA)

### Présidents

- Olivier Simonin, INSA Lyon, CITI lab. & Inria

### Membres

- Flavien Balbo ISCOD / Henri Fayol Institute, MINES Saint-Etienne
- Valérie Camps University of Toulouse - IRIT, France
- Olivier Boissier Mines Saint-Etienne, Institut Henri Fayol, Laboratoire Hubert Curien UMR CNRS 5516
- Zahia Guessoum LIP6, Université de Paris 6 and CReSTIC, Université de Reims Champagne Ardenne
- Guillaume Hutzler Evry University,
- Aurélie Beynier University of Pierre and Marie Curie (Paris 6), LIP6
- Maxime Morge Université de Lille
- Yves Demazeau CNRS - LIG
- Philippe Mathieu University of Lille 1
- Stéphane Galland UBFC - UTBM
- Bruno Mermet GREYC-CNRS - Université du Havre
- Gildas Morvan LIG2A, Université d'Artois
- Nicolas Marilleau UMI UMMISCO - Institut de recherche pour le développement (IRD)
- Rene Mandiau LAMIH, Université de Valenciennes
- Julien Saunier LITIS, INSA-Rouen
- Michel Ocello Université Grenoble Alpes - LCIS
- Frederic Migeon IRIT
- Emmanuelle Grislin LAMIH - Université de Valenciennes
- Nicolas Sabouret LIMSI-CNRS
- Lilia Rejeb SMART LAB, ISG Tunis, Université de Tunis
- Vincent Chevrier Université de Lorraine, LORIA
- Ely Kaddoum IRIT Université Toulouse III
- François Charpillet LORIA, Inria, Nancy
- Suzanne Pinson Université Paris-Dauphine
- Grégory Bonnet Université de Caen Normandie
- Jean-Pierre Muller CIRAD
- Jean-Paul Jamont LCIS, Université de Grenoble
- Mahdi Zargayouna IFSTTAR Marne-la-Vallée
- Denis Payet Université de la Réunion
- Christophe Lang FEMTO-ST Besançon
- Emmanuel Adam LAMIH, Université de Valenciennes

## Journées d'Intelligence Artificielle Fondamentale (JIAF)

### Présidents

- Tiago de Lima (CRIL, Université d'Artois)
- Sylvie Doutre (IRIT, Université Toulouse 1 Capitole)

### Membres

- Francesco Belardinelli (IBISC, Université d'Évry)
- Elise Bonzon (LIPADE, Université Paris Descartes)
- Tristan Cazenave (LAMSADE, Université Paris Dauphine)
- Nadia Creignou (LIS, Aix-Marseille Université)
- Jérôme Euzenat (LIG, INRIA)
- George Katsirelos (MIAT, INRA)
- Sébastien Konieczny (CRIL, CNRS)
- Jérôme Lang (LAMSADE, Université Paris Dauphine)
- Jean Lieber (LORIA, INRIA)
- Pierre Marquis (CRIL, Université d'Artois)
- Amedeo Napoli (LORIA, CNRS)
- Odile Papini (LIS, Aix-Marseille Université)
- Laurent Perrussel (IRIT, Université Toulouse 1 Capitole)
- Sophie Pinchinat (IRISA, INRIA)
- Stéphanie Roussel (ONERA)
- Serena Villata (I3S, CNRS)
- Christel Vrain (LIFO, Université d'Orléans)
- Bruno Zanuttini (GREYC, UNICAEN)

## Rencontres des Jeunes Chercheurs en Intelligence Artificielle (RJCIA)

### Présidents

- Maxime Lefrançois (MINES Saint-Étienne)

### Membres

- Céline Alec (Université Caen-Normandie)
- Jonathan Bonnet (Continental)
- Mathieu Chollet (University of Glasgow)
- Nicolas Cointe (TU Delft)
- Antoine Cully (Imperial College London)
- Jérôme Delobelle (INRIA Sophia-Antipolis - Méditerranée)
- Dennis Diefenbach (University Jean Monet)
- Géraud Fokou Pelap (INRIA / EDUCLEVER)
- Hugo Gilbert (Gran Sasso Science Institute)
- Pierre-François Gimenez (LAAS)
- Maxime Gueriau (Trinity College Dublin)
- Robin Lamarche-Perrin (CNRS / ISC-PIF / LIP6)
- Marc Law (University of Toronto)
- Amélie Levray (University of Edinburgh)
- Juan Antonio Lossio-Ventura (University of Florida)
- Jean-Guy Mailly (LIPADE, Université Paris Descartes, France)
- Wafa Mefteh (Laboratoire MIRACL, Université de Gabès, Tunisie)
- Amro Najjar (Umeå University, Sweden)
- Abdelkader Ouali (INRIA Nancy - Grand Est)
- Michaël Perrot (Max Planck Institute for Intelligent Systems)
- Eric Piette (DKE)
- Abdallah Saffidine (The University of New South Wales)
- Michael Sioutis (Aalto University)
- Nicolas Szczepanski (CRIL)
- Michaël Thomazo (Inria, DIENS, ENS, CNRS, PSL University)
- Thibaut Vallée (GREYC)
- Nicolas Verstaevael (SMART Infrastructure Facility, University of Wollongong)

# Conférence sur le Traitement Automatique des Langues Naturelles (TALN-RECITAL)

## Présidents de TALN

- Emmanuel Morin (LS2N, Université de Nantes)
- Sophie Rosset (LIMSI, CNRS, Université Paris-Saclay)
- Pierre Zweigenbaum (LIMSI, CNRS, Université Paris-Saclay)

## Membres du CP de TALN

- Delphine Bernard (LiLPa, Université de Strasbourg)
- Chloé Braud (LORIA, CNRS)
- Nathalie Camelin (LIUM, Université du Maine)
- Peggy Cellier (IRISA, INSA Rennes)
- Benoît Crabbé (LLF, Université Paris Diderot)
- Iris Eshkol-Taravella (MoDyCo, Université Paris Nanterre)
- Cécile Fabre (ERSS, Université Toulouse 2)
- Núria Gala (LPL, Aix Marseille Université)
- Thierry Hamon (LIMSI, Université Paris Nord)
- Philippe Langlais (RALI/DIRO, Université de Montréal)
- Gwénoél Lecorvé (IRISA, Université de Rennes 1)
- Aurélie Névéol (LIMSI, CNRS, Université Paris-Saclay)
- Damien Nouvel (ERTIM, INaLCO)
- Didier Schwab (LIG, Université Grenoble Alpes)
- Xavier Tannier (LIMICS, Université Pierre et Marie Curie)

## Présidentes de RECITAL

- Anne-Laure Ligozat (LIMSI, CNRS, Université Paris-Saclay)
- Sahar Ghannay (LIMSI, CNRS, Université Paris-Saclay)

## Membres du CP de RECITAL

- Jean-Yves Antoine, LI, Université François Rabelais de Tours
- Ismail Badache, ESPE / LIS, Aix-Marseille Université
- Amira Barhoumi, LIUM, Université du Maine - MIRACL Sfax
- Rachel Bawden, University of Edinburgh
- Aurélien Bossard, LIASD, Université Paris 8
- Chloé Braud, LORIA - CNRS
- Nathalie Camelin, LIUM, Université du Maine
- Rémi Cardon, STL, Lille
- Peggy Cellier, IRISA/INSA Rennes
- Antoine Doucet, L3i, University of La Rochelle
- Maha Elbayad, LIG/ Inria
- Arnaud Ferré, LIMSI-CNRS/MaIAGE-INRA, Université Paris-Saclay
- Amel Fraisse, Gériico, Lille
- Thomas François, CENTAL, Université catholique de Louvain
- Nicolas Hernandez, LS2N, Université de Nantes
- Yann Mathet, Greyc, University of Caen
- Alice Millour, STIH, Université Paris-Sorbonne
- Anne-Lyse Minard, LLL, Orléans
- Jose Moreno, IRIT, UPS
- Tsanta Randriatsitohaina, LIMSI, Orsay
- Loïc Vial, LIG, Université Grenoble Alpes

# Comité d'organisation

## Institut de Recherche en Informatique de Toulouse (IRIT)

### Présidents

- Sylvie DOUTRE, IRIT, Université Toulouse 1 Capitole
- Dominique LONGIN, IRIT, CNRS
- Frédéric MARIS, IRIT, Université Toulouse 3 Paul Sabatier

### Gestion administrative

- Marlène GIAMPORCARO, SAIC-INPACT, INPT
- Telma NARDI, SAIC-INPACT, INPT

### Membres

- Julien ALIGON, IRIT, Université Toulouse 1 Capitole
- Jean-Christophe BUISSON, IRIT, INPT
- Valérie CAMPS, IRIT, Université Toulouse 3 Paul Sabatier
- Yannick CHEVALIER, IRIT, Université Toulouse 3 Paul Sabatier
- Stéphanie COMBETTES, IRIT, Université Toulouse 3 Paul Sabatier
- Romain GUILLAUME, IRIT, Université Toulouse 2 Jean Jaurès
- Nathalie HERNANDEZ, IRIT, Université Toulouse 2 Jean Jaurès
- Marie-José HUGUET, LAAS, INSA
- Jean-Michel LOUBES, IMT, Université Toulouse 3 Paul Sabatier
- Frédéric MIGEON, IRIT, Université Toulouse 3 Paul Sabatier
- Philippe MULLER, IRIT, Université Toulouse 3 Paul Sabatier
- Arianna NOVARO, IRIT, Université Toulouse 3 Paul Sabatier
- Laurent PERRUSSEL, IRIT, Université Toulouse 1 Capitole
- Emmanuel RACHELSON, ISAE-SUPAERO
- Stéphanie ROUSSEL, ONERA
- Patrick TAILLANDIER, MIAT, INRA
- Ludovic TANGUY, CLLE-ERSS, Université Toulouse 2 Jean Jaurès
- Elise VAREILLES, IMT Mines Albi et ETS Montréal, Canada
- Pascale ZARATÉ, IRIT, Université Toulouse 1 Capitole

## Sponsors et soutiens institutionnels

### SPONSORS PLATINE



### SPONSORS OR



### SPONSORS ARGENT



### SPONSORS BRONZE



### SUPPORTS INSTITUTIONNELS







# Étude de la motivation intrinsèque en apprentissage par renforcement

A. Aubret<sup>1</sup>L. Matignon<sup>1</sup>S. Hassas<sup>1</sup><sup>1</sup> Univ Lyon, Université Lyon 1, CNRS, LIRIS, F-69622, Villeurbanne, France

arthur.aubret@liris.cnrs.fr

## Résumé

*Malgré les nombreux travaux existants en apprentissage par renforcement (AR) et les récents succès obtenus notamment en le combinant avec l'apprentissage profond, l'AR fait encore aujourd'hui face à de nombreux défis. Certains d'entre eux, comme la problématique de l'abstraction temporelle des actions ou la difficulté de concevoir une fonction de récompense sans connaissances expertes, peuvent être adressées par l'utilisation de récompenses intrinsèques. Dans cet article, nous proposons une étude du rôle de la motivation intrinsèque en AR et de ses différents usages, en détaillant les intérêts et les limites des approches existantes. Notre analyse suggère que la compression d'information est centrale à la plupart des travaux utilisant la motivation intrinsèque en AR. Celle-ci, combinée aux algorithmes d'AR profond, permet d'apprendre des comportements plus complexes et plus généralisables que ce que permet l'AR traditionnel.*

## Mots Clef

Apprentissage par renforcement, motivation intrinsèque, curiosité, acquisition de connaissances, options, génération d'objectifs, méta-récompense.

## Abstract

*Despite many existing works in reinforcement learning (RL) and the recent successes obtained by combining it with deep learning, RL is facing many challenges. Some of them, like the ability to abstract the action or the difficulty to conceive a reward function without expert knowledge, can be addressed by the use of intrinsic motivation. In this article, we provide a survey on the role of intrinsic motivation in RL and its different usages by detailing interests and limits of existing approaches. Our analysis suggests that compression of information is central to most of the work using intrinsic motivation in RL. The combination of deep RL and intrinsic motivation enables to learn more complicated and more generalisable behaviours than what enables standard RL.*

## Keywords

Reinforcement learning, intrinsic motivation, curiosity, knowledge acquisition, options, generation of objectives, meta-reward.

## 1 Introduction

En apprentissage par renforcement (AR), un agent apprend par essais-erreurs à maximiser l'espérance des récompenses reçues suite aux actions effectuées dans son environnement [57].

Traditionnellement, pour apprendre une tâche, un agent maximise une récompense définie selon la tâche à accomplir : cela peut être un score lorsque l'agent doit apprendre à gagner à un jeu ou une fonction de distance lorsque l'agent apprend à atteindre un objectif. Nous parlons alors de récompense extrinsèque (ou *feedback*) car la fonction de récompense est fournie de manière experte spécifiquement pour la tâche. Avec une récompense extrinsèque, plusieurs résultats spectaculaires ont été obtenus sur les jeux Atari [7] avec le Deep Q-network (DQN) [34] ou sur le jeu du go avec AlphaGo Zero [54]. Cependant, ces approches se révèlent le plus souvent infructueuses lorsque les récompenses sont trop éparses dans l'environnement, et l'agent est alors incapable d'apprendre le comportement recherché pour la tâche [17]. D'autre part, les comportements appris par l'agent sont difficilement réutilisables, aussi bien au sein d'une même tâche que pour plusieurs tâches [17] : il est difficile pour un agent de généraliser ses compétences de manière à prendre des décisions abstraites dans l'environnement. Par exemple, une décision abstraite (ou de haut-niveau) pourrait être *aller jusqu'à la porte* en utilisant des actions primitives (ou de bas-niveau) consistant à se déplacer dans les quatre directions cardinales ; ou encore de *se déplacer en avant* en contrôlant les différentes articulations d'un robot humanoïde comme dans le simulateur de robotique MuJoCo [60]. Ces actions abstraites sont souvent appelées *options* [58].

Contrairement à l'AR, l'apprentissage développemental [45] s'inspire de la tendance qu'ont les bébés, ou plus généralement tout organisme, à explorer spontanément leur environnement [19] : c'est ce que nous appelons une motivation intrinsèque, laquelle peut être issue d'une récompense intrinsèque.

Dans cet article nous proposons une étude de l'usage de la motivation intrinsèque dans le framework de l'apprentissage profond par renforcement, plus particulièrement nous souhaitons répondre aux questions suivantes :

- Comment caractériser la motivation intrinsèque ?

- Comment la motivation intrinsèque peut-elle s'intégrer au framework d'AR?
- Quel rôle joue-t-elle vis-à-vis des défis énoncés ci-dessus?
- Quel lien existe-t-il entre la motivation intrinsèque et la théorie de l'information?
- Quelles sont les limites actuelles de l'utilisation de récompenses intrinsèques en AR?

Nous ne prétendons pas faire une étude exhaustive mais plutôt donner les axes de recherches courants et des perspectives à exploiter.

Dans un premier temps, nous définirons les éléments clés de l'article qui sont les processus de décision markovien et la motivation intrinsèque (Partie 2). Ensuite nous mettrons en avant les problématiques de l'AR (Partie 3). Nous aurons alors les éléments pour détailler les trois différents types de travaux intégrant l'AR et la motivation intrinsèque (Partie 4). Puis, nous prendrons du recul et analyserons les points communs entre les différents travaux (Partie 5). Pour terminer, nous mettrons en avant les défis actuels des modèles intégrant la motivation intrinsèque à l'AR (Partie 6). Suite à des limitations d'espace, une version complète de l'article peut être trouver sur HAL<sup>1</sup>

## 2 Définitions

### 2.1 Processus de décision markovien

L'objectif d'un processus de décision markovien (MDP) est de maximiser l'espérance de récompense reçue via une suite d'interaction. Il est défini par  $S$  l'ensemble des états possibles du système,  $A$  l'ensemble des actions possibles, la fonction de transition  $P : S \times A \times S \rightarrow \mathbb{R}$ , la fonction de récompense  $R : S \times A \rightarrow \mathbb{R}$ ;  $\gamma \in [0, 1]$  le facteur d'atténuation et  $\rho_0 : S \rightarrow \mathbb{R}$  la distribution initiale de l'état du système.

L'agent démarre dans un état  $s_0$  donné par  $\rho_0$  puis effectue une action  $a_0$ . Il attend ensuite la réponse de l'environnement qui lui renverra un nouvel état  $s'$ , donné par la fonction de transition  $P$ , et une récompense  $r$  évaluée par la fonction de récompense  $R$ . L'agent peut répéter la boucle d'interactions jusqu'à la fin d'un épisode. Un algorithme de renforcement permet d'associer des actions  $a$  aux états  $s$  via une politique  $\pi$ . L'objectif de l'agent est alors de trouver la politique  $\pi^*$  maximisant la récompense sur le long terme :

$$\pi^* = \arg \max_{\pi} \mathbb{E} \left[ \sum_{t=0}^{\infty} \gamma^t R(s_t, \pi(s_t)) \right]. \quad (1)$$

De manière à trouver l'action maximisant la récompense sur le long-terme dans un état  $s$ , il est courant de maximiser l'espérance de gain atténuée depuis cet état, noté  $V(s)$ , ou l'espérance de gain atténuée depuis le couple (état,action)  $Q(s, a)$  (c.f équation 2). Cela permet d'adresser le *credit assignment problem* en mesurant le rôle du

couple (état,action) dans l'obtention de la récompense cumulée [57].

$$Q_{\pi}(s, a) = E_{a_t \sim \pi(s_t)} \left( \sum_{t=0}^{\infty} \gamma^t R(s_t, a_t) \Big|_{s_0=s, a_0=a} \right). \quad (2)$$

Pour calculer ces valeurs, il est possible d'utiliser l'équation de Bellman [57] :

$$Q(s_t, a_t) = R(s_t, a_t) + \gamma Q(P(s_t, a_t), a_{t+1}) \quad (3)$$

$Q$  et/ou  $\pi$  sont souvent approximés via des réseaux de neurones lorsque l'espace d'état est continu ou de très grande taille [33].

### 2.2 Motivation intrinsèque

L'idée de la motivation intrinsèque est d'inciter un agent à avoir un certain type de comportement sans que l'environnement intervienne directement. Plus simplement, il s'agit de faire quelque chose pour son inhérente satisfaction plutôt que pour une récompense assignée par l'environnement [48]. Ce type de motivation renvoie à l'apprentissage développemental, lequel s'inspire de la tendance des bébés à explorer leur environnement [19]. Historiquement, la motivation intrinsèque est issue de la tendance des organismes à jouer ou explorer leur environnement sans qu'aucune récompense externe ne leur soit attribuée [63].

Plus rigoureusement, Oudeyer [41] explique qu'*une situation est intrinsèquement motivée pour une entité autonome si son intérêt dépend principalement de la collecte ou comparaison d'information depuis différents stimulus indépendamment de leur sémantique*. Le point principal est que l'agent ne doit avoir aucun *a priori* sur la sémantique des observations qu'il reçoit. On remarque que le terme de comparaison d'information renvoie directement à la théorie de l'information ci-dessus. Berlyne [8] et Oudeyer [41] proposent plusieurs types de motivation pouvant être caractérisées comme intrinsèques :

- la nouveauté et la complexité comme étant quelque chose que l'agent ne connaît pas;
- la surprise et l'incongruité peuvent attirer l'agent car cela remet en question ses précédents apprentissages;
- l'ambiguïté et l'indistinction renvoient à l'incompréhension de l'agent vis-à-vis des observations.

Typiquement, un étudiant qui fait ses devoirs de mathématique car il les trouve intéressants est intrinsèquement motivé tandis que son camarade qui les fait pour avoir une bonne note est extrinsèquement motivé. De même, jouer avec des jouets pour s'amuser est une motivation intrinsèque tandis que jouer à un jeu télévisé pour gagner de l'argent est une motivation extrinsèque. La notion d'**intrinsèque/extrinsèque** renvoie au *pourquoi de l'action*, à ne pas confondre avec l'internalité/externalité qui renvoie à la localisation de la récompense [41].

1. <https://hal.archives-ouvertes.fr/hal-02272091>

En pratique, la récompense  $r$  est souvent calculée comme une somme pondérée de récompense intrinsèque  $r_{int}$  et extrinsèque  $r_{ext}$  :  $r = \alpha r_{int} + \beta r_{ext}$  [11]. Nous pouvons aussi parler de la récompense intrinsèque comme un bonus intrinsèque. Notons que la fonction de récompense évolue au cours du temps, cela ne respecte pas la propriété d'invariance d'un MDP.

### 3 Problématiques de l'AR

Dans cette partie, nous détaillons tout d'abord les deux principaux défis que cherchent à résoudre les travaux combinant récompenses intrinsèques et AR.

#### 3.1 Les récompenses éparées

Les algorithmes classiques d'AR fonctionnent dans des environnements où les récompenses sont **denses**, *i.e.* que l'agent reçoit une récompense après presque chaque action réalisée. Dans ce type d'environnements, des politiques d'explorations naïves telles que l'exploration  $\epsilon$ -greedy [57] ou l'ajout de bruit gaussien [30] sont efficaces. Des méthodes plus élaborées peuvent aussi être utilisées comme l'exploration Boltzmann [12] ou une exploration dans l'espace des paramètres [46].

Dans les environnements à récompenses **éparées**, l'agent reçoit un signal de récompense seulement après avoir exécuté une longue séquence spécifique d'actions. Le jeu *Montezuma's revenge* [7] est un environnement de référence pour illustrer le cas des récompenses éparées. Dans ce jeu, un agent doit se déplacer de salles en salles en y récupérant des objets (clés pour ouvrir les portes, torches, ...). L'agent reçoit une récompense uniquement lorsqu'il trouve des objets ou la sortie d'une salle. Plusieurs actions spécifiques doivent donc être réalisées avant l'obtention d'une récompense. Ce type d'environnements à récompenses éparées est pratiquement impossible à résoudre avec les méthodes d'exploration mentionnées ci-dessus, l'agent parvenant difficilement à apprendre une bonne politique vis-à-vis de la tâche [34].

Plutôt que de travailler sur la politique d'exploration, il est courant de construire une récompense intermédiaire dense qui s'ajoute à celle de la tâche pour faciliter l'apprentissage de l'agent [56]. Cependant, la construction d'une fonction de récompense fait souvent apparaître des erreurs inattendues [39] et nécessite le plus souvent des compétences expertes. Par exemple, il est difficile de concevoir une récompense locale pour des tâches de navigation. En effet, il faudrait être capable de calculer le plus court chemin entre l'agent et son objectif, ce qui revient à résoudre le problème de navigation. D'un autre côté l'automatisation de la construction d'une récompense locale (sans faire appel à un expert) demande de trop grandes capacités de calcul [13].

#### 3.2 L'abstraction temporelle des actions

L'abstraction temporelle des actions consiste à utiliser des actions de haut niveau, aussi appelées *options*, pouvant

avoir des durées d'exécution différentes [58]. A chaque *option* est associée une politique intra-option qui définit les actions (de bas-niveau ou d'autres *options*) à réaliser dans chaque état lorsque l'*option* est exécutée. Abstraire les actions est un élément clé pour accélérer l'apprentissage. En effet, le nombre de choix à réaliser pour atteindre un objectif peut être fortement diminué si des *options* sont utilisées. Cela facilite aussi le renforcement des actions qui sont déterminantes pour l'obtention de la récompense (c'est le *credit assignment problem* [57]). Par exemple supposons qu'un robot essaye d'accéder à un gâteau sur une table. Si le robot a une *option* se rendre à la table et qu'il la suit, il ne lui restera qu'à prendre le gâteau. Il sera alors facile d'assimiler l'obtention du gâteau à l'*option* se rendre à la table. À l'inverse, si le robot doit apprendre à gérer chacune de ses articulations (actions de bas-niveau), il aura du mal à déterminer quelles actions de bas-niveau lui ont permis d'obtenir le gâteau, parmi toutes celles qu'il a réalisées.

Utiliser des *options* peut par ailleurs faciliter l'exploration lorsque les récompenses sont éparées. Pour illustrer cela, supposons que l'agent ait accès à l'*option* `Aller chercher la clé dans Montezuma's revenge`. Le problème devient trivial car une seule action d'exploration peut mener à une récompense là où il faudrait sans *option* une séquence spécifique d'actions de bas-niveau.

Concernant la politique intra-option, elle peut être définie manuellement [58], mais cela nécessite des compétences expertes, ou être apprises avec la fonction de récompenses [4], mais les *options* ne seront alors pas réutilisables pour d'autres tâches.

Pour résumer, l'utilisation de la motivation intrinsèque peut aider d'une part, à définir des politiques d'explorations plus élaborées permettant d'améliorer l'AR dans le cas de récompenses éparées; et d'autre part, à établir une abstraction des décisions ne dépendant pas de l'objectif.

### 4 Principales motivations intrinsèques

Dans cet article, nous proposons de catégoriser les différentes motivations intrinsèques utilisées en AR en deux classes de méta-compétences : l'acquisition de connaissances et la génération d'objectifs. Bien que l'*empowerment* soit primordial, nous n'avons pas l'espace pour développer la notion, nous renvoyons le lecteur intéressé vers la version complète de l'article<sup>2</sup>.

#### 4.1 Acquisition de connaissances

Les motivations intrinsèques basées sur l'acquisition de connaissances sont les plus utilisées en AR, de part leur capacité à rendre accessible des récompenses éparées. Nous allons étudier les trois principales méthodes existantes pour implémenter une récompense d'acquisition de connais-

2. <https://hal.archives-ouvertes.fr/hal-02272091>

sances. La première utilise le gain d’information, la seconde l’erreur de prédiction des états et la troisième l’évaluation de la nouveauté d’un état. Dans les trois cas, cette récompense intrinsèque permet de compléter la politique d’exploration. Nous focalisons notre étude sur des travaux récents et renvoyons le lecteur vers [53] pour une revue sur des travaux plus anciens concernant l’acquisition de connaissances.

**Le gain d’information.** Le gain d’information est une récompense basée sur la réduction de l’incertitude sur la dynamique de l’environnement suite à une action [42], ce qui peut aussi être assimilé au progrès d’apprentissage [51] ou à la surprise bayésienne [52]. Cela permet d’une part de diriger l’agent vers les zones déterministes qu’il ne connaît pas, d’autre part de l’empêcher d’aller dans les zones fortement stochastiques. En effet, si la zone est déterministe, les transitions de l’environnement sont prédictibles et son incertitude à propos des dynamiques de la zone peut baisser. Au contraire, lorsque les transitions sont stochastiques, l’agent se révèle incapable de prédire les transitions et ne réduit pas son incertitude. La stratégie d’exploration VIME [25] formalise le progrès d’apprentissage de manière bayésienne, l’intérêt de ces approches étant de pouvoir mesurer l’incertitude sur le modèle appris. L’agent apprend donc la dynamique de l’environnement via un réseau de neurone bayésien [20], et utilise la réduction d’incertitude sur la dynamique de l’environnement comme bonus intrinsèque. Autrement dit, l’agent essaye de faire des actions qui sont informatives sur le modèle des dynamiques. Similairement [2] propose de remplacer le modèle bayésien par un réseau de neurones classique suivi d’une distribution de probabilité gaussienne factorisée. Deux approches sont proposées : la première (NLL) utilise comme bonus intrinsèque l’entropie croisée de la prédiction et la seconde (AKL) l’amélioration de prédiction entre un instant  $t$  et après  $k$  améliorations à  $t + k$ . Bien que ces deux méthodes soient plus simples que VIME, les bénéfiques en terme de performance sont mitigés.

**L’erreur de prédiction.** L’idée est ici de diriger l’agent vers des zones pour lesquelles la prédiction de l’état suivant est difficile. Plutôt que de considérer la réduction d’erreur dans le modèle des dynamiques, l’agent utilise l’erreur directement comme récompense intrinsèque. Ainsi, au lieu d’utiliser des modèles probabilistes comme précédemment, Dynamic AE [55] calcule la distance entre l’état prédit et l’état réel dans un espace compressé via un auto-encodeur [24]. Cette distance fait ensuite office de récompense intrinsèque. Cependant cette approche est incapable de gérer la stochasticité locale de l’environnement [10]. Par exemple, il s’avère qu’ajouter une télévision affichant aléatoirement des images dans un environnement 3D attire l’agent; il va alors regarder passivement la télé puisqu’il sera incapable de prédire la prochaine observation. Ce problème est aussi appelé le *problème du bruit blanc* [44; 53]. Une solution serait de s’assurer que les transitions entre états soient apprenables, *i.e.* que la transition n’est

pas trop stochastique, mais cette problématique est difficile à résoudre en pratique.

Le module de curiosité intrinsèque (ICM) [44] apprend la dynamique de l’environnement dans un espace de caractéristiques. Il construit d’abord une représentation des états en apprenant à prédire les actions réalisées dans un état à partir de cet état et de l’état d’arrivée. Cela restreint la représentation à ce qui peut être contrôlé par l’agent. Il prédit ensuite dans cet espace l’état suivant. L’erreur de prédiction est alors utilisée comme récompense intrinsèque, ainsi l’erreur n’incorpore pas le bruit blanc puisque celui-ci ne dépend pas des actions. ICM permet notamment à un agent d’explorer son environnement dans les jeux *VizDoom* et *Super Mario Bros*. Dans *Super Mario Bros*, l’agent franchit 30% du premier niveau sans récompense extrinsèque. Finalement, en considérant toujours l’erreur de prédiction comme bonus intrinsèque, [10] propose un résumé des différentes manières de définir un espace de caractéristiques et montre, d’une part, qu’utiliser des caractéristiques aléatoires peut être performant mais peu généralisable lorsque l’environnement change, d’autre part qu’utiliser l’espace brut d’états (e.g. les pixels) est inefficace.

**Nouveauté d’un état.** Une première idée a été d’ajouter un bonus intrinsèque lorsque l’agent se dirige dans un état dans lequel il ne va jamais [9; 26]; ces méthodes sont dites basées sur le comptage. Au fur et à mesure qu’il visite un état, la récompense intrinsèque liée à cet état baisse. Bien que cette méthode soit efficace dans un environnement tabulaire (avec un espace d’états discretisé), elle est difficilement applicable lorsque les états sont très nombreux ou continus puisqu’on ne retourne alors jamais dans un même état. Une première solution proposée par [59], nommée TRPO-AE-hash, est de faire un hashage de l’espace d’états lorsqu’il est continu ou très grand. Cependant, les résultats ne sont que légèrement meilleurs que ceux d’une politique d’exploration classique. D’autres tentatives d’adaptation à un très grand espace d’états ont été proposées, comme DDQN-PC [5], A3C+ [5] ou DQN-PixelCNN [40], qui reposent sur des modèles de densité [61; 6]. Ces modèles permettent de calculer le *pseudo-count* [5], adaptation du comptage permettant la généralisation du décompte d’un état auprès des états avoisinants. Bien que ces algorithmes fonctionnent sur des environnements avec des récompenses éparées, les modèles de densité rajoutent une importante couche de complexité calculatoire [40]. Par ailleurs, même si ces modèles gèrent des espaces d’états de très haute dimension (e.g. pixels), ils ne peuvent pas être utilisés avec des espaces d’états continus.

Afin de diminuer la complexité calculatoire induite par les modèles de densité,  $\phi$ -EB [31] propose de ne pas modéliser la densité sur l’espace d’états brut, mais sur un espace de caractéristiques dénombrable induit par le calcul de  $V(s)$ . RND [11] mesure la nouveauté d’un état en distillant un réseau de neurones aléatoire (dont les poids sont figés) dans un autre réseau de neurones apprenant. Pour chaque état, le réseau de neurones aléatoire génère des caractéristiques

aléatoires. Le second réseau apprend à reproduire la sortie du réseau aléatoire pour chaque état. L'erreur de prédiction fait office de récompense. Cela revient à récompenser la nouveauté d'un état puisque l'erreur sera importante tant que le second réseau aura peu visité l'état en question, et sera faible lorsqu'il aura beaucoup appris dessus. L'agent n'arrive cependant pas à apprendre une exploration à long terme. Par exemple dans *Montezuma's revenge*, l'agent utilise ses clés pour ouvrir les premières portes qu'il voit mais n'arrive pas à accéder aux deux dernières portes. De plus, les caractéristiques aléatoires peuvent être insuffisantes pour représenter la richesse d'un environnement.

**Nouveauté comme écart aux autres états.** Une autre manière d'évaluer la nouveauté d'un état est de l'estimer comme la distance aux états habituellement parcourus. EX<sup>2</sup> [18] apprend un discriminateur pour différencier les états entre eux : lorsque le discriminateur n'arrive pas à différencier l'état courant de ceux d'un buffer, cela veut dire qu'il est peu allé dans cet état et l'agent recevra un bonus intrinsèque, et inversement lorsqu'il arrive à différencier l'état. Le module de curiosité épisodique ECO [49] approfondit cette idée en s'inspirant de la mémoire épisodique. Le modèle proposé contient un module de comparaison capable de renvoyer un bonus si l'agent est proche ou loin des états contenus dans un buffer. Ainsi, il calcule la probabilité que le nombre d'actions nécessaire pour aller de l'état sélectionné dans le buffer à l'état courant soit inférieure à un seuil. En stockant des états éparses dans le buffer, l'agent pose des points de repères dans l'environnement et essaye de s'éloigner de ceux-ci, cela revient à partitionner l'environnement. La probabilité que l'agent soit écarté de chaque état du buffer est utilisée pour calculer la récompense intrinsèque. Cependant pour calculer la récompense intrinsèque, l'agent doit comparer son observation courante à tous ses états en mémoire. Un passage à l'échelle de cette méthode risque d'être difficile lorsque la richesse de l'espace d'état deviendra plus grande, il serait en effet plus compliqué de partitionner efficacement l'espace d'état. D'un autre côté, l'avantage de cette méthode est que l'agent ne subit pas l'effet du bruit blanc (cf. §4.1).

## 4.2 Génération d'objectifs

La génération d'objectifs est la capacité d'un agent à apprendre des compétences diverses de manière non supervisée. Les compétences ou objectifs générés par l'agent sont des *options* (cf. §3). Comparé à l'AR multi-objectifs, les compétences sont ici générées de manière non supervisée. Dans les travaux utilisant la génération d'objectifs, l'agent apprend généralement sur deux échelles de temps : il génère des *options* et apprend les politiques intra-option associées en utilisant une récompense intrinsèque ; si un objectif global existe, il apprend à utiliser ces compétences pour réaliser cet objectif global, appelé tâche, en utilisant la récompense extrinsèque associée à la tâche. L'intérêt

est d'une part d'apprendre des compétences intéressantes pouvant servir à plusieurs tâches, or elles le seront d'autant plus qu'elles sont décorréliées de la tâche apprise [23]. D'autre part, l'abstraction temporelle des actions réalisée via les compétences facilite l'apprentissage. Prenons l'exemple de MuJoCo [60] qui est un environnement souvent utilisé dans les travaux sur la génération d'objectifs. Dans cet environnement, les articulations d'un robot peuvent être contrôlées par un agent pour accomplir par exemple des tâches de locomotion. L'idée de certains travaux est donc de générer des compétences de type *avancer* ou *reculer* avec une récompense intrinsèque. Ces compétences peuvent alors servir à une tâche de navigation.

Classiquement, l'AR a un seul objectif et n'apprend pas à réaliser plusieurs objectifs. Une manière de généraliser l'AR profond à l'apprentissage de plusieurs objectifs, voir à tous les objectifs possibles dans l'espace d'état, est d'utiliser l'approximateur de fonction de valeur universelle (UVFA) [50]. UVFA intègre la représentation de l'état objectif dans l'observation de l'agent. La politique trouvée est alors conditionnée sur l'objectif :  $\pi(s)$  devient  $\pi(s, g)$  où  $g$  est un objectif. Ainsi, les travaux apprenant les *options* avec une motivation intrinsèque apprennent des politiques  $\pi(s, g)$ . Bien que l'espace d'exploration augmente alors, [3] améliore l'efficacité des données en apprenant sur plusieurs objectifs à la fois via une seule interaction. En effet, via une interaction  $(s, s', r_g, a, g)$ , il est possible de créer une nouvelle interaction avec un nouvel objectif (lequel serait réussi)  $(s, s', r_{g'}, a, g')$  tant qu'une fonction de récompense  $R(s, a, s', g)$  est accessible, ce qui est généralement le cas lorsque la récompense est intrinsèque.

Dans la suite, nous étudierons les deux principaux ensembles de travaux portant sur l'auto-génération d'objectifs. La première approche utilise l'espace d'états pour générer les objectifs et calculer la récompense intrinsèque ; la seconde utilise la théorie de l'information.

**Distance entre objectifs.** Certains travaux utilisent l'espace d'états pour créer un espace d'objectif, l'intérêt est de pouvoir se servir de chaque état comme objectif. Ainsi, *Hierarchical Actor-Critic* (HAC) [29] se sert directement de l'espace d'états comme espace d'objectif pour apprendre trois niveaux d'options (les *options* du second niveau sont choisies de manière à répondre à l'*option* du troisième niveau). Une distance entre l'objectif et l'état final fait donc office de récompense intrinsèque. Au contraire, HIRO [37] utilise comme objectif la différence entre l'état initial et l'état à la fin de l'objectif ; la récompense intrinsèque est alors une distance entre la direction prise et l'objectif. Les objectifs permettent ainsi d'orienter les compétences vers certaines zones spatiales. Cependant, il y a deux problèmes dans l'utilisation de l'espace d'états comme espace d'objectifs. D'une part une distance (comme L2) a peu de sens dans un espace très grand comme une image

composée de pixels, d'autre part il est difficile de faire fonctionner un algorithme d'AR sur un espace d'action trop grand. Concrètement, un algorithme ayant comme espace d'objectif des images peut impliquer pour la politique d'*options* un espace d'action de 84x84 dimensions. Un espace d'action aussi large est actuellement inconcevable, aussi, ces algorithmes ne fonctionnent que sur des espaces d'états de faible dimension.

Le problème est finalement de construire une bonne représentation de l'espace d'états faisant office d'espace d'objectifs, i.e. choisir les informations à ne pas perdre lors de la compression des états en une nouvelle représentation. Ainsi, la distance entre deux objectifs aurait un sens et serait une bonne récompense intrinsèque. Pour construire un espace de caractéristiques, RIG [38] se sert d'un auto-encodeur variationnel (VAE) [27], mais ce type d'approche peut-être très sensible à des distracteurs (i.e. des caractéristiques inutiles à la tâche ou l'objectif, présentes dans les états) et ne permet donc pas de pondérer les caractéristiques correctement. Sub-optimal representation learning [36] essaye de borner la sous-optimalité de la représentation des objectifs, offrant des garanties théoriques. L'agent s'avère capable d'apprendre à aller partout en sélectionnant les caractéristiques importantes pour la tâche. [15] reformule l'équation de Bellman et présente des perspectives intéressantes en apprenant une représentation d'états pour laquelle la distance L2 entre deux états correspond au nombre d'actions à effectuer pour aller d'un état à l'autre. Il manque cependant des expérimentations montrant son intérêt.

#### Information mutuelle entre objectif et trajectoire.

Une deuxième approche ne nécessite pas de fonction de distance mais consiste à maximiser l'information mutuelle entre un objectif et sa trajectoire associée. Informellement, il s'agit d'apprendre des compétences selon la capacité de l'agent à les distinguer entre elles à partir de la trajectoire (i.e. les états parcourus) de la politique de la compétence choisie. Dans cette section, nous appelons  $I(g; c)$  l'information mutuelle entre  $g$ , un objectif, et  $c$ , une partie (changeante selon les travaux) de la trajectoire. L'information mutuelle est définie comme l'information que contient une variable aléatoire sur une autre variable aléatoire, i.e. avec  $H(x|y)$  désignant l'entropie d'une variable  $x$  sachant la variable  $y$  :

$$I(g; c) = H(g) - H(g|c). \quad (4)$$

[35] propose de maximiser l'information mutuelle en calculant une borne inférieure à cette information mutuelle, il en résulte :

$$I(g, c) \geq H(g) + \mathbb{E}_{p(s'|a,s)\pi(a|s)} \log q_\xi(g|c). \quad (5)$$

où  $q_\xi$  approxime  $p(g|c)$ . Plus d'information à ce sujet est disponible dans la version complète de l'article<sup>3</sup>.

3. <https://hal.archives-ouvertes.fr/hal-02272091>

SNN4HRL [16] apprend des compétences en maximisant l'information mutuelle  $I(g; c)$  où  $c$  est un agrégé de la trajectoire. Chaque objectif est généré de manière uniforme donc maximiser l'information mutuelle revient à minimiser  $H(g|c)$  (cf. équation 4). Or, c'est équivalent de maximiser la récompense intrinsèque  $\log q(g|c)$  (équation 5), où  $q$  est la probabilité prédite par un modèle. Pour calculer cet élément, il discrétise l'espace d'états et calcule la probabilité d'accomplir son objectif courant dans son état courant. Pour calculer cette probabilité, l'agent compte le nombre de fois où il a parcouru cette partition pour chaque objectif. Ensuite, l'agent ayant appris les compétences est intégré dans une structure hiérarchique dans laquelle un manager choisit les objectifs à accomplir. Notons que l'espace d'objectifs est ici discret.

VALOR [1] et DIAYN [14] reprennent la même idée, mais se distinguent des travaux précédents en utilisant un réseau de neurone plutôt qu'une discrétisation pour calculer  $\log q(g|c)$  et en choisissant  $c$  comme une partie de la trajectoire de la compétence dans l'environnement. Ils réussissent à faire apprendre à un agent des tâches de locomotion dans des espaces d'état ayant plus de 100 degrés de liberté. De plus, ils montrent l'intérêt de cette méthode utilisée comme pré-entraînement pour de l'apprentissage hiérarchique et comme initialisation pour l'apprentissage d'une tâche. DIAYN choisit  $c$  comme un état de la trajectoire et calcule la récompense intrinsèque à chaque itération de la trajectoire. VALOR se distingue en considérant  $c$  comme un agrégé de la trajectoire d'états et en assignant la récompense à la fin de la trajectoire. Avec VALOR, l'agent parvient à apprendre jusqu'à 10 compétences différentes et jusqu'à 100 en augmentant peu à peu le nombre d'objectifs via un curriculum [1]. VIC [21] avait déjà fait quelques expériences avec la même approche, mais sur des environnements plus simples et sans exhiber la même diversité de comportements.

Deux principales limites à cette approche peuvent être distinguées : l'agent est incapable d'apprendre à générer des objectifs sans désapprendre ses compétences. Ainsi, la distribution d'objectifs générée par l'agent doit rester uniforme [14]. Le calcul de  $\log q(g|c)$  implique par ailleurs que l'espace d'objectifs soit discrétisé. Il est donc impossible d'utiliser des *options* continues. DISCERN [62] s'attaque à ce problème en considérant l'espace d'objectifs comme l'espace d'états. Ensuite il fait une approximation de  $\log q(g|c)$  en essayant de classifier l'état final de la trajectoire auprès du bon objectif parmi d'autres objectifs tirés de la même distribution que le vrai objectif. Cela revient à apprendre à trouver l'objectif le plus proche de l'état final depuis une liste d'objectifs.

**Conclusion.** Pour résumer, il existe deux principaux ensembles de travaux portant sur l'auto-génération d'objectifs. Le premier ensemble apprend ses objectifs à l'aide de l'espace d'états, l'avantage est alors d'avoir un espace continu et lisse d'objectifs permettant l'interpolation. Le

désavantage est qu'il faut trouver la bonne métrique de comparaison et la juste manière de compresser l'espace d'états. Le second ensemble de travaux utilise la théorie de l'information, cela force l'utilisation d'un espace discret d'objectifs mais évite les complications associées à un espace continu qui est originellement de haute dimension.

## 5 Motivation intrinsèque comme compression de l'information

### 5.1 Compression de l'information

Schmidhuber propose que l'organisme est guidé par le désir de compresser l'information qu'il reçoit [52]. Ainsi, plus nous arrivons à compresser les données reçues de l'environnement, plus la récompense intrinsèque reçue est élevée. Il note toutefois que c'est l'amélioration de la compression qui est importante et non le degré de compression en lui-même, sous peine qu'un agent soit inactif devant du bruit ou devant une obscurité uniforme. Or la compression de données est fortement liée à l'observation de régularités dans ces mêmes données. Par exemple ce que nous appelons visage est, dans notre environnement, un ensemble, apparaissant de manière récurrente, composé d'une forme ovale contenant deux yeux, un nez et une bouche. De même, un état de l'environnement peut être décrit avec quelques-unes des caractéristiques les plus pertinentes. Cela implique que la motivation intrinsèque se traduit par une recherche de nouvelles régularités dans l'environnement.

### 5.2 Lien avec les travaux

Il a été montré dans [52; 25] que les travaux sur le gain d'information sont directement liés au progrès de la compression d'information; le module de curiosité épisodique [49] essaye d'encoder l'environnement en sauvegardant les états les plus diversifiés possibles; et les modèles prédictifs [10] encodent les dynamiques de l'environnement dans un modèle paramétré (souvent un réseau de neurone). Les travaux sur la génération d'objectifs cherchent explicitement à compresser des trajectoires dans un espace d'objectifs. Enfin, une partie des travaux [36; 44] repose sur la qualité d'une compression de l'espace d'état.

### 5.3 Conclusion

Pour résumer, l'ensemble des travaux s'attache à compresser les nouvelles régularités détectées dans les trajectoires de l'agent. Pour cela, la théorie de l'information est un puissant outil de mesure.

## 6 Limites et challenges

Plusieurs travaux sont limités par des problématiques sortant du cadre de l'AR, telles que les performances des modèles de densité [5; 40], la difficulté d'approximation de l'information mutuelle entre deux variables aléatoires continues [21] ou les performances des modèles prédictifs [37; 38]. Ces limites dépassent le cadre de cet article.

Aussi, malgré l'hétérogénéité des travaux sur la motivation intrinsèque en AR, et les limitations propres à chacune de ses méthodes, nous avons identifié et présentons dans cette section quatre problématiques majeures qui sont communes à l'ensemble des approches.

### 6.1 Stochasticité de l'environnement

Nous avons vu dans la partie précédente qu'il était intéressant de maximiser le progrès de compression et que la plupart des travaux étaient relatifs à la compression d'information, et non au progrès de la compression. Cet écart explique la difficulté de plusieurs travaux [10] à gérer l'effet du bruit blanc [53] ou plus généralement la stochasticité de l'environnement. Certains travaux de l'état de l'art gèrent cette problématique [49; 44; 11; 25], mais chacun avec ses défauts.

### 6.2 Acquisition de connaissance sur le long-terme

A notre connaissance, aucune approche existante n'est capable de gérer la recherche d'information long-terme [11]. Dans *Montezuma's revenge*, il s'agit d'éviter d'utiliser un clé trop rapidement pour pouvoir l'utiliser plus tard. Dans la vie de tous les jours, il peut s'agir d'éviter de dépenser son argent trop vite. Cette difficulté pourrait être résolue avec une approche utilisant la planification [22]. L'apprentissage hiérarchique de compétences pourrait apporter une solution, en transformant le long terme en court terme via une hiérarchie de compétences multi-niveaux [47]. Finalement, il manque d'autres environnements que *Montezuma's revenge* mettant vraiment en avant cette problématique.

### 6.3 Construire une représentation des états

Nous avons vu que construire de bonnes caractéristiques d'état est important dans la découverte d'objectifs afin de travailler sur un espace d'objectifs réduit. C'est aussi primordial dans les travaux sur l'acquisition de connaissances pour avoir une erreur de prédiction significative. Le module de curiosité intrinsèque ICM [44] propose une représentation des états intéressante, restreinte à ce qui peut être contrôlé par l'agent, mais sa limite est que le module apprend une partie suffisante des caractéristiques permettant de déterminer l'action, et non l'ensemble des caractéristiques déterminées par l'action. Il manque encore un moyen de compresser parfaitement l'espace d'états dans l'espace des caractéristiques contrôlées par l'agent. [15] est allé dans ce sens en apprenant une représentation d'états pour laquelle la distance L2 entre deux états correspond au nombre d'actions à effectuer pour aller d'un état à l'autre, mais il manque des expérimentations. Par ailleurs, [28] présente plusieurs pistes pouvant améliorer la représentation des états d'un agent.

### 6.4 Décorrélérer les objectifs de la tâche

L'avantage de décorrélérer l'apprentissage des objectifs de l'apprentissage d'une tâche est de favoriser l'exploration et

le transfert d'apprentissage. On parle alors d'apprentissage *bottom-up* car on apprend les compétences avant la tâche. Si cet apprentissage a fait des progrès significatifs, il est encore impossible d'apprendre des tâches spécifiques en même temps que les compétences permettant de les réaliser sans subir d'oubli catastrophique [32]. En effet, lorsque l'agent apprend séquentiellement des tâches, il oublie les premières tâches en apprenant les suivantes. Des travaux adressant le problème d'oubli catastrophique existent déjà [43] mais ils n'ont, à notre connaissance, pas été évalués avec la motivation intrinsèque et un large nombre de tâches.

## 7 Conclusion

L'AR fait face à plusieurs défis, comme l'apprentissage avec des récompenses éparées ou l'abstraction des actions de l'agent en décisions de plus haut niveau. Nous avons vu que la motivation intrinsèque pouvait être utilisée en AR et que ses nombreuses applications pouvait résoudre partiellement ces problématiques. Plusieurs types de motivations intrinsèques existent comme méta-compétences, chacune avec leur littérature. Parmi celles-ci, l'acquisition de connaissances est effectuée via des modèles prédictifs, des modèles bayésiens ou des modèles de densité pour inciter l'agent à explorer l'environnement. La génération d'objectifs permet d'abstraire les actions de l'agent, aidant à résoudre le *credit assignment problem*. Lorsque l'apprentissage est de type *bottom-up*, l'abstraction des décisions facilite l'exploration et le transfert d'apprentissage. Plusieurs défis restent cependant à adresser : les mécanismes d'acquisition de connaissance gèrent difficilement la stochasticité de l'environnement et ont encore des difficultés à générer des trajectoires utiles à l'exploration sur le long terme ; avoir des représentations d'état plus significatives pourrait ouvrir de nouvelles perspectives pour la génération d'objectifs ; les travaux apprenant des compétences multi-tâches souffrent encore de l'oubli catastrophique. Notre analyse suggère que la compression de l'information contenue dans les séquences d'interactions de l'agent, devrait jouer un rôle prédominant dans la résolution des défis mentionnés.

## Références

- [1] Joshua Achiam, Harrison Edwards, Dario Amodei, and Pieter Abbeel. Variational option discovery algorithms. *arXiv preprint arXiv :1807.10299*, 2018.
- [2] Joshua Achiam and Shankar Sastry. Surprise-based intrinsic motivation for deep reinforcement learning. *arXiv preprint arXiv :1703.01732*, 2017.
- [3] Marcin Andrychowicz, Filip Wolski, Alex Ray, Jonas Schneider, Rachel Fong, Peter Welinder, Bob McGrew, Josh Tobin, OpenAI Pieter Abbeel, and Wojciech Zaremba. Hindsight experience replay. In *Advances in Neural Information Processing Systems*, pages 5048–5058, 2017.
- [4] Pierre-Luc Bacon, Jean Harb, and Doina Precup. The option-critic architecture. In *AAAI*, pages 1726–1734, 2017.
- [5] Marc Bellemare, Sriram Srinivasan, Georg Ostrovski, Tom Schaul, David Saxton, and Remi Munos. Unifying count-based exploration and intrinsic motivation. In *Advances in Neural Information Processing Systems*, pages 1471–1479, 2016.
- [6] Marc Bellemare, Joel Veness, and Erik Talvitie. Skip context tree switching. In *International Conference on Machine Learning*, pages 1458–1466, 2014.
- [7] Marc G. Bellemare, Yavar Naddaf, Joel Veness, and Michael Bowling. The arcade learning environment : An evaluation platform for general agents (extended abstract). In *IJCAI*, pages 4148–4152. AAAI Press, 2015.
- [8] Daniel E Berlyne. Structure and direction in thinking. 1965.
- [9] Ronen I Brafman and Moshe Tennenholtz. R-max-a general polynomial time algorithm for near-optimal reinforcement learning. *Journal of Machine Learning Research*, 3(Oct) :213–231, 2002.
- [10] Yuri Burda, Harri Edwards, Deepak Pathak, Amos Storkey, Trevor Darrell, and Alexei A. Efros. Large-scale study of curiosity-driven learning. In *International Conference on Learning Representations*, 2019.
- [11] Yuri Burda, Harrison Edwards, Amos Storkey, and Oleg Klimov. Exploration by random network distillation. *arXiv preprint arXiv :1810.12894*, 2018.
- [12] Nicolò Cesa-Bianchi, Claudio Gentile, Gábor Lugosi, and Gergely Neu. Boltzmann exploration done right. In *Advances in Neural Information Processing Systems*, pages 6284–6293, 2017.
- [13] Hao-Tien Lewis Chiang, Aleksandra Faust, Marek Fiser, and Anthony Francis. Learning navigation behaviors end-to-end with autorl. *IEEE Robotics and Automation Letters*, 4(2) :2007–2014, 2019.
- [14] Benjamin Eysenbach, Abhishek Gupta, Julian Ibarz, and Sergey Levine. Diversity is all you need : Learning skills without a reward function. *CoRR*, abs/1802.06070, 2018.
- [15] Carlos Florensa, Jonas Degraeve, Nicolas Heess, Jost Tobias Springenberg, and Martin Riedmiller. Self-supervised learning of image embedding for continuous control. *arXiv preprint arXiv :1901.00943*, 2019.
- [16] Carlos Florensa, Yan Duan, and Pieter Abbeel. Stochastic neural networks for hierarchical reinforcement learning. In *5th International Conference on Learning Representations, ICLR 2017, Toulon, France, April 24-26, 2017, Conference Track Proceedings*, 2017.



- [17] Vincent François-Lavet, Peter Henderson, Riashat Islam, Marc G Bellemare, Joelle Pineau, et al. An introduction to deep reinforcement learning. *Foundations and Trends® in Machine Learning*, 11(3-4) :219–354, 2018.
- [18] Justin Fu, John Co-Reyes, and Sergey Levine. Ex2 : Exploration with exemplar models for deep reinforcement learning. In *Advances in Neural Information Processing Systems*, pages 2577–2587, 2017.
- [19] Alison Gopnik, Andrew N Meltzoff, and Patricia K Kuhl. *The scientist in the crib : Minds, brains, and how children learn*. William Morrow & Co, 1999.
- [20] Alex Graves. Practical variational inference for neural networks. In *Advances in neural information processing systems*, pages 2348–2356, 2011.
- [21] Karol Gregor, Danilo Jimenez Rezende, and Daan Wierstra. Variational intrinsic control. *arXiv preprint arXiv :1611.07507*, 2016.
- [22] Danijar Hafner, Timothy P. Lillicrap, Ian Fischer, Ruben Villegas, David Ha, Honglak Lee, and James Davidson. Learning latent dynamics for planning from pixels. *CoRR*, abs/1811.04551, 2018.
- [23] Nicolas Heess, Greg Wayne, Yuval Tassa, Timothy Lillicrap, Martin Riedmiller, and David Silver. Learning and transfer of modulated locomotor controllers. *arXiv preprint arXiv :1610.05182*, 2016.
- [24] Geoffrey E Hinton and Ruslan R Salakhutdinov. Reducing the dimensionality of data with neural networks. *science*, 313(5786) :504–507, 2006.
- [25] Rein Houthoofd, Xi Chen, Yan Duan, John Schulman, Filip De Turck, and Pieter Abbeel. Vime : Variational information maximizing exploration. In *Advances in Neural Information Processing Systems*, pages 1109–1117, 2016.
- [26] Michael Kearns and Satinder Singh. Near-optimal reinforcement learning in polynomial time. *Machine learning*, 49(2-3) :209–232, 2002.
- [27] Diederik P Kingma and Max Welling. Auto-encoding variational bayes. *arXiv preprint arXiv :1312.6114*, 2013.
- [28] Timothée Lesort, Natalia Díaz-Rodríguez, Jean-François Goudou, and David Filliat. State representation learning for control : An overview. *Neural Networks*, 2018.
- [29] Andrew Levy, Robert Platt, and Kate Saenko. Hierarchical reinforcement learning with hindsight. In *International Conference on Learning Representations*, 2019.
- [30] Timothy P Lillicrap, Jonathan J Hunt, Alexander Pritzel, Nicolas Heess, Tom Erez, Yuval Tassa, David Silver, and Daan Wierstra. Continuous control with deep reinforcement learning. *arXiv preprint arXiv :1509.02971*, 2015.
- [31] Jarryd Martin, Suraj Narayanan Sasikumar, Tom Everitt, and Marcus Hutter. Count-based exploration in feature space for reinforcement learning. In *Proceedings of the Twenty-Sixth International Joint Conference on Artificial Intelligence, IJCAI 2017, Melbourne, Australia, August 19-25, 2017*, pages 2471–2478, 2017.
- [32] Michael McCloskey and Neal J Cohen. Catastrophic interference in connectionist networks : The sequential learning problem. In *Psychology of learning and motivation*, volume 24, pages 109–165. Elsevier, 1989.
- [33] Volodymyr Mnih, Adria Puigdomenech Badia, Mehdi Mirza, Alex Graves, Timothy Lillicrap, Tim Harley, David Silver, and Koray Kavukcuoglu. Asynchronous methods for deep reinforcement learning. In *International conference on machine learning*, pages 1928–1937, 2016.
- [34] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fidfjeld, Georg Ostrovski, et al. Human-level control through deep reinforcement learning. *Nature*, 518(7540) :529, 2015.
- [35] Shakir Mohamed and Danilo Jimenez Rezende. Variational information maximisation for intrinsically motivated reinforcement learning. In *Advances in neural information processing systems*, pages 2125–2133, 2015.
- [36] Ofir Nachum, Shixiang Gu, Honglak Lee, and Sergey Levine. Near-optimal representation learning for hierarchical reinforcement learning. In *International Conference on Learning Representations*, 2019.
- [37] Ofir Nachum, Shixiang (Shane) Gu, Honglak Lee, and Sergey Levine. Data-efficient hierarchical reinforcement learning. In S. Bengio, H. Wallach, H. Larochelle, K. Grauman, N. Cesa-Bianchi, and R. Garnett, editors, *Advances in Neural Information Processing Systems 31*, pages 3303–3313. 2018.
- [38] Ashvin V Nair, Vitchyr Pong, Murtaza Dalal, Shikhar Bahl, Steven Lin, and Sergey Levine. Visual reinforcement learning with imagined goals. In *Advances in Neural Information Processing Systems*, pages 9209–9220, 2018.
- [39] Andrew Y Ng, Daishi Harada, and Stuart Russell. Policy invariance under reward transformations : Theory and application to reward shaping. In *ICML*, volume 99, pages 278–287, 1999.
- [40] Georg Ostrovski, Marc G Bellemare, Aaron van den Oord, and Rémi Munos. Count-based exploration with neural density models. *arXiv preprint arXiv :1703.01310*, 2017.
- [41] Pierre-Yves Oudeyer and Frederic Kaplan. How can we define intrinsic motivation? In *Proceedings of*

- the 8th International Conference on Epigenetic Robotics : Modeling Cognitive Development in Robotic Systems, Lund University Cognitive Studies, Lund : LUCS, Brighton.* Lund University Cognitive Studies, Lund : LUCS, Brighton, 2008.
- [42] Pierre-Yves Oudeyer and Frederic Kaplan. What is intrinsic motivation ? a typology of computational approaches. *Frontiers in neurorobotics*, 1 :6, 2009.
- [43] German I Parisi, Ronald Kemker, Jose L Part, Christopher Kanan, and Stefan Wermter. Continual lifelong learning with neural networks : A review. *Neural Networks*, 2019.
- [44] Deepak Pathak, Pulkit Agrawal, Alexei A Efros, and Trevor Darrell. Curiosity-driven exploration by self-supervised prediction. In *International Conference on Machine Learning (ICML)*, volume 2017, 2017.
- [45] Jean Piaget and Margaret Cook. *The origins of intelligence in children*, volume 8. International Universities Press New York, 1952.
- [46] Matthias Plappert, Rein Houthoofd, Prafulla Dhariwal, Szymon Sidor, Richard Y Chen, Xi Chen, Tamim Asfour, Pieter Abbeel, and Marcin Andrychowicz. Parameter space noise for exploration. *arXiv preprint arXiv :1706.01905*, 2017.
- [47] Matthew Riemer, Miao Liu, and Gerald Tesauro. Learning abstract options. In *Advances in Neural Information Processing Systems*, pages 10445–10455, 2018.
- [48] Richard M Ryan and Edward L Deci. Intrinsic and extrinsic motivations : Classic definitions and new directions. *Contemporary educational psychology*, 25(1) :54–67, 2000.
- [49] Nikolay Savinov, Anton Raichuk, Raphaël Marinier, Damien Vincent, Marc Pollefeys, Timothy Lillicrap, and Sylvain Gelly. Episodic curiosity through reachability. *arXiv preprint arXiv :1810.02274*, 2018.
- [50] Tom Schaul, Daniel Horgan, Karol Gregor, and David Silver. Universal value function approximators. In *International Conference on Machine Learning*, pages 1312–1320, 2015.
- [51] Jürgen Schmidhuber. Curious model-building control systems. In *[Proceedings] 1991 IEEE International Joint Conference on Neural Networks*, pages 1458–1463. IEEE, 1991.
- [52] Jürgen Schmidhuber. Driven by compression progress : A simple principle explains essential aspects of subjective beauty, novelty, surprise, interestingness, attention, curiosity, creativity, art, science, music, jokes. In *Workshop on Anticipatory Behavior in Adaptive Learning Systems*, pages 48–76. Springer, 2008.
- [53] Jürgen Schmidhuber. Formal theory of creativity, fun, and intrinsic motivation (1990–2010). *IEEE Transactions on Autonomous Mental Development*, 2(3) :230–247, 2010.
- [54] David Silver, Julian Schrittwieser, Karen Simonyan, Ioannis Antonoglou, Aja Huang, Arthur Guez, Thomas Hubert, Lucas Baker, Matthew Lai, Adrian Bolton, et al. Mastering the game of go without human knowledge. *Nature*, 550(7676) :354, 2017.
- [55] Bradley C Stadie, Sergey Levine, and Pieter Abbeel. Incentivizing exploration in reinforcement learning with deep predictive models. *arXiv preprint arXiv :1507.00814*, 2015.
- [56] Pei-Hao Su, David Vandyke, Milica Gasic, Nikola Mrksic, Tsung-Hsien Wen, and Steve Young. Reward shaping with recurrent neural networks for speeding up on-line policy learning in spoken dialogue systems. *arXiv preprint arXiv :1508.03391*, 2015.
- [57] Richard S Sutton and Andrew G Barto. *Reinforcement learning : An introduction*, volume 1. MIT press Cambridge, 1998.
- [58] Richard S Sutton, Doina Precup, and Satinder Singh. Between mdps and semi-mdps : A framework for temporal abstraction in reinforcement learning. *Artificial intelligence*, 112(1-2) :181–211, 1999.
- [59] Haoran Tang, Rein Houthoofd, Davis Foote, Adam Stooke, OpenAI Xi Chen, Yan Duan, John Schulman, Filip DeTurck, and Pieter Abbeel. # exploration : A study of count-based exploration for deep reinforcement learning. In *Advances in neural information processing systems*, pages 2753–2762, 2017.
- [60] Emanuel Todorov, Tom Erez, and Yuval Tassa. Mujoco : A physics engine for model-based control. In *2012 IEEE/RSJ International Conference on Intelligent Robots and Systems*, pages 5026–5033. IEEE, 2012.
- [61] Aaron Van den Oord, Nal Kalchbrenner, Lasse Espeholt, Oriol Vinyals, Alex Graves, et al. Conditional image generation with pixelcnn decoders. In *Advances in Neural Information Processing Systems*, pages 4790–4798, 2016.
- [62] David Warde-Farley, Tom Van de Wiele, Tejas Kulkarni, Catalin Ionescu, Steven Hansen, and Volodymyr Mnih. Unsupervised control through non-parametric discriminative rewards. *arXiv preprint arXiv :1811.11359*, 2018.
- [63] Robert W White. Motivation reconsidered : The concept of competence. *Psychological review*, 66(5) :297, 1959.

## Stratégie situationnelle pour l'équilibrage de charge

Quentin Baert<sup>1</sup>

Anne-Cécile Caron<sup>1</sup>

Maxime Morge<sup>1</sup>

Jean-Christophe Routier<sup>1</sup>

Kostas Stathis<sup>2</sup>

<sup>1</sup> Centre de Recherche en Informatique, Signal et Automatique de Lille, Université de Lille, France

<sup>2</sup> Department of Computer Science, Royal Holloway, University of London, Egham, United Kingdom

{quentin.baert,anne-cecile.caron,maxime.morge,jean-christophe.routier}@univ-lille.fr  
kostas.stathis@rhul.ac.uk

### Résumé

*Nous étudions une stratégie qui tient compte de la localité des ressources pour équilibrer les charges dans un système distribué. Cette stratégie permet aux agents coopératifs d'identifier une allocation non équilibrée, voire de déclencher des enchères concurrentes pour réallouer localement certaines des tâches. Les tâches sont réallouées en tenant compte de l'accessibilité des ressources pour les agents. Nous évaluons cette stratégie dans le cadre du déploiement multi-agents de MapReduce. Ce patron de conception permet le traitement distribué de données massives. Les résultats empiriques démontrent que notre stratégie améliore significativement le temps d'exécution du traitement d'un jeu de données.*

### Mots Clef

Systeme multi-agents, négociation.

### Abstract

*We study a novel location-aware strategy for distributed systems where cooperating agents perform the load-balancing. The strategy allows agents to identify opportunities within a current unbalanced allocation. The tasks are reallocated according to the proximity of the resources. We evaluate the strategy in a multi-agent deployment of the MapReduce design pattern for processing large data-sets. Empirical results demonstrate that our strategy significantly improves the overall runtime of the data processing.*

### Keywords

Multi-agent system, negotiation.

## 1 Introduction

Le problème d'équilibrage des charges et d'allocation des tâches dans un système distribué apparaît dans de nom-

breuses applications telles que l'informatique en nuage, les réseaux pair-à-pair, les réseaux sociaux et le traitement de données massives. Dans cet article, nous abordons les problèmes applicatifs où (a) certaines ressources, par exemple des données, nécessaires à l'exécution d'une tâche sont distribuées sur différents nœuds du réseau, et (b) certains des nœuds sont susceptibles de rencontrer des perturbations lors de l'exécution, par exemple une chute de performance ou des latences réseau. Comme plusieurs ressources sont requises pour exécuter une tâche, toute allocation implique inévitablement le transfert de ces ressources entre les nœuds de calculs, ce qui induit un délai supplémentaire lors du traitement de la tâche [14]. Dans cette classe de problèmes, l'allocation de tâches peut être remise en cause lors de l'exécution des tâches et tirer parti de la distribution des ressources dans le système.

Afin d'aborder le problème d'équilibrage des charges et d'allocation des tâches dans des applications telles que celles qui motivent ce travail, les technologies multi-agents ont fait l'objet d'une grande attention, particulièrement celles qui portent sur l'ordonnancement à l'aide d'enchères [24, 21, 23]. En plus de la décentralisation qui permet d'éviter les goulots d'étranglement en terme de performance, nous montrons ici qu'une approche multi-agents pour l'allocation située de tâches répond à deux autres exigences cruciales : (a) la co-occurrence de la réallocation des tâches et de leur exécution ; (b) l'adaptation de l'allocation, c'est-à-dire le déclenchement d'une réallocation quand une perturbation se produit. Nous supposons que les agents sont coopératifs. Ils partagent le même objectif : diminuer le temps d'exécution global des tâches. Nous supposons également que les agents ne partagent aucune connaissance, y compris l'état de l'allocation courante. Néanmoins, les agents possèdent un modèle de leurs pairs : ils sont ainsi capables de calculer le coût des tâches pour leurs pairs. Nous supposons également qu'une tâche peut être exécutée par n'importe quel agent seul, et ce sans préemption ni contrainte de priorité. Une tâche est indivisible, sans date butoir et non partageable, c'est-à-dire qu'une tâche n'appartient qu'à un agent à la fois.

Nous formalisons ici le problème d'allocation de tâches situées. En adoptant une approche fondée sur le marché, nous décentralisons totalement le processus d'équilibre des charges grâce auquel les agents coopératifs essaient de minimiser le temps d'exécution de la dernière tâche, i.e. le *makespan*. À cette fin, nous proposons une stratégie situationnelle afin de procéder à l'équilibrage des charges<sup>1</sup>. Quand les agents identifient des opportunités pour rééquilibrer l'allocation courante, ils initient des enchères de manière concurrente pour réallouer localement certaines tâches. Ces dernières sont réallouées en considérant la proximité entre les ressources nécessaires à leur exécution et le nœud sur lequel elles sont effectivement exécutées. Ce processus de réallocation de tâches est dynamique et continu. Il est concurrent à l'exécution des tâches, ce qui rend le système distribué adaptatif aux événements perturbateurs (consommation de tâche, chute de performance d'un nœud).

Notre application pratique est la distribution de MapReduce. Ce patron de conception, dont Hadoop est une implémentation, permet de traiter de larges volumes de données sur des grappes de calculs [9]. Plusieurs biais apparaissent sur des jeux de données réels et mènent à un déséquilibre des charges [16]. Un ordonnanceur centralisé ne peut pas être utilisé comme point de référence de par le grand nombre de tâches à traiter. Cependant, un processus de négociation multi-agents permet de réallouer les tâches lors de la phase de *reduce* pour réduire le temps d'exécution du *job*. Nos résultats empiriques préliminaires montrent que la stratégie situationnelle permet d'améliorer significativement le temps de traitement d'un jeu de données.

Cet article est structuré comme suit. Après une présentation des travaux connexes en section 2, nous formalisons le problème d'allocation de tâches situées à l'aide d'un système multi-agents en section 3. Cette section définit également la délégation socialement rationnelle de tâches, ce qui permet aux agents d'améliorer localement une allocation de tâches. La section 4 illustre le processus de négociations itérées qui a lieu en même temps que l'exécution des tâches. La section 5 décrit la stratégie situationnelle, c'est-à-dire comment les agents choisissent quelle tâche exécuter/négocier. Notre application pratique et notre validation empirique sont décrites dans la section 6. Enfin, la section 7 résume notre contribution et présente nos perspectives.

## 2 Travaux connexes

Les problèmes d'ordonnancement classiques ont fait l'objet d'un grand nombre d'études et de recherches. Ces dernières ont abouti à des ordonnanceurs hors-ligne pour des modèles simples [6]. Le problème de la minimisation du *makespan* (le temps auquel la dernière tâche est achevée)

<sup>1</sup>. Cet article est une version courte de [5] où, dans la continuité de [4], la localité des ressources est prise en compte.

pour  $n$  tâches sur  $m$  machines hétérogènes (avec des capacités différentes), appelé  $R||C_{max}$ , est NP-difficile [12]. Même si l'heuristique *earliest completion time* [13] (ECT) est un algorithme d'approximation qui atteint des résultats acceptables avec un coût computationnel faible, les algorithmes centralisés ne peuvent pas être appliqués à notre scénario où les tâches sont nombreuses (par exemple 82,283 clés dans la section 6). La stratégie situationnelle est une heuristique de recherche locale décentralisée.

L'ordonnancement multi-agents [14] a suscité beaucoup d'intérêt dans le cadre du problème d'équilibrage de charge pour les systèmes distribués. Ce problème est différent des problèmes d'ordonnancement classiques de par ses exigences :

- **passage à l'échelle.** Un contrôleur global constitue un goulot d'étranglement en terme de performance car il doit collecter les informations d'état de l'ensemble du système en temps réel. Au lieu de cela, la répartition des tâches peut être négociée par des agents représentant les nœuds [24, 1] ;
- **réactivité.** Les problèmes d'ordonnancement classiques sont statiques. L'estimation inexacte du temps d'exécution des tâches, aggravée par des perturbations (consommation de tâches, ralentissement des nœuds, etc.), peuvent nécessiter d'importantes modifications de l'allocation existante pour qu'elle reste optimale [23]. Plutôt que de recalculer en continu une allocation optimale, quelques modifications locales au cours de l'exécution des tâches peuvent améliorer l'équilibre de charge [3, 25, 15].

La plupart des travaux existants adoptent une approche fondée sur le marché [24, 11, 18] qui modélise l'équilibrage de charge comme un jeu non coopératif afin d'optimiser des métriques centrées sur l'utilisateur plutôt que des métriques centrées sur le système, telles que le temps d'achèvement global que nous utilisons dans cet article.

Contrairement à notre travail, [21] assignent les tâches à des coalitions d'agents car ils ne peuvent pas réaliser les tâches seuls. Dans cet article, nous supposons qu'une tâche, qui peut être exécutée par n'importe quel agent seul sans préemption ni contrainte de précedence, est indivisible, non partageable (i.e. une tâche n'appartient qu'à un seul agent à la fois) et sans date butoir. [8] ont étudié l'allocation multi-agents de ressources mais leur travail se limite à l'affectation d'une seule ressource par agent. Nous supposons ici qu'un lot de tâches peut être assigné à chaque agent. Dans la continuité de [2, 10], [19] proposent des méthodes potentiellement distribuables pour l'allocation de ressources qui s'appuient sur la négociation multi-agents. Même si la topologie du réseau n'entre pas dans le cadre de notre travail (nous partons du principe que les agents sont pleinement connectés), nous allons un peu plus loin en décentralisant effectivement la négociation et en remettant en jeu cette allocation pendant le traitement des tâches.

Il est bien connu que le traitement de jeux de données réels via le déploiement distribué du patron de conception MapReduce comporte souvent des biais de données qui peuvent conduire à un déséquilibre des charges [16]. En particulier, le biais de partitionnement se produit lorsqu'un *reducer* traite un plus grand nombre de clés que les autres. Puisque le *job* se termine lorsque toutes les tâches *reduce* sont terminées, le temps d'exécution du *job* est pénalisé par le *reducer* le plus chargé. Ce biais de données est abordé par [7, 17] à l'aide d'un paramétrage dépendant de connaissances préalables sur des données et sur l'environnement d'exécution. Nous abordons cette question à l'aide d'une réallocation de tâches dynamique et adaptative qui est concurrente à la consommation des tâches. Ainsi, la réallocation s'adapte au traitement des données. Cela nous permet d'aborder les problèmes réels suivants : (a) l'absence de connaissance préalable sur les données, (b) l'estimation inexacte du temps d'exécution des tâches, et (c) les aléas d'exécutions (chute de performance d'un nœud, latence réseau). À notre connaissance, aucune autre proposition n'est, comme la nôtre, réactive et capable de passer à l'échelle.

### 3 Allocation de tâches situées

Nous formalisons maintenant le problème de l'allocation multi-agents de tâches situées (MASTA). Ici les tâches ont des coûts différents selon les agents, en fonction de la localité des ressources.

**Définition 1** (MASTA). *Un problème d'allocation multi-agents de tâches situées de taille  $(k, m, n)$  avec  $k \geq 1$ ,  $m \geq 2$  et  $n \geq 1$  est un  $n$ -uplet MASTA =  $\langle Node, \mathcal{A}, \mathcal{T}, l, d, c \rangle$  tel que :*

- $Node = \{node_1, \dots, node_k\}$  est un ensemble de  $k$  nœuds ;
- $\mathcal{A} = \{1, \dots, m\}$  est un ensemble de  $m$  agents ;
- $\mathcal{T} = \{\tau_1, \dots, \tau_n\}$  est un ensemble de  $n$  tâches à traiter ;
- $l : \mathcal{A} \mapsto Node$  donne la localisation d'un agent ;
- $d : \mathcal{T} \times Node \mapsto \mathbb{N}^+$  donne le nombre de ressources d'une tâche  $\tau$  situées sur un nœud  $x$  ;
- $c : \mathcal{T} \times Node \mapsto \mathbb{R}_+^*$  donne le coût d'une tâche  $\tau$  lorsqu'elle est exécutée sur un nœud. Plus une tâche est locale, moins elle coûte chère :

$$\forall i, j \in \mathcal{A}, d(\tau, l(i)) > d(\tau, l(j)) \Rightarrow c_i(\tau, l(i)) \leq c(\tau, l(j)) \quad (1)$$

Dans la suite de l'article, on écrira  $l_i$  (resp.  $c_i(\tau)$ ,  $d_i(\tau)$ ) à la place de  $l(i)$  (resp. de  $c(\tau, l_i)$ ,  $d(\tau, l_i)$ ). De la même façon, on écrira  $d_\tau$  pour  $\sum_{node \in Node} d(\tau, node)$ . On dit que  $\tau$  est locale (resp. semi-locale, distante) pour l'agent  $i$  si  $d_i(\tau) = d_\tau$  (resp.  $d_i(\tau) < d_\tau$ ,  $d_i(\tau) = 0$ ).

Étant donné un problème MASTA particulier, nous évaluons d'un point de vue collectif une allocation de tâches

répondant à ce problème, en considérant le temps nécessaire pour achever la dernière tâche, i.e. le *makespan*.

**Définition 2** (Allocation de tâches/Charge de travail/Makespan). *Une allocation de tâches  $P$  est une partition des tâches parmi les agents, i.e un ensemble de  $m$  lots de tâches  $\{P(1), \dots, P(m)\}$  tel que :*

$$\cup_{i \in \mathcal{A}} P(i) = \mathcal{T} \quad (2)$$

$$\forall i \in \mathcal{A}, \forall j \in \mathcal{A} \setminus \{i\}, P(i) \cap P(j) = \emptyset \quad (3)$$

*La charge de travail de l'agent  $i \in \mathcal{A}$  pour l'allocation  $P$  est définie par :*

$$w_i(P) = \sum_{\tau \in P(i)} c_i(\tau) \quad (4)$$

*Le makespan de  $P$  est défini par :*

$$C_{max}(P) = \max\{w_i(P) \mid i \in \mathcal{A}\} \quad (5)$$

Les agents réalisent des délégations de tâches qui sont socialement rationnelles afin d'améliorer l'allocation.

**Définition 3** (Délégation socialement rationnelle). *Soit  $P$  une allocation de tâches. La délégation  $\delta$  de la tâche  $\tau$  par l'agent  $i$  à l'agent  $j$  permet d'obtenir l'allocation  $\delta(P) = \{P'(1), \dots, P'(m)\}$  telle que :*

$$\forall k \in \mathcal{A} \setminus \{i, j\}, P'(k) = P(k) \quad (6)$$

$$P'(i) = P(i) \setminus \{\tau\} \wedge P'(j) = P(j) \cup \{\tau\} \quad (7)$$

*La délégation est socialement rationnelle si et seulement si :*

$$w_j(P) + c_j(\tau) < w_i(P) \quad (8)$$

Comme la délégation socialement rationnelle  $\delta$  décroît strictement le *makespan* local entre deux agents, elle ne peut pas augmenter le *makespan* global ( $C_{max}(\delta(P)) \leq C_{max}(P)$ ).

Nous pouvons maintenant noter  $\Gamma_i(P)$  l'ensemble des délégations socialement rationnelles pour l'agent  $i$  :

$$\Gamma_i(P) = \{\tau \in P(i) \mid w_j(P) + c_j(\tau) < w_i(P)\} \quad (9)$$

Une allocation de tâches  $P$  est dite stable si aucun agent ne peut réaliser de délégation socialement rationnelle, i.e.  $\forall i \in \mathcal{A}, \Gamma_i(P) = \emptyset$ .

**Propriété 1.** *On peut toujours aboutir à une allocation stable à partir d'une allocation non stable en utilisant un nombre fini de délégations socialement rationnelles.*

Voici un exemple pour illustrer les différentes définitions.

**Exemple 1.** *Considérons le problème MASTA<sup>ex</sup> =  $\langle Node, \mathcal{A}, \mathcal{T}, l, d, c \rangle$  de taille  $(2, 2, 7)$ , où  $Node = \{node_1, node_2\}$ ,  $\mathcal{A} = \{1, 2\}$  et  $\mathcal{T} = \{\tau_1, \dots, \tau_7\}$  avec  $l(1) = node_1$ ,  $l(2) = node_2$ . Les localisations et coûts des tâches sont donnés dans*

	$\tau_1$	$\tau_2$	$\tau_3$	$\tau_4$	$\tau_5$	$\tau_6$	$\tau_7$
$d_1(\tau_k)$	1	0	3	6	1	6	0
$d_2(\tau_k)$	0	4	6	12	4	1	7
$c_1(\tau_k)$	1	8	15	30	9	8	14
$c_2(\tau_k)$	2	4	12	24	6	13	7

TABLE 1 – Localisations et coûts des tâches

le tableau 1. Soient les allocations  $Pmks$  et  $P$  telles que  $Pmks = \{\{\tau_1, \tau_3, \tau_5, \tau_6\}, \{\tau_2, \tau_4, \tau_7\}\}$  et  $P = \{\{\tau_2, \tau_4, \tau_6\}, \{\tau_1, \tau_3, \tau_5, \tau_7\}\}$ .  $Pmks$  est optimale puisque  $C_{max}(Pmks) = 35$  ( $w_1(Pmks) = 33$  et  $w_2(Pmks) = 35$ ). Ce n'est pas le cas de  $P$  puisque  $w_1(P) = 46$ ,  $w_2(P) = 27$ , et donc  $C_{max}(P) = 46$ . De plus,  $\Gamma_1(P) = \{\tau_2, \tau_6\}$  et  $\Gamma_2(P) = \emptyset$ .

Soit  $P' = \delta_1(P)$  l'allocation obtenue par la délégation  $\delta_1$  de la tâche  $\tau_6$  par l'agent 1 à l'agent 2. Comme  $w_1(P') = 38$  et  $w_2(P') = 40$ ,  $\delta_1$  améliore le makespan (i.e.  $C_{max}(P') = 40$ ). Cependant,  $P'$  n'est pas stable car  $\Gamma_2(P') = \{\tau_1\}$ .

Soit  $P'' = \delta_2(P')$  l'allocation obtenue via la délégation  $\delta_2$  de la tâche  $\tau_1$  par l'agent 2 à l'agent 1. Même si  $P''$  n'est pas optimale ( $C_{max}(P'') > C_{max}(Pmks)$ ),  $P''$  est stable car  $\Gamma_1(P'') = \Gamma_2(P'') = \emptyset$ .

## 4 Processus de négociation

Cette section présente le processus de négociation qui est itéré et concurrent à la consommation des tâches. Quand une tâche est exécutée, elle est retirée de l'ensemble des tâches et le système multi-agents cherche à minimiser le makespan d'un nouveau problème MASTA. En effet, l'accomplissement d'une tâche est un événement disruptif qui modifie les paramètres du problème et l'allocation des tâches.

**Définition 4** (Consommation de tâche). Soit  $P$  l'allocation de tâches courante pour le problème MASTA  $= \langle Node, \mathcal{A}, \mathcal{T}, l, d, c \rangle$ . La consommation  $\gamma$  de la tâche  $\tau$  par l'agent  $i$  aboutit à l'allocation  $P' = \gamma(P)$  pour le problème MASTA'  $= \langle Node, \mathcal{A}, \mathcal{T}', l, d, c \rangle$  telle que :

$$\mathcal{T}' = \mathcal{T} \setminus \{\tau\} \quad (10)$$

$$P'(i) = P(i) \setminus \{\tau\} \quad (11)$$

$$\forall j \in \mathcal{A} \setminus \{i\}, P'(j) = P(j) \quad (12)$$

À l'évidence, la consommation d'une tâche diminue le makespan. La séquence des consommations, qui retirent peu à peu toutes les tâches de l'allocation initiale jusqu'à l'allocation vide  $\perp$ , consiste en une itération de problèmes MASTA.

**Processus décentralisé de délégations de tâches.** Pour déléguer des tâches, les agents réalisent des négociations concurrentes. Chaque négociation est basée sur le protocole Contract Net [22]. Il y a 3 étapes de décision durant

une négociation : (a) l'initiateur de l'enchère choisit une tâche selon une des stratégies décrites dans la section 5, (b) chaque agent propose/refuse de prendre la tâche selon que la délégation est socialement rationnelle ou non, et (c) l'initiateur sélectionne le vainqueur de l'enchère, e.g. l'enchérisseur qui a la plus petite charge de travail. Plusieurs négociations impliquant différents groupes d'agents peuvent se dérouler simultanément et les agents peuvent participer à plusieurs enchères en même temps. La concurrence des délégations de tâches améliore la réactivité de la réallocation des tâches.

Comme il n'y a pas de partage de connaissances, un agent a des croyances partielles éventuellement erronées concernant l'allocation courante  $P$ . En effet, l'agent  $i$  connaît sa charge de travail  $w_i(P)$  et dispose de sa propre base de croyances :

$$\mathcal{B}_i(P) = \langle w_1^i(P), \dots, w_{i-1}^i(P), w_{i+1}^i(P), \dots, w_m^i(P) \rangle \quad (13)$$

où  $w_j^i(P)$  est ce que l'agent  $i$  croit connaître de la charge de travail de l'agent  $j$  dans l'allocation  $P$ .

L'ensemble des délégations qui sont potentiellement socialement rationnelles  $\Gamma_i^{\mathcal{B}}(P)$  et que l'agent  $i$  peut initier à partir de l'allocation  $P$  s'appuie sur sa base de croyances  $\mathcal{B}_i(P)$ . Formellement,

$$\Gamma_i^{\mathcal{B}}(P) = \{\tau \in P(i) \mid w_j^i(P) + c_j(\tau) < w_i(P)\}. \quad (14)$$

Lorsqu'un agent souhaite initier une enchère, le calcul du makespan local s'appuie sur sa base de croyances, éventuellement erronée. C'est le prix à payer pour la décentralisation. Cependant, un agent informe ses pairs de sa charge de travail au début du traitement (i.e. lorsqu'il se fait connaître de ses pairs), lorsqu'il envoie des messages en tant qu'initiateur ou enchérisseur durant la négociation, et à chaque fois qu'il consomme une tâche. Ainsi la base de croyances d'un agent est mise à jour quand il reçoit un appel d'offre, et l'agent refusera une délégation de tâche qui n'est pas socialement rationnelle. Une négociation réussie aboutit nécessairement à une délégation de tâche socialement rationnelle qui ne peut pas détériorer le makespan. En adoptant une approche conservatrice [20], nous nous assurons qu'une négociation couronnée de succès améliore strictement l'allocation de tâches et donc que le processus de négociation converge.

**Consommations et délégations concurrentes.** Les délégations et consommations de tâches sont concurrentes et complémentaires, puisque la disparition d'une tâche peut permettre de nouvelles délégations socialement rationnelles. La figure 1 représente leur impact sur l'allocation, jusqu'à ce que toutes les tâches soient exécutées. À partir de l'allocation initiale  $P_0$ , les agents réalisent des délégations socialement rationnelles pour améliorer le makespan (e.g. le chemin de  $P_0$  à  $P_k$ ) jusqu'à la consommation d'une tâche (e.g. l'arc de  $P_k$  à  $P'_0$ ), ce qui interrompt un

chemin vers une allocation stable (e.g. le chemin en gris de  $P_k$  à  $P_{stable}$ ). Une consommation de tâche peut aussi se produire après avoir atteint une allocation stable (e.g. après  $P'_{stable}$ ).

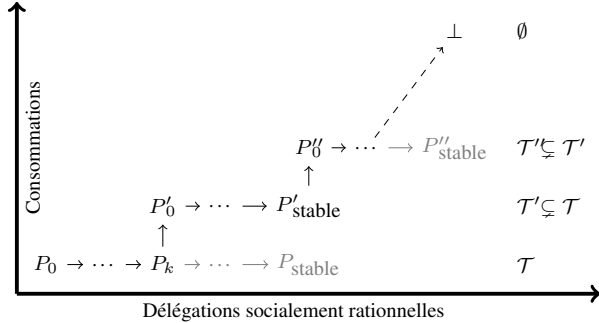


FIGURE 1 – Consommations de tâches (arcs verticaux) et délégations de tâches (arcs horizontaux) concurrentes.

## 5 Stratégie situationnelle

Au cours du processus, les négociations et les consommations de tâches sont concurrentes. La stratégie d'un agent consiste à choisir la prochaine délégation potentiellement socialement rationnelle et la prochaine tâche à exécuter.

**Stratégie agnostique vis-à-vis de la localité.** Dans une première approche, un agent  $i$  peut exécuter les tâches les plus grandes de son propre lot et négocier les plus petites parmi les délégations potentiellement socialement rationnelles. Formellement,

$$\hat{\tau} = \operatorname{argmax}_{\tau \in P(i)} \{c_i(\tau)\} \quad (15)$$

$$\check{\tau} = \operatorname{argmin}_{\tau \in \Gamma_i^B(P)} \{c_i(\tau)\} \quad (16)$$

où  $\hat{\tau}$  (resp.  $\check{\tau}$ ) représente la prochaine tâche à exécuter (resp. à négocier). Cette stratégie nécessite que l'agent trie son lot de tâches selon leurs coûts. Cependant, acquérir une ressource consomme du temps et représente un coût supplémentaire lors de l'exécution.

Afin de mesurer la localité des tâches, on définit le ratio de possession d'une tâche par un agent  $i$  comme le ratio entre le nombre de ressources locales à l'agent pour cette tâche et le nombre total de ressources de la tâche :

$$o_i(\tau) = \frac{d_i(\tau)}{d_\tau} \quad (17)$$

Le ratio maximal de possession de  $\tau$  est :

$$\hat{o}(\tau) = \max_{i \in \mathcal{A}} \{o_i(\tau)\} \quad (18)$$

**Stratégie situationnelle.** Intuitivement, un agent devrait d'abord exécuter les tâches qui peuvent coûter plus pour

les autres que pour lui-même, et déléguer les tâches qui peuvent coûter moins pour les autres. Selon cette stratégie, un agent exécute d'abord les grandes tâches locales et négocie d'abord les grandes tâches distantes, en fonction de ses connaissances et croyances locales, i.e. le coût de chaque tâche et son ratio de possession. Cette stratégie utilise une structure de données appelée lot par possession.

Le **lot par possession** de l'agent  $i$  (voir figure 2) est partagé en trois paquets selon le ratio de possession de l'agent pour ces tâches.

1. Le *paquet local* (noté  $MB$ ) contient les tâches dont l'agent possède au moins une ressource et dont il est le plus gros propriétaire. Formellement,

$$\forall \tau \in MB, o_i(\tau) \neq 0 \wedge o_i(\tau) = \hat{o}(\tau) \quad (19)$$

Dans  $MB$ , les tâches sont triées par ordre décroissant de coût (cf. gauche de la figure 2).

2. Le *paquet semi-local* (noté  $IB$ ) contient les tâches partiellement locales. Formellement,

$$\forall \tau \in IB, 0 < o_i(\tau) < \hat{o}(\tau). \quad (20)$$

Dans  $IB$ , les tâches sont triées par ordre décroissant de ratio de possession et les tâches ayant le même ratio de possession sont triées par ordre décroissant de coût (cf. centre de la figure 2).

3. Le *paquet distant* (noté  $DB$ ) contient les tâches distantes. Formellement,

$$\forall \tau \in DB, o_i(\tau) = 0. \quad (21)$$

Dans  $DB$ , les tâches sont triées par ordre croissant de coût. (cf. droite de la figure 2).

Quand un agent recherche une tâche à exécuter, il commence par le début du paquet local, i.e. par la tâche locale la plus grande. Quand un agent cherche une tâche à négocier, il commence par la fin du paquet distant (i.e. la plus grande tâche distante) et il sélectionne la première qui constitue une délégation potentiellement socialement rationnelle selon ses croyances,  $\mathcal{B}_i(P)$ . Formellement,

- Un agent cherche une tâche à exécuter  $\hat{\tau} \in MB$  telle que

$$\forall \tau \in MB \setminus \{\hat{\tau}\}, c_i(\hat{\tau}) \geq c_i(\tau). \quad (22)$$

Si  $MB = \emptyset$ , il cherche  $\hat{\tau} \in IB$  telle que

$$\begin{aligned} \forall \tau \in IB \setminus \{\hat{\tau}\}, \\ o_i(\hat{\tau}) > o_i(\tau) \vee \\ (o_i(\hat{\tau}) = o_i(\tau) \wedge c_i(\hat{\tau}) \geq c_i(\tau)). \end{aligned} \quad (23)$$

Si  $IB = \emptyset$ , il cherche  $\hat{\tau} \in DB$  telle que

$$\forall \tau \in DB \setminus \{\hat{\tau}\}, c_i(\hat{\tau}) \leq c_i(\tau). \quad (24)$$

Finalement, si  $MB = IB = DB = \emptyset$ , l'agent n'a aucune tâche à exécuter.

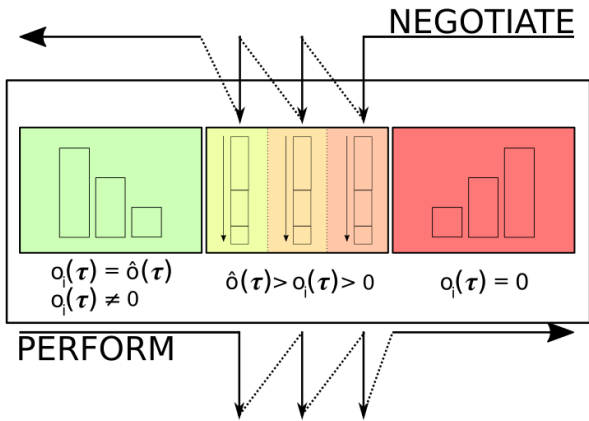


FIGURE 2 – Le lot par possession contient le paquet local (à gauche), le paquet semi-local (au centre) et le paquet distant (à droite). Les tâches sont représentées par des rectangles dont la taille est proportionnelle au coût de la tâche. Les flèches désignent l'ordre de parcours selon lequel un agent va chercher une tâche à exécuter/négocier.

- Un agent cherche une tâche à négocier  $\tilde{\tau} \in DB$  telle que

$$\forall \tau \in (DB \cap \Gamma_i^B(P)) \setminus \{\tilde{\tau}\}, c_i(\tilde{\tau}) \geq c_i(\tau). \quad (25)$$

- Si  $DB = \emptyset$ , il cherche  $\tilde{\tau} \in IB$  telle que

$$\begin{aligned} \forall \tau \in (IB \cap \Gamma_i^B(P)) \setminus \{\tilde{\tau}\}, \\ o_i(\tilde{\tau}) < o_i(\tau) \vee \\ (o_i(\tilde{\tau}) = o_i(\tau) \wedge c_i(\tilde{\tau}) \geq c_i(\tau)). \end{aligned} \quad (26)$$

- Si  $IB = \emptyset$ , il cherche  $\tilde{\tau} \in MB$  telle que

$$\begin{aligned} \forall \tau \in (MB \cap \Gamma_i^B(P)) \setminus \{\tilde{\tau}\}, \\ c_i(\tilde{\tau}) \leq c_i(\tau). \end{aligned} \quad (27)$$

Finalement, si l'agent ne trouve pas une telle tâche, il arrête d'initier des enchères.

Il est important de remarquer que l'ordre leximax sur le paquet semi-local est conforme au principe selon lequel on exécute d'abord les grandes tâches locales et on négocie d'abord les grandes tâches distantes. Quand un agent explore le paquet semi-local, il commence par les tâches qui ont un ratio de possession élevé afin de les exécuter, et par celles qui ont un ratio de possession faible afin de les négocier. Quand un agent examine les tâches qui ont le même ratio de possession, il le fait toujours dans le même ordre : des plus coûteuses vers les moins coûteuses (cf. figure 2). Donc, quand un agent explore le paquet semi-local à la recherche d'une tâche à exécuter, il considère la plus locale et plus coûteuse ; et quand un agent explore le paquet semi-local à la recherche d'une tâche à négocier, il considère la moins locale et plus coûteuse.

	$\tau_1$	$\tau_2$	$\tau_3$	$\tau_4$	$\tau_5$	$\tau_6$	$\tau_7$
$o_1(\tau_k)$	1	0	0.33	0.33	0.2	0.85	0
$o_2(\tau_k)$	0	1	0.66	0.66	0.8	0.14	1

TABLE 2 – Ratios de possession pour l'exemple 2

**Exemple 2** (Lot par possession). Reprenons l'exemple 1 où nous supposons qu'une ressource distante est deux fois plus coûteuse qu'une ressource locale. Formellement,

$$c_i(\tau) = d_i(\tau) + 2(d_\tau - d_i(\tau)) = 2d_\tau - d_i(\tau) \quad (28)$$

Le tableau 2 donne le ratio de possession de chaque agent pour chaque tâche. Considérons l'allocation où toutes les tâches sont affectées à l'agent 1. Selon la stratégie agnostique, l'agent 1 ne fait pas de différence entre  $\tau_2$  et  $\tau_6$  pour une délégation puisque ces tâches ont le même coût pour lui. Selon la stratégie situationnelle, l'agent 1 dont le lot par possession est donné figure 3 considère :

- l'exécution des tâches  $\tau_6, \tau_1, \tau_4, \tau_3, \tau_5, \tau_2$ , puis  $\tau_7$  ;
- la négociation des tâches  $\tau_7, \tau_2, \tau_5, \tau_4, \tau_3, \tau_1$ , puis  $\tau_6$ .

La stratégie situationnelle amène l'agent 1 à exécuter  $\tau_6$  et à négocier  $\tau_7$  afin de diminuer le makespan, puisque  $\tau_6$  est en grande partie locale et  $\tau_7$  est distante. En particulier, l'agent 1 exécute  $\tau_6$  avant  $\tau_1$  car  $\tau_6$  est plus coûteuse (cf. table 1).

Nous soulignons que les tâches dans le paquet semi-local sont triées d'abord selon le ratio de possession puis selon leur coût : que l'agent 1 cherche à négocier ou à exécuter une tâche, il considère celles qui ont le même ratio maximal de possession et dans le même ordre (par ex.  $\tau_4$ , puis  $\tau_3$ ).

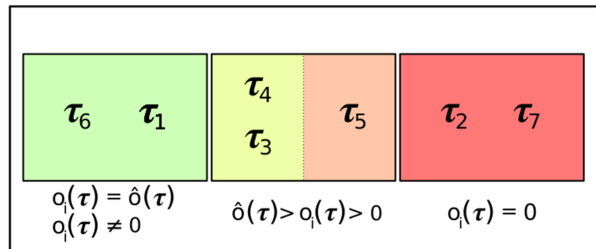


FIGURE 3 – Lot par possession de l'agent 1

## 6 Application pratique

Comme application pratique, nous considérons le déploiement distribué du patron de conception MapReduce pour le traitement de jeux de données de grandes tailles sur une grappe de serveurs [9]. [16] identifient plusieurs biais qui pénalisent le temps nécessaire au traitement des données. En particulier, le biais de partitionnement se produit quand un *reducer* doit traiter un plus grand nombre de clés (et donc de tâches) que les autres. Ce biais produit un déséquilibre dans les charges de travail des *reducers*. La stratégie



situationnelle permet aux agents d'équilibrer les charges et ainsi d'améliorer le temps d'exécution.

Nous avons déployé MapReduce à l'aide d'un système distribué multi-agents. Notre prototype<sup>2</sup> a été implémenté en Scala avec la boîte à outils Akka<sup>3</sup> adaptée aux applications orientées messages, fortement concurrentes, distribuées et robustes.

Nos expériences<sup>4</sup> utilisent un jeu de données de 8 Gio (82, 283 clés) qui a été généré de telle sorte que l'allocation initiale des tâches (cf. figure 4) soit déséquilibrée et de manière à pouvoir vérifier que la proximité entre les données et les nœuds de traitement a un impact sur le *make-span* et donc sur le temps d'exécution. L'allocation initiale des tâches est le résultat de la phase de partitionnement du processus MapReduce. En d'autres termes, quand il n'y a pas de négociation, notre système a le même comportement que Hadoop.

Nous comparons, à partir de 10 exécutions pour chacune des (ré)allocations, les temps d'exécutions médians quand les agents adoptent une stratégie agnostique ou situationnelle. Nous observons que la stratégie situationnelle améliore significativement le temps d'exécution, d'environ -7.6%. Pour comparaison, le temps d'exécution sans remise en cause du partitionnement par défaut d'Hadoop est de 853s (environ +100%). Nous en déduisons que le coût de la négociation peut être négligé au regard du gain obtenu par équilibrage des charges. De plus, la négociation permet de diminuer le temps d'exécution, tout particulièrement avec la stratégie situationnelle.

En raison du non-déterminisme de l'exécution distribuée, nous présentons un *run* typique. La figure 4 compare l'allocation de tâches quand toutes les tâches ont été traitées, qu'elles aient été négociées ou pas entre les 16 agents<sup>5</sup>, selon les stratégies agnostique ou situationnelle. Les deux stratégies sont comparées *ex-post*. Nous remarquons que le *makespan* de l'allocation initiale vaut approximativement  $3.3 \cdot 10^8$ , il est de  $2.5 \cdot 10^8$  (-24%) en cas de négociation avec la stratégie agnostique locale et de  $2 \cdot 10^8$  (-30.7%) avec la stratégie situationnelle. Nous en déduisons que la négociation, en particulier avec la stratégie situationnelle, permet d'améliorer l'équilibre des charges.

Puisqu'elle favorise l'exécution des tâches par les nœuds qui sont les plus proches des chunks, la stratégie situationnelle améliore l'équilibre des charges et diminue le temps d'exécution.

2. <https://github.com/cristal-smac/mas4data>

3. <http://akka.io>

4. Nos expériences ont été réalisées sur 16 PCs quadricœurs Intel(R) i7 avec 16GB RAM chacun.

5. Une analyse empirique préliminaire montre qu'il n'y a pas de valeur ajoutée à avoir plus d'un *reducer* par nœud/disque.

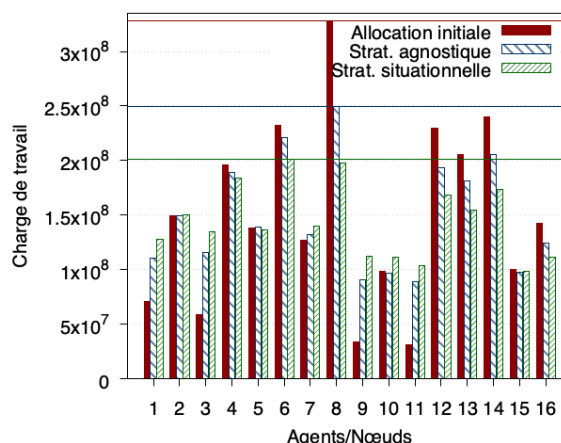


FIGURE 4 – Allocation initiale des tâches (Hadoop) et allocation *ex-post* avec les stratégies agnostique et situationnelle

## 7 Conclusion

Dans cet article, nous avons formalisé le problème de l'allocation multi-agents de tâches situées et nous avons proposé une stratégie situationnelle pour traiter le problème de l'équilibrage des charges dans les systèmes distribués. Cette stratégie adopte une approche fondée sur le marché qui nous permet de décentraliser l'équilibrage de charge pour minimiser le *makespan*, i.e. le temps nécessaire pour achever la dernière tâche. Au cours du processus, les négociations de tâches permettent aux agents qui utilisent la stratégie situationnelle de réallouer les tâches en même temps qu'ils en consomment. Cette stratégie permet également à un agent de déterminer la prochaine délégation socialement rationnelle et la prochaine tâche à traiter. Conformément à leur propre base de croyances et à leurs connaissances, les agents traitent en priorité les grandes tâches locales et négocient en priorité les grandes tâches distantes. Afin de valider notre approche, nous avons développé un prototype dans lequel les agents négocient les tâches de *reduce* du patron de conception MapReduce dans une configuration distribuée. Nos résultats expérimentaux montrent que la stratégie situationnelle proposée améliore l'équilibre de charge et donc le temps d'exécution pour de telles applications.

Actuellement, nous évaluons notre prototype dans différentes configurations matérielles, avec de nombreuses expérimentations qui se basent sur des jeux de données réels. Certaines configurations nous amènent à considérer que nous devons étendre notre stratégie pour négocier (a) des échanges de tâches afin d'améliorer le *makespan* des allocations stables, et (b) des lots de tâches pour accélérer les négociations.

**Remerciements.** Ce travail est soutenu par l'AAP ULille « Internationalisation Actions bilatérales ».

## Références

- [1] Bo An, Victor Lesser, David Irwin, and Michael Zink. Automated negotiation with decommitment for dynamic resource allocation in cloud computing. In *Proc. of AAMAS*, pages 981–988, 2010.
- [2] Martin R. Andersson and Tuomas W. Sandholm. Contract types for satisficing task allocation : Ii experimental results. In *Proc. of the AAAI spring symposium : Satisficing models*, pages 1–7, 1998.
- [3] Gamal Attiya and Yskandar Hamam. Task allocation for maximizing reliability of distributed systems : A simulated annealing approach. *Journal of Parallel and Distributed Computing*, 66(10) :1259–1266, 2006.
- [4] Quentin Baert, Anne-Cécile Caron, Maxime Morge, and Jean-Christophe Routier. Stratégie de découpe de tâche pour le traitement de données massives. In *Actes des JFSMA*, pages 65–74, 2017.
- [5] Quentin Baert, Anne-Cécile Caron, Maxime Morge, Jean-Christophe Routier, and Kostas Stathis. Stratégie situationnelle pour l'équilibrage de charge. In *Actes des JFSMA*, pages 9–18, 2019.
- [6] Bo Chen, Chris N. Potts, and Gerhard J Woeginger. *Handbook of combinatorial optimization*, chapter A Review of Machine Scheduling : Complexity, Algorithms and Approximability, pages 1493–1641. Springer, 1998.
- [7] Quan Chen, Daqiang Zhang, Minyi Guo, Qianni Deng, and Song Guo. SAMR : A self-adaptive MapReduce scheduling algorithm in heterogeneous environment. In *International Conference on Computer and Information Technology*, pages 2736–2743. IEEE, 2010.
- [8] Anastasia Damamme, Aurélie Beynier, Yann Chevalleyre, and Nicolas Maudet. The Power of Swap Deals in Distributed Resource Allocation. In *Proc. of AAMAS*, pages 625–633, 2015.
- [9] J. Dean and S. Ghemawat. MapReduce : Simplified data processing on large clusters. In *Symposium on Operating Systems Design and Implementation*, pages 137–150, 2004.
- [10] Ulle Endriss, Nicolas Maudet, Fariba Sadri, and Francesca Toni. Negotiating socially optimal allocations of resources. *JAIR*, 25(1) :315–348, 2006.
- [11] Saurabh Kumar Garg, Srikumar Venugopal, James Broberg, and Rajkumar Buyya. Double auction-inspired meta-scheduling of parallel applications on global grids. *Journal of Parallel and Distributed Computing*, 73(4) :450–464, 2013.
- [12] Ellis Horowitz and Sartaj Sahni. Exact and approximate algorithms for scheduling nonidentical processors. *JACM*, 23(2) :317–327, 1976.
- [13] Oscar H Ibarra and Chul E Kim. Heuristic algorithms for scheduling independent tasks on nonidentical processors. *JACM*, 24(2) :280–289, 1977.
- [14] Yichuan Jiang. A survey of task allocation and load balancing in distributed systems. *IEEE Transactions on Parallel and Distributed Systems*, 27(2) :585–599, 2016.
- [15] Qin-Ma Kang, Hong He, Hui-Min Song, and Rong Deng. Task allocation for maximizing reliability of distributed computing systems using honeybee mating optimization. *Journal of Systems and Software*, 83(11) :2165–2174, 2010.
- [16] YongChul Kwon, Kai Ren, Magdalena Balazinska, and Bill Howe. Managing skew in Hadoop. *IEEE Data Eng. Bull.*, 36(1) :24–33, 2013.
- [17] Miguel Liroz-Gistau, Reza Akbarinia, and Patrick Valduriez. FP-Hadoop : efficient execution of parallel jobs over skewed data. *VLDB Endowment*, 8(12) :1856–1859, 2015.
- [18] Amro Najjar, Olivier Boissier, and Gauthier Picard. Négociation adaptative pour l'acceptabilité des services d'un fournisseur SaaS. In *Actes des JFSMA*, pages 85–94, 2017.
- [19] Antoine Nongaillard and Philippe Mathieu. Reallocation problems in agent societies : A local mechanism to maximize social welfare. *JASSS*, 14(3) :21, 2011.
- [20] Michael Schillo, Christian Kray, and Klaus Fischer. The eager bidder problem : a fundamental problem of DAI and selected solutions. In *Proc. of AAMAS*, pages 599–606. ACM, 2002.
- [21] Onn Shehory and Sarit Kraus. Methods for task allocation via agent coalition formation. *Artif. Intell.*, 101(1-2) :165–200, 1998.
- [22] Reid G. Smith. The contract net protocol : High-level communication and control in a distributed problem solver. *IEEE Transactions on computers*, 29(12) :1104–1113, December 1980.
- [23] Joanna Turner, Qinggang Meng, Gerald Schaefer, and Andrea Soltoggio. Distributed strategy adaptation with a prediction function in multi-agent task allocation. In *Proc. of AAMAS*, pages 739–747, 2018.
- [24] William E Walsh and Michael P Wellman. A market protocol for decentralized task allocation. In *Proc. of ICMAS*, pages 325–332, 1998.
- [25] Peng-Yeng Yin, Shiuh-Sheng Yu, Pei-Pei Wang, and Yi-Te Wang. Task allocation for maximizing reliability of a distributed system using hybrid particle swarm optimization. *Journal of Systems and Software*, 80(5) :724–735, 2007.

# Elect : Une méthode de gestion des incohérences dans des ontologies légères partiellement préordonnées\*

Sihem Belabbes<sup>1</sup>Salem Benferhat<sup>1</sup>Jan Chomicki<sup>2</sup><sup>1</sup> CRIL, Université d'Artois, CNRS-UMR 8188, Lens, France<sup>2</sup> SUNY at Buffalo, Buffalo, NY, USA

{belabbes, benferhat}@cril.fr, chomicki@buffalo.edu

## Résumé

Nous nous intéressons au problème de la gestion des incohérences dans des ontologies légères. Nous supposons que les bases de connaissances terminologiques (TBox) sont spécifiées dans les logiques de description DL-Lite. Nous supposons aussi que les bases assertionnelles (ABox) sont partiellement préordonnées et peuvent être incohérentes avec la base terminologique TBox. Une contribution importante de cet article est la proposition d'une méthode efficace et saine, appelée *Elect*, pour restaurer la cohérence de la base ABox par rapport à la base TBox. Notre méthode permet de retrouver la sémantique dite IAR (Intersection ABox Repair) lorsque la base ABox est non-ordonnée. Elle permet également de retrouver la sémantique dite non-contestée lorsque la base ABox est totalement préordonnée. La justification sémantique de la méthode *Elect* est basée d'abord sur le fait qu'une base ABox partiellement préordonnée est interprétée comme une famille de bases ABox totalement préordonnées. Ensuite, la méthode *Elect* applique l'inférence non-contestée à chacune de ces ABox totalement préordonnées. Dans la deuxième partie de l'article, nous introduisons le concept d'assertions, dites élues, qui permet de donner une caractérisation équivalente de la méthode *Elect* sans avoir à générer explicitement toutes les bases ABox totalement préordonnées. Nous montrons que le calcul des assertions élues se fait en temps polynomial.

## Mots Clef

Incohérence. Ontologies légères. Bases de connaissances partiellement préordonnées

## Abstract

We focus on the problem of handling inconsistency in lightweight ontologies. We assume that terminological knowledge bases (TBox) are specified in DL-Lite and that assertional facts (ABox) are partially preordered and may be inconsistent with respect to the TBox. One of the main contributions of this paper is the provision of an efficient and safe

\*Cet article est paru en anglais dans les actes de 15<sup>th</sup> International Conference on Logic Programming and Nonmonotonic Reasoning (LPNMR), Philadelphie, USA, 3–7 Juin 2019, pages 210–223, Springer.

method, called *Elect*, to restore consistency of the ABox with respect to the TBox. In the case where the ABox is flat (no priorities are associated with the assertions) or totally preordered, our method collapses with the well-known IAR semantics and non-defeated semantics, respectively. The semantic justification of *Elect* is obtained by first viewing a partially preordered ABox as a family of totally preordered ABoxes, and then applying non-defeated inference to each of the totally preordered ABoxes. We introduce the concept of elected assertions which allows us to provide an equivalent characterisation of *Elect* without explicitly generating all totally preordered ABoxes. Finally we show that the computation of *Elect* is done in polynomial time.

## Keywords

Inconsistency. Lightweight ontologies. Partially preordered knowledge bases

## 1 Introduction

Dans cet article, nous nous intéressons à la gestion des incohérences dans des ontologies spécifiées en DL-Lite [11], une famille de fragments légers des Logiques de Description (DL) ayant de bonnes propriétés calculatoires. Dans le contexte des Logiques de Description, une base de connaissances (KB) comporte deux composantes : une base TBox qui contient la connaissance terminologique et une base ABox qui est une base assertionnelle. Il est raisonnable et d'usage de supposer que le contenu de la base TBox est correct et sans conflits, donc les éléments de la base TBox ne sont pas remis en cause en présence de conflits. Cependant, les assertions de la base ABox peuvent être discutables lorsque la base de connaissances est incohérente. Plusieurs stratégies ont été proposées pour permettre de raisonner avec des bases de connaissances incohérentes [2, 8, 12, 21]. Cela revient souvent à calculer des réparations pour la base ABox, où une réparation est définie comme un sous-ensemble maximal de la base ABox qui est cohérent avec la base TBox.

L'approche dite *ABox Repair* (AR) [16] revient à réparer la base ABox de manière minimale (en termes d'inclusion ensembliste) sans modifier la base TBox. La réponse à une

requête est basée sur les réponses obtenues de chacune des réparations. L’approche AR est souvent considérée comme un moyen fiable de gérer les conflits. Cependant, le calcul des réponses aux requêtes avec la méthode AR est coûteux, même pour une logique d’ontologie légère telle que DL-Lite. La méthode connue sous le nom *Intersection ABox Repair* (IAR) [16] est plus prudente. Elle interroge une seule sous-base cohérente de la base ABox obtenue à partir de l’intersection de toutes les réparations. L’avantage de la réparation IAR est son calcul en temps polynomial.

La notion de réparation non-contestée (en anglais *non-defeated*) d’une base ABox incohérente et totalement préordonnée a été proposée dans [4] dans le cadre des logiques DL-Lite et dans [6] dans le cadre de la logique propositionnelle. L’approche suppose que la base ABox est stratifiée du fait de l’application d’un préordre total sur les assertions. Intuitivement, la réparation non-contestée est basée sur l’application itérative de la sémantique IAR à des ensembles formés par un nombre de strates incrémenté à chaque itération. Ce calcul est également effectué en temps polynomial [19] pour DL-Lite.

Dans cet article, nous nous intéressons à la recherche d’un moyen efficace pour calculer les réparations d’une base de connaissances incohérente spécifiée en DL-Lite, et dans laquelle la relation de priorité sur les assertions est un préordre partiel. Pour ce faire, nous proposons une méthode efficace et saine, appelée *Elect*, pour restaurer la cohérence de la base ABox par rapport à la base TBox. Nous montrons que la méthode *Elect* généralise à la fois la sémantique IAR lorsque les bases assertionnelles sont non-ordonnées et aussi la sémantique “non-contestée” pour des bases ABox totalement préordonnées.

La justification sémantique de la méthode *Elect* consiste d’abord à interpréter un préordre partiel associé à une base ABox comme une famille de préordres totaux, puis d’appliquer dans un deuxième temps la sémantique “non-contestée” à chacune des bases ABox totalement préordonnées, et finalement de calculer leur intersection pour produire une réparation unique. La méthode *Elect* est saine puisqu’elle ne fait pas de choix arbitraire entre les préordres totaux associés à un préordre partiel. De ce fait, tous les préordres totaux sont pris en compte pour définir la méthode *Elect*. Nous proposons le concept d’assertion élue comme étant une assertion strictement préférée à toute autre assertion avec laquelle elle est en conflit. Cela nous permet de proposer une caractérisation équivalente de la méthode *Elect*, et donc d’obtenir la réparation sans avoir à calculer explicitement l’ensemble des préordres totaux. Enfin, nous montrons que le calcul de la réparation avec la méthode *Elect* se fait en temps polynomial. Ainsi, *Elect* préserve la calculabilité de la sémantique IAR et de la sémantique “non-contestée” pour des bases ABox partiellement préordonnées.

Cet article est structuré comme suit. La section 2 contient des rappels sur la logique DL-Lite. La section 3 présente la sémantique IAR pour des bases ABox non-ordonnées. La

section 4 aborde la réparation non-contestée pour des bases ABox totalement préordonnées. La section 5 concerne les bases ABox partiellement préordonnées. Nous introduisons notre méthode appelée *Elect* et en proposons une caractérisation. La section 6 contient une discussion sur trois extensions possibles de la méthode *Elect*, avant de conclure l’article dans la section 7.

## 2 Logique de Description DL-Lite

Les Logiques de Description (DL) [1] sont une famille de formalismes de représentation des connaissances ayant de nombreuses applications, notamment pour la formalisation d’ontologies. Les fragments légers de DL tels que DL-Lite [11] sont particulièrement intéressants car ils offrent un bon compromis entre le pouvoir expressif et la complexité calculatoire. En effet, répondre à des requêtes en DL-Lite s’effectue de manière efficace. Il existe quelques variantes de DL-Lite, telles que DL-Lite<sub>R</sub> que nous utilisons dans le présent article.

Le langage DL-Lite<sub>R</sub> est construit sur un ensemble fini de *noms de concepts* C, un ensemble fini de *noms de rôles* R et un ensemble fini de *noms d’individus* I, tels que C, R et I sont mutuellement disjoints. Le langage DL-Lite<sub>R</sub> est défini selon les règles suivantes :  $R \rightarrow P \mid P^-; E \rightarrow R \mid \neg R; B \rightarrow A \mid \exists R; C \rightarrow B \mid \neg B$ .

Dans ces règles,  $A \in C$  dénote un nom de concept,  $P \in R$  dénote un nom de rôle, et  $P^-$  est la relation *inverse* associée à  $P$ . De plus,  $R$  dénote un *rôle de base*, alors que  $E$  représente un *rôle complexe*. Par ailleurs,  $B$  dénote un *concept de base* et  $C$  est un *concept complexe*.

**Exemple 1** *Considérons l’exemple suivant :*

- $C = \{Danse, DM, DT, DA, DS, Acc\}$ , représentant respectivement les noms de concept : *danse, danse moderne, danse traditionnelle, danse avec des accessoires, danse sans accessoires ainsi que les accessoires utilisés dans certaines danses.*

- $R = \{UA\}$ , représentant les accessoires utilisés dans certaines danses et qui peuvent être des fleurs (fl), un chapeau (ch), ou des écharpes (ec).

- $I = \{d_1, d_2, d_3, d_4, d_5\} \cup \{fl, ch, ec\}$ , où chaque  $d_i$ , pour  $i = 1, \dots, 5$ , représente une danse et le reste représente des accessoires.

Des concepts complexes sont :  $\neg DA$  et  $\neg \exists UA$ . □

Un *axiome d’inclusion* sur des concepts (respectivement des rôles) est un énoncé de la forme  $B \sqsubseteq C$  (respectivement,  $R \sqsubseteq E$ ). Une inclusion de concept ayant le symbole  $\neg$  du côté droit de l’inclusion est appelée *axiome négatif d’inclusion*, sinon elle est appelée *axiome positif d’inclusion*. Des exemples d’axiomes d’inclusion sur des concepts sont :  $DS \sqsubseteq \neg DA$  et  $\exists UA^- \sqsubseteq Acc$ .

Une base TBox  $\mathcal{T}$  en DL-Lite<sub>R</sub> est un ensemble fini d’axiomes d’inclusion (positifs et négatifs). Une *assertion* est un énoncé de la forme  $A(a)$  ou  $P(a, b)$ , avec  $a, b \in I$ . Des exemples d’assertions sont  $DM(d_1)$  et  $UA(d_3, ch)$ . Une base ABox  $\mathcal{A}$  en DL-Lite<sub>R</sub> est un ensemble fini d’as-

sertions. Pour des bases  $\mathcal{T}$  et  $\mathcal{A}$ , nous notons une base de connaissances (KB) par  $\mathcal{K} = \langle \mathcal{T}, \mathcal{A} \rangle$ .

Dans cet article, nous utilisons l'exemple suivant.

**Exemple 1** (suite) Soit la base TBox suivante :

$$\mathcal{T} = \left\{ \begin{array}{ll} 1. DM \sqsubseteq \text{Danse} & 2. DT \sqsubseteq \text{Danse} \\ 3. DT \sqsubseteq DA & 4. DM \sqsubseteq DS \\ 5. DS \sqsubseteq \neg DA & 6. DS \sqsubseteq \neg \exists UA \\ 7. \exists UA^- \sqsubseteq \text{Acc} & 8. \exists UA \sqsubseteq DA \end{array} \right\}$$

Les deux premiers axiomes indiquent seulement que les danses modernes (DM) et les danses traditionnelles (DT) sont des danses. L'axiome 3 signifie que les danses traditionnelles sont des danses qui utilisent des accessoires (DA). L'axiome 4 énonce que les danses modernes n'utilisent pas d'accessoires (DS). L'axiome 5 indique que l'ensemble de danses modernes et l'ensemble de danses traditionnelles sont disjoints. L'axiome 6 énonce qu'une danse moderne n'utilise pas d'accessoires. L'axiome 7 indique que les éléments utilisés par les danses, donnés par le nom de rôle "UA", doivent appartenir à l'ensemble des éléments spécifiés par le nom de concept "Acc". L'axiome 8 spécifie que tout ce qui utilise des accessoires doit être une danse avec accessoires.

Soit la base ABox donnée par les assertions suivantes :

$$\mathcal{A} = \left\{ \begin{array}{l} DM(d_1), DM(d_2), DT(d_2), DT(d_3), \\ DT(d_4), DA(d_3), DA(d_5), DS(d_5), \\ UA(d_2, fl), UA(d_3, ch), UA(d_4, ec) \end{array} \right\} \square$$

Une base de connaissances  $\mathcal{K}$  est dite *cohérente* si elle admet au moins un modèle, sinon elle est *incohérente*. Une base TBox  $\mathcal{T}$  est *incohérente* s'il existe un nom de concept  $A \in \mathcal{C}$  tel que  $A$  est vide dans chaque modèle de  $\mathcal{T}$ , sinon elle est *cohérente*. Notons que la base KB donnée par  $\mathcal{K} = \langle \mathcal{T}, \mathcal{A} \rangle$  dans l'Exemple 1 est incohérente.

Pour de plus amples détails sur la famille DL-Lite de Logiques de Description, nous invitons le lecteur à consulter les travaux de Calvanese et al. [11]. Dans la suite de cet article, nous notons DL-Lite<sub>R</sub> simplement par DL-Lite.

### 3 Sémantique IAR pour des Bases Assertionnelles Non-Ordonnées

Dans cette section, nous considérons une base KB spécifiée en DL-Lite  $\mathcal{K} = \langle \mathcal{T}, \mathcal{A} \rangle$  qui peut être incohérente. Nous supposons que la base TBox  $\mathcal{T}$  est cohérente et fiable (c-à-d. validée par les concepteurs de l'ontologie). De ce fait, les éléments de  $\mathcal{T}$  ne sont pas remis en cause en présence de conflits, contrairement aux assertions de  $\mathcal{A}$  qui peuvent être discutables. De plus, nous supposons que la base ABox  $\mathcal{A}$  est non-ordonnée, donc toutes les assertions ont le même degré de priorité. Une manière standard de gérer l'incohérence consiste à d'abord calculer l'ensemble des sous-ensembles maximaux cohérents de  $\mathcal{A}$ , appelés réparations, puis de les utiliser pour l'inférence (c'est-à-dire pour répondre à des requêtes). Formellement, une réparation est définie comme suit [16] :

**Définition 1** Soit  $\mathcal{K} = \langle \mathcal{T}, \mathcal{A} \rangle$  une base KB non-ordonnée et incohérente. Une sous-base  $\mathcal{R} \subseteq \mathcal{A}$  est une réparation si  $\langle \mathcal{T}, \mathcal{R} \rangle$  est cohérente, et  $\forall \mathcal{R}' \subseteq \mathcal{A}$  t.q.  $\mathcal{R} \subsetneq \mathcal{R}'$ ,  $\langle \mathcal{T}, \mathcal{R}' \rangle$  est incohérente. De plus, si  $\langle \mathcal{T}, \mathcal{A} \rangle$  est cohérente, alors il existe une seule réparation  $\mathcal{R} = \mathcal{A}$ .

De ce fait, lorsque  $\mathcal{K}$  est incohérente, ajouter une assertion  $f$  de  $(\mathcal{A} \setminus \mathcal{R})$  à  $\mathcal{R}$  entraîne l'incohérence de  $\langle \mathcal{T}, \mathcal{R} \cup \{f\} \rangle$ . Nous notons par  $MAR(\mathcal{A})$  (de l'anglais *Maximal Assertional-based Repair*) l'ensemble des réparations de  $\mathcal{A}$  par rapport à  $\mathcal{T}$ . Grâce à la notion de réparation, l'incohérence d'une base KB non-ordonnée peut être gérée par l'application d'un mécanisme standard de réponse aux requêtes, soit à l'ensemble des réparations (conséquence universelle ou conséquence AR [16]), soit à une seule réparation (conséquence "brave" [9]). Il est bien connu que la sémantique dite brave est aventureuse et peut entraîner des conclusions contestables, tandis que la sémantique AR est saine mais coûteuse sur le plan calculatoire.

Une alternative à cela est d'utiliser la sémantique IAR [16] qui sélectionne une seule sous-base cohérente de  $\mathcal{A}$ , dénotée  $IAR(\mathcal{A})$ . Avant d'introduire la sémantique IAR, nous présentons d'abord la notion de conflit assertionnel. Il s'agit d'un sous-ensemble d'assertions minimal incohérent qui contredit la base TBox. Formellement :

**Définition 2** Soit  $\mathcal{K} = \langle \mathcal{T}, \mathcal{A} \rangle$  une base KB incohérente. Une sous-base  $\mathcal{C} \subseteq \mathcal{A}$  est un conflit assertionnel dans  $\mathcal{K}$  ssi  $\langle \mathcal{T}, \mathcal{C} \rangle$  est incohérente et  $\forall f \in \mathcal{C}$ ,  $\langle \mathcal{T}, \mathcal{C} \setminus \{f\} \rangle$  est cohérente.

Nous notons par  $\mathcal{C}(\mathcal{A})$  l'ensemble des conflits de  $\mathcal{A}$ . D'après la Définition 2, nous constatons qu'enlever une assertion  $f$  de  $\mathcal{C}$  restaure la cohérence de  $\langle \mathcal{T}, \mathcal{C} \rangle$ . Une propriété intéressante de DL-Lite est que le calcul de l'ensemble des conflits se fait en temps polynomial [12]. Par ailleurs, un conflit  $\mathcal{C}$  concerne exactement deux assertions [12]. Dans ce cas, si  $f$  et  $g$  sont deux assertions qui appartiennent à un conflit, nous notons le conflit par une paire  $\{f, g\}$  et nous disons que  $f$  et  $g$  sont en conflit. Nous introduisons à présent la notion d'éléments libres.

**Définition 3** Soit  $\mathcal{K} = \langle \mathcal{T}, \mathcal{A} \rangle$  une base KB incohérente. Une assertion  $f \in \mathcal{A}$  est dite libre ssi  $\forall \mathcal{C} \in \mathcal{C}(\mathcal{A}) : f \notin \mathcal{C}$ .

Intuitivement, les assertions libres correspondent aux éléments qui n'apparaissent dans aucun conflit. A l'origine, la notion d'éléments libres avait été proposée dans le cadre de la logique propositionnelle [5].

Dans la suite de cet article, nous notons par  $IAR(\mathcal{A})$  (de l'anglais *Intersection ABox Repair*) l'ensemble des éléments libres de  $\mathcal{A}$ . Formellement :

**Définition 4**  $IAR(\mathcal{A}) = \{f : f \in \mathcal{A} \text{ t.q. } f \text{ est libre}\}$ .

La Définition 4 est une réécriture équivalente de la définition standard de  $IAR(\mathcal{A})$  donnée par l'intersection de toutes les réparations :  $IAR(\mathcal{A}) = \bigcap \{\mathcal{R} \mid \mathcal{R} \in MAR(\mathcal{A})\}$  [5, 16].

Répondre à des requêtes avec la sémantique IAR revient à appliquer un mécanisme standard de réponse aux requêtes à la base  $\langle \mathcal{T}, IAR(\mathcal{A}) \rangle$  (puisque'elle est cohérente).

**Exemple 2** L'ensemble des conflits dans  $\langle \mathcal{T}, \mathcal{A} \rangle$  est :

$$\mathcal{C}(\mathcal{A}) = \{\{DM(d_2), DT(d_2)\}, \{DA(d_5), DS(d_5)\}, \{DM(d_2), UA(d_2, fl)\}\}.$$

Afin d'obtenir  $IAR(\mathcal{A})$ , il suffit d'enlever de  $\mathcal{A}$  toutes les assertions de  $\mathcal{C}(\mathcal{A})$ . Cela donne :

$$IAR(\mathcal{A}) = \{DM(d_1), DT(d_3), DT(d_4), DA(d_3), UA(d_3, ch), UA(d_4, ec)\}.$$

## 4 Réparation Non-Contestée pour des Bases Assertionnelles Totalement Préordonnées

Dans cette section, nous considérons des bases KB spécifiées en DL-Lite, incohérentes et totalement préordonnées. Donc une relation de préordre total  $\geq$  est appliquée seulement à la composante ABox que nous notons par  $(\mathcal{A}, \geq)$ . La relation  $\geq$  est réflexive, transitive et vérifie :  $\forall f \in \mathcal{A}, \forall g \in \mathcal{A}$ , soit  $f \geq g$  ou bien  $g \geq f$ . Soient  $>$  la relation de préférence stricte et  $\equiv$  la relation d'équivalence associées à la relation  $\geq$ . De plus, nous représentons  $(\mathcal{A}, \geq)$  par la partition bien-ordonnée de  $\mathcal{A}$  induite par  $\geq$ . Ainsi, étant donné  $(\mathcal{A}, \geq)$ , nous considérons que  $\mathcal{A}$  est partitionnée en  $n$  strates de la forme  $\mathcal{A} = (\mathcal{S}_1, \dots, \mathcal{S}_n)$ , où :

- $\mathcal{S}_1 = \{f \in \mathcal{A} : \forall g \in \mathcal{A}, f \geq g\}$ , et
- $\mathcal{S}_i = \{f \in \mathcal{A} : \forall g \in \mathcal{A} \setminus (\mathcal{S}_1 \cup \dots \cup \mathcal{S}_{i-1}), f \geq g\}$ , pour  $i = 2, \dots, n$ .

En d'autres termes, les assertions d'une strate  $\mathcal{S}_i$  ont le même degré de priorité  $i$ , et elles sont plus fiables que celles d'une strate  $\mathcal{S}_j$  pour  $j > i$ . Donc  $\mathcal{S}_1$  contient les assertions les plus importantes, alors que  $\mathcal{S}_n$  contient les assertions les moins importantes, et  $\mathcal{A} = \mathcal{S}_1 \cup \dots \cup \mathcal{S}_n$ .

Plusieurs travaux prennent en compte la notion de priorité pour interroger des bases de données incohérentes (comme [17, 18]) ou des bases de connaissances DL (comme [8, 14]). La plupart de ces formalismes se basent sur les notions de réparation et de sémantique AR, ils sont donc coûteux sur le plan calculatoire. En particulier, la sémantique "des réparations préférées" a été introduite dans [8] (dans l'esprit de ce qui avait été proposé en logique propositionnelle pondérée [10]). Cette sémantique revisite les sémantiques AR et IAR en remplaçant la notion de réparation par différents types de réparations préférées basées sur : le cardinal d'un ensemble, des degrés de priorité sur la base ABox et des poids sur les assertions. Cependant, ce formalisme entraîne souvent une augmentation de la complexité pour les sémantiques proposées. Notamment, la complexité de la sémantique IAR qui est polynomiale dans un cadre sans priorités explose lorsqu'un préordre total est appliqué à la base ABox.

Dans [4], une attention particulière a été accordée aux approches qui sélectionnent une seule réparation. L'une de ces approches est la réparation non-contestée qui a une

complexité polynomiale sans être aventureuse (c'est-à-dire qui génère des conclusions contestables). Intuitivement, la réparation non-contestée consiste à récupérer itérativement, strate par strate, l'ensemble des assertions libres.

**Définition 5** Soit  $\mathcal{K}$  une base KB incohérente dont la base ABox  $(\mathcal{A}, \geq)$  est totalement préordonnée. Soit  $\mathcal{A} = (\mathcal{S}_1, \dots, \mathcal{S}_n)$  la partition bien-ordonnée associée à  $\geq$ . La réparation non-contestée  $nd(\mathcal{A}, \geq) = \mathcal{S}'_1 \cup \dots \cup \mathcal{S}'_n$  t.q. :

$$\forall i = 1, \dots, n : \mathcal{S}'_i = IAR(\mathcal{S}_1 \cup \dots \cup \mathcal{S}_i)$$

où  $\forall i : IAR(\mathcal{S}_1 \cup \dots \cup \mathcal{S}_i)$  est la base IAR de  $(\mathcal{S}_1 \cup \dots \cup \mathcal{S}_i)$ , donnée par la Définition 4.

Notons qu'une réécriture (similaire à celle de  $IAR(\mathcal{A})$ ) est donnée pour  $nd(\mathcal{A}, \geq)$  dans [3]. En fait, une assertion  $f \in \mathcal{S}_i$  est dite contestée s'il existe une assertion  $g \in \mathcal{S}_j$  telle que  $j \leq i$ , et  $\{f, g\}$  est un conflit. Il a été démontré que  $nd(\mathcal{A}, \geq)$  contient toutes les assertions non-contestées [3].

**Exemple 3** Reprenons notre exemple et considérons un préordre total  $\geq$  sur les assertions de la base ABox représenté par la Figure 1, où :  $f \equiv g$  signifie que les deux assertions ont le même degré de priorité, et la flèche  $f \rightarrow g$  signifie que  $f$  est plus prioritaire que  $g$  (c-à-d.  $f > g$ ).

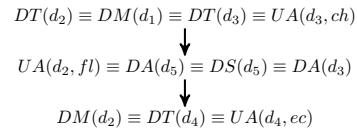


FIGURE 1 – Un préordre total sur la base ABox

A partir de cette ABox totalement préordonnée, il est possible de calculer la sous-classe non-contestée de  $\mathcal{A}$ . La partition bien-ordonnée correspondante est :

- $\mathcal{S}_1 = \{DM(d_1), DT(d_2), DT(d_3), UA(d_3, ch)\}$ .
- $\mathcal{S}_2 = \{UA(d_2, fl), DA(d_3), DA(d_5), DS(d_5)\}$ .
- $\mathcal{S}_3 = \{DM(d_2), DT(d_4), UA(d_4, ec)\}$ .

Nous avons :

$nd(\mathcal{A}, \geq) = IAR(\mathcal{S}_1) \cup IAR(\mathcal{S}_1 \cup \mathcal{S}_2) \cup IAR(\mathcal{S}_1 \cup \mathcal{S}_2 \cup \mathcal{S}_3)$ , tel que :

- $IAR(\mathcal{S}_1) = \{DT(d_2), DM(d_1), DT(d_3), UA(d_3, ch)\}$ .
- $IAR(\mathcal{S}_1 \cup \mathcal{S}_2) = \{DT(d_2), DM(d_1), DT(d_3), UA(d_3, ch), UA(d_2, fl), DA(d_3)\}$ .
- $IAR(\mathcal{S}_1 \cup \mathcal{S}_2 \cup \mathcal{S}_3) = IAR(\mathcal{A})$  (donné dans l'exemple 2).

Par conséquent :

$$nd(\mathcal{A}, \geq) = \{DM(d_1), DT(d_2), DT(d_3), DT(d_4), DA(d_3), UA(d_2, fl), UA(d_3, ch), UA(d_4, ec)\}.$$

## 5 Bases Assertionnelles Partiellement Préordonnées

La sémantique IAR (ABox non-ordonnée) et la sémantique "non-contestée" (ABox totalement préordonnée) sont connues pour leur efficacité pour gérer l'incohérence. En



FIGURE 2 – Un préordre partiel sur la base ABox

effet, elles calculent une réparation unique pour la base ABox et le font en temps polynomial. Dans cette section, nous proposons une méthode pour calculer une réparation unique lorsqu'un préordre partiel noté  $\trianglerighteq$  est appliqué aux assertions de la base ABox notée  $(\mathcal{A}, \trianglerighteq)$ . Soient  $\triangleright$  l'ordre strict (irréflexif et transitif) et  $\equiv$  la relation d'équivalence associés à la relation  $\trianglerighteq$ .

Nous appelons notre méthode "Elect" et notons par  $Elect(\mathcal{A}, \trianglerighteq)$  la réparation qu'elle produit. Comme nous le verrons par la suite, la méthode Elect généralise la sémantique IAR dans le cas où la relation  $\trianglerighteq$  est non-ordonnée, et aussi la sémantique "non-contestée" lorsque la relation  $\trianglerighteq$  est totalement préordonnée.

### 5.1 Une famille de préordres totaux à partir d'un préordre partiel

Dans cette section, nous considérons qu'un préordre partiel  $\trianglerighteq$  est interprété comme une famille de préordres totaux, de sorte que chaque préordre total est une extension totale de la relation  $\trianglerighteq$  définie comme suit :

**Définition 6** *Un préordre total  $\geq$  sur  $\mathcal{A}$  est une extension totale de  $\trianglerighteq$  sur  $\mathcal{A}$  ssi  $\forall f, g \in \mathcal{A}$ , si  $f \trianglerighteq g$  alors  $f \geq g$ .*

Notons que l'interprétation d'une base KB partiellement préordonnée comme une famille de bases KB totalement préordonnées est une représentation naturelle qui a été utilisée dans d'autres cadres tels que la logique possibiliste partiellement ordonnée [7, 20] et les réseaux probabilistes crédaux [13]. Cette démarche n'a de sens que si les éléments peuvent être comparés (le fait de dire que deux assertions  $f$  et  $g$  sont incomparables signifie que soit  $f$  est strictement préférée à  $g$ , soit  $f$  est aussi préférée que  $g$ , soit  $g$  est strictement préférée à  $f$ , mais on ignore lequel).

**Exemple 4** *Soit un préordre partiel  $\trianglerighteq$  sur les assertions de la base ABox qui permet de les répartir en quatre sous-ensembles comme suit :*

- $A = \{DM(d_1) \equiv DT(d_2) \equiv DT(d_3) \equiv UA(d_3, ch)\}$
- $B = \{UA(d_2, fl) \equiv DA(d_3) \equiv DS(d_5) \equiv DA(d_5)\}$
- $C = \{DM(d_2)\}$
- $D = \{DT(d_4) \equiv UA(d_4, ec)\}$

où  $f \equiv g$  signifie que les assertions  $f$  et  $g$  ont le même degré de priorité.

Ce préordre partiel est représenté par la Figure 2, où la flèche  $A \rightarrow B$  (par exemple) signifie  $\forall f \in A, \forall g \in B$ , l'assertion  $f$  est plus prioritaire que  $g$  (c-à-d.  $f \triangleright g$ ).

De ce fait, l'ensemble  $A$  (resp.  $D$ ) contient les assertions les plus (resp. les moins) prioritaires. Les assertions des

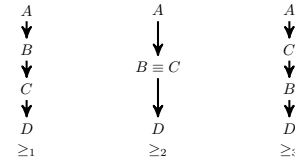


FIGURE 3 – Les extensions totales du préordre partiel  $\trianglerighteq$

ensembles  $B$  et  $C$  sont incomparables.

Il s'en suit que le préordre partiel  $\trianglerighteq$  peut être interprété comme une famille de trois préordres totaux  $\geq_1, \geq_2$  et  $\geq_3$  représentés par la Figure 3, de sorte que :

- Selon  $\geq_1$ , les assertions de l'ensemble  $B$  sont strictement préférées à celles de l'ensemble  $C$ .
- Selon  $\geq_2$ , les assertions des ensembles  $B$  et  $C$  sont préférées de manière égale.
- Selon  $\geq_3$ , les assertions de l'ensemble  $C$  sont strictement préférées à celles de l'ensemble  $B$ .  $\square$

Une question qui se pose est comment gérer cette famille de ABox totalement préordonnées. Nous voulons éviter le choix arbitraire qui consiste à sélectionner un seul préordre total parmi tous les autres. Il faudrait donc prendre en compte tous les préordres totaux. Une méthode prudente pour obtenir une sous-base assertionnelle cohérente unique est de considérer l'intersection de toutes les réparations non-contestées associées à tous les préordres totaux.

**Définition 7** *Soit  $\mathcal{K}$  une base KB spécifiée en DL-Lite, incohérente et dont la base ABox  $(\mathcal{A}, \trianglerighteq)$  est partiellement préordonnée.*

- $Elect(\mathcal{A}, \trianglerighteq) = \bigcap_{\geq} \{nd(\mathcal{A}, \geq) \text{ t.q. } \geq \text{ est une extension totale de } \trianglerighteq\}$ , où  $nd(\mathcal{A}, \geq)$  est donné par la Définition 5.
- Soit  $q$  une requête. Alors  $q$  est une conséquence de  $\mathcal{K}$  par la méthode Elect si  $q$  s'ensuit de  $Elect(\mathcal{A}, \trianglerighteq)$  (en utilisant l'inférence standard de DL-lite).

Nous illustrons cette idée sur notre exemple.

**Exemple 5** *Les réparations non-contestées des ABox totalement préordonnées  $(\mathcal{A}, \geq_1)$ ,  $(\mathcal{A}, \geq_2)$  et  $(\mathcal{A}, \geq_3)$  sont :*

- $nd(\mathcal{A}, \geq_1) = A \cup \{UA(d_2, fl), DA(d_3)\} \cup D.$
- $nd(\mathcal{A}, \geq_2) = A \cup \{DA(d_3)\} \cup D.$
- $nd(\mathcal{A}, \geq_3) = A \cup \{DA(d_3)\} \cup D.$

Par conséquent :

$$Elect(\mathcal{A}, \trianglerighteq) = nd(\mathcal{A}, \geq_1) \cap nd(\mathcal{A}, \geq_2) \cap nd(\mathcal{A}, \geq_3) = A \cup \{DA(d_3)\} \cup D. \quad \square$$

Dans ce qui suit, nous étudions quelques caractéristiques de notre méthode.

Un résultat important énoncé dans la Proposition 1 est que le calcul de la réparation  $Elect(\mathcal{A}, \trianglerighteq)$  peut s'effectuer en temps polynomial. En effet, il n'est pas nécessaire d'exhiber toutes les extensions totales de la relation  $\trianglerighteq$ .

**Proposition 1** *Le calcul de la réparation  $Elect(\mathcal{A}, \trianglerighteq)$  se fait en temps polynomial (par rapport à la taille de  $\mathcal{A}$ ).*

La proposition suivante établit qu'une base KB ayant  $Elect(\mathcal{A}, \triangleright)$  comme base ABox est cohérente.

**Proposition 2**  $\langle \mathcal{T}, Elect(\mathcal{A}, \triangleright) \rangle$  est cohérente.

Une autre caractéristique de la méthode Elect est qu'elle correspond à la sémantique IAR lorsque la base ABox est non-ordonnée, et à la sémantique "non-contestée" lorsque la base ABox est totalement préordonnée.

**Proposition 3** Si le préordre partiel  $\triangleright$  est non-ordonné, alors  $Elect(\mathcal{A}, \triangleright) = IAR(\mathcal{A})$ . Si le préordre partiel  $\triangleright$  est un préordre total, alors  $Elect(\mathcal{A}, \triangleright) = nd(\mathcal{A}, \triangleright)$ .

Les preuves des Propositions 1 à 3 sont établies grâce à la définition d'une caractérisation équivalente de la réparation  $Elect(\mathcal{A}, \triangleright)$ , présentée dans la prochaine section.

## 5.2 Caractérisation de $Elect(\mathcal{A}, \triangleright)$

Dans cette section, nous proposons une caractérisation de la réparation  $Elect(\mathcal{A}, \triangleright)$  qui permet d'éviter de calculer l'ensemble des extensions totales de la relation  $\triangleright$ . Pour cela, nous introduisons le concept d'assertions élues. Intuitivement, une assertion est élue dans  $(\mathcal{A}, \triangleright)$  si elle est strictement préférée à toutes les assertions avec lesquelles elle est en conflit. Formellement :

**Définition 8** Une assertion  $f \in \mathcal{A}$  est élue ssi  $\forall g \in \mathcal{A}$ , si  $\{f, g\}$  est un conflit, alors  $f \triangleright g$  (c-à-d.  $f$  est strictement préférée à  $g$ ).

La Définition 8 généralise le concept d'assertions libres donné dans la Définition 3. En effet, si la relation  $\triangleright$  est non-prioritaire (c-à-d.  $\forall f \in \mathcal{A}, \forall g \in \mathcal{A}, f \triangleright g$  et  $g \triangleright f$ ), alors  $f$  est élue dans  $(\mathcal{A}, \triangleright)$  ssi  $f$  est libre. Evidemment, le contraire n'est en général pas vrai (lorsque la relation  $\triangleright$  est prioritaire). En effet, une assertion élue peut ne pas être une assertion libre, mais son degré de fiabilité est strictement plus important que celui de ses opposants. Cette définition généralise également la notion d'assertions acceptées donnée pour des réparations non-contestées dans des bases KB totalement préordonnées [3]. Enfin, le concept d'assertions élues est dans le même esprit que la notion de croyances acceptées proposée en théorie de l'incertitude [15].

La Proposition 4 énonce que les assertions élues sont exactement celles de la réparation  $Elect(\mathcal{A}, \triangleright)$ .

**Proposition 4** Une assertion  $f \in \mathcal{A}$  est élue dans  $(\mathcal{A}, \triangleright)$  ssi  $f \in Elect(\mathcal{A}, \triangleright)$ .

Nous démontrons cette proposition comme suit.

**Preuve 1** Soit  $(\mathcal{A}, \triangleright)$  une base assertionnelle partiellement préordonnée.

1. Soit  $f \in \mathcal{A}$  une assertion élue. Montrons que pour chaque extension totale  $(\mathcal{A}, \geq)$  de  $(\mathcal{A}, \triangleright)$ , nous avons  $f \in nd(\mathcal{A}, \geq)$ . Soit  $(\mathcal{S}_1, \dots, \mathcal{S}_n)$  la partition bien-ordonnée associée à  $\geq$ . Soit  $i$  la première strate où  $f \in \mathcal{S}_i$ . Rappelons que  $f$  est élue dans  $(\mathcal{A}, \triangleright)$  signifie  $\forall g \in \mathcal{A}$ ,

si  $\{f, g\}$  est un conflit, alors  $f \triangleright g$  (c-à-d.  $f$  est strictement préférée à  $g$  selon  $\triangleright$ ). Et puisque  $\geq$  est une extension totale de  $\triangleright$ , alors cela signifie que  $f > g$ . Cela signifie aussi  $\forall g \in \mathcal{A}$  tel que  $\{f, g\}$  est un conflit,  $g \in \mathcal{S}_j$  avec  $j > i$ . Donc,  $f \in IAR(\mathcal{S}_1 \cup \dots \cup \mathcal{S}_i)$ . Il s'en suit que  $f \in nd(\mathcal{A}, \geq)$ .

2. Montrons à présent l'inverse. Supposons que  $f \in \mathcal{A}$  n'est pas élue et construisons une extension totale  $(\mathcal{A}, \geq)$  de  $(\mathcal{A}, \triangleright)$  telle que  $f \notin nd(\mathcal{A}, \geq)$ .

L'assertion  $f$  n'est pas élue signifie  $\exists g \in \mathcal{A}$  tel que  $\{f, g\}$  est un conflit et  $f \triangleright g$  n'est pas vrai. Donc il existe une extension totale  $\geq$  de  $\triangleright$  où  $g \geq f$ . Si  $\{f, g\}$  est un conflit et  $(\mathcal{S}_1, \dots, \mathcal{S}_n)$  est la partition bien-ordonnée associée à  $\geq$ , alors si  $f \in \mathcal{S}_i$ , il s'en suit que  $g \in \mathcal{S}_j$  où  $j \leq i$ . Par conséquent,  $\forall k \in \{1, \dots, n\}$ ,  $f \notin IAR(\mathcal{S}_1 \cup \dots \cup \mathcal{S}_k)$ , donc  $f \notin nd(\mathcal{A}, \geq)$ .  $\square$

A partir de ce résultat, nous pouvons démontrer les Propositions 1, 2 et 3 données dans la section précédente.

### Preuve 2

1. En ce qui concerne la complexité calculatoire, nous rappelons que le calcul de l'ensemble des conflits  $\mathcal{C}(\mathcal{A})$  se fait en temps polynomial par rapport à la taille de  $\mathcal{A}$ . Donc, calculer  $Elect(\mathcal{A}, \triangleright)$  se fait aussi en temps polynomial. En effet, vérifier si une assertion  $f \in \mathcal{A}$  est élue revient à parcourir tous les conflits assertionnels dans  $\mathcal{C}(\mathcal{A})$ . Ceci se fait en temps linéaire par rapport à la taille de  $\mathcal{C}(\mathcal{A})$  (cette taille étant bornée par  $\mathcal{O}(|\mathcal{A}|^2)$ ).

2. Montrons que la base  $Elect(\mathcal{A}, \triangleright)$  est cohérente par rapport à  $\mathcal{T}$ . Supposons que ce n'est pas le cas. Donc  $\exists f \in Elect(\mathcal{A}, \triangleright), \exists g \in Elect(\mathcal{A}, \triangleright), g \neq f$ , tel que  $\{f, g\}$  est un conflit. Puisque  $f$  et  $g$  sont dans  $Elect(\mathcal{A}, \triangleright)$ , alors cela signifie que  $f \triangleright g$  et  $g \triangleright f$ , ce qui est impossible.

3. Finalement, par construction de la base  $Elect(\mathcal{A}, \triangleright)$ , il est facile de vérifier que lorsque  $\triangleright$  est un préordre total, alors  $Elect(\mathcal{A}, \triangleright)$  est la réparation non-contestée de  $\triangleright$ . Et lorsque  $\triangleright$  est non-ordonnée (donc  $\forall f \in \mathcal{A}, \forall g \in \mathcal{A}, f \triangleright g$  et  $g \triangleright f$ ), alors  $Elect(\mathcal{A}, \triangleright) = IAR(\mathcal{A}) = \{f \in \mathcal{A} : \nexists g \in \mathcal{A}, t.q. \{f, g\} \text{ est un conflit}\}$ .  $\square$

**Exemple 6** Nous utilisons la notion d'assertions élues de la Définition 8 pour recalculer la réparation  $Elect(\mathcal{A}, \triangleright)$ . Il est aisé de vérifier que :

- L'assertion  $DM(d_2)$  n'est pas élue car  $\{DM(d_2), DT(d_2)\}$  est un conflit et la préférence stricte  $DM(d_2) \triangleright DT(d_2)$  n'est pas vérifiée.
- Les assertions  $DS(d_5)$  et  $DA(d_5)$  ne sont pas élues car elles sont en conflit et ont le même degré de priorité.
- L'assertion  $UA(d_2, fl)$  n'est pas élue car  $\{DM(d_2), UA(d_2, fl)\}$  est un conflit et la préférence stricte  $UA(d_2, fl) \triangleright DM(d_2)$  n'a pas lieu.
- Les assertions restantes sont toutes élues. A savoir :

$Elect(\mathcal{A}, \triangleright) = \{DM(d_1), DT(d_2), DT(d_3), DT(d_4), DA(d_3), UA(d_3, ch), UA(d_4, ec)\}$ .

Ce résultat correspond à celui de l'Exemple 5 en considérant toutes les extensions totales du préordre partiel  $\triangleright$ .



## 6 Extensions de la Méthode Elect

Dans cette section, nous proposons une discussion brève de trois extensions possibles de la méthode Elect. Leur étude approfondie fera l'objet de travaux futurs.

### 6.1 Au-delà de la méthode Elect

Une question qui se pose est comment obtenir une base plus grande (plus productive) qu'une base calculée par la méthode Elect, tout en produisant une réparation saine. Une solution immédiate est d'utiliser la notion de fermeture déductive positive, selon laquelle la fermeture de la base ABox est définie en termes des axiomes positifs de la base TBox. Il y a deux façons d'appliquer la fermeture déductive positive, soit sur la base ABox initiale, soit sur l'ensemble des réparations non-contestées calculées pour toutes les extensions totales du préordre partiel.

Dans la première option, l'application de la fermeture positive à la base ABox initiale, dans l'esprit de la sémantique ICAR pour des bases ABox non-ordonnées [16], soulève deux questions. Premièrement, en termes de sémantique, l'approche ICAR peut ne pas être fiable puisqu'elle permet de dériver des conséquences à partir d'assertions discutables. Deuxièmement, il peut y avoir plusieurs façons de définir la fiabilité des éléments dérivés. Par exemple, supposons que la base TBox contient  $\{A \sqsubseteq B, E \sqsubseteq B\}$  et que la base ABox contient  $\{A(x), E(x)\}$ . Supposons que  $A(x)$  et  $E(x)$  sont incomparables. Donc  $B(x)$  peut être dérivée de  $A(x)$  mais aussi de  $E(x)$ . La question est alors où positionner  $B(x)$ . L'intuition est de considérer  $B(x)$  comme *au moins aussi plausible que*  $A(x)$  et  $E(x)$ , mais cela n'est pas évident à définir de manière générale (en particulier pour des logiques de description expressives).

La deuxième option consiste à définir la fermeture de toutes les réparations non-contestées, ce qui est plus sain. En effet, les conclusions supplémentaires sont obtenues uniquement d'assertions appartenant à des réparations non-contestées, qui sont connues pour ne contenir que des assertions saines. Ainsi, la réparation associée à un préordre partiel est calculée comme l'intersection des réparations non-contestées fermées [4], une méthode que nous appelons CElect. Une base calculée avec la méthode CElect est alors plus grande qu'une base calculée avec la méthode Elect. De plus, pour des bases ABox non-ordonnées, une réparation calculée avec la méthode CElect est équivalente à la fermeture des réparations IAR (ce qui est différent d'une réparation calculée avec la méthode ICAR). Et pour des bases ABox totalement préordonnées, une réparation calculée avec la méthode CElect est équivalente à la fermeture des réparations non-contestées. Dans des travaux futurs, nous envisageons d'exhiber des cas spécifiques où la complexité de la méthode CElect serait également polynomiale en DL-Lite.

### 6.2 Au-delà de la réparation non-contestée

La question traitée dans cette partie concerne la possibilité ou non d'utiliser une sémantique autre que la sémantique

“non-contestée” pour définir la méthode Elect. D'un point de vue sémantique, la réponse est positive. Par exemple, il est possible d'utiliser l'une des réparations préférées définies dans [8] au lieu de la réparation non-contestée dans notre définition de  $Elect(\mathcal{A}, \triangleright)$ . Rappelons d'abord la notion de réparation préférée pour des bases ABox totalement préordonnées.

Soit  $\mathcal{A} = (\mathcal{S}_1, \dots, \mathcal{S}_n)$  une base ABox totalement préordonnée. Soient  $\mathcal{R}_1$  et  $\mathcal{R}_2$  deux sous-bases cohérentes de  $\mathcal{S}_1 \cup \dots \cup \mathcal{S}_n$ . Une base  $\mathcal{R}_1$  est aussi préférée que  $\mathcal{R}_2$  ssi  $\forall i, 1 \leq i \leq n, \mathcal{R}_1 \cap \mathcal{S}_i = \mathcal{R}_2 \cap \mathcal{S}_i$ . Une base  $\mathcal{R}_1$  est strictement préférée à  $\mathcal{R}_2$  ssi  $\exists i, 1 \leq i \leq n$ , t.q.  $\mathcal{R}_2 \cap \mathcal{S}_i \subsetneq \mathcal{R}_1 \cap \mathcal{S}_i$  et  $\forall j, 1 \leq j < i, \mathcal{R}_1 \cap \mathcal{S}_j = \mathcal{R}_2 \cap \mathcal{S}_j$ . Donc une base  $\mathcal{R} \subseteq \mathcal{A}$  est une réparation préférée ssi  $\nexists \mathcal{R}' \subseteq \mathcal{A}$  t.q.  $\mathcal{R}'$  est strictement préférée à  $\mathcal{R}$ .

(Se référer à [8] et aussi [6] pour de plus amples détails.)

La notion de réparation préférée définie pour un préordre total peut alors servir de base pour définir une réparation en présence d'un préordre partiel  $\triangleright$ .

Nous appelons ce nouveau formalisme  $Partial_{PR}(\mathcal{A}, \triangleright)$  (où  $PR$  désigne en anglais *Preferred Repairs*). Tout comme une réparation  $Elect(\mathcal{A}, \triangleright)$ , une réparation  $Partial_{PR}(\mathcal{A}, \triangleright)$  considère toutes les extensions totales  $\geq$  du préordre partiel  $\triangleright$ . Cependant, au lieu de considérer l'intersection des réparations non-contestées comme pour calculer  $Elect(\mathcal{A}, \triangleright)$ , nous considérons l'intersection des réparations préférées, notée  $IPR(\mathcal{A}, \geq)$  (de l'anglais *Intersection of Preferred Repairs*), comme suit :  $IPR(\mathcal{A}, \geq) = \bigcap \{\mathcal{R} \mid \mathcal{R} \text{ est une réparation préférée de } \geq\}$ . Donc  $Partial_{PR}(\mathcal{A}, \triangleright) = \bigcap_{\geq} \{IPR(\mathcal{A}, \geq)\}$  t.q.  $\geq$  est une extension totale de  $\triangleright$ .

La réparation  $Partial_{PR}(\mathcal{A}, \triangleright)$  est plus grande que la réparation  $Elect(\mathcal{A}, \triangleright)$ . Cependant, la base  $Partial_{PR}(\mathcal{A}, \triangleright)$  n'est pas calculable en temps raisonnable puisque la complexité de  $IPR(\mathcal{A}, \geq)$  est coNP [8] si la relation  $\geq$  est un simple préordre total. Dans ce cas, la Proposition 1 n'est pas vérifiée.

### 6.3 Au-delà de DL-Lite

Une autre question concerne le fait de pouvoir ou non généraliser la méthode Elect à des bases ABox partiellement préordonnées exprimées dans des logiques autres que DL-Lite. D'un point de vue sémantique, il n'y a pas de limitations et les résultats obtenus correspondent également à la réparation IAR (base ABox non-ordonnée) et à la réparation non-contestée (base ABox totalement préordonnée).

Cependant, d'un point de vue calculatoire, il est impératif d'avoir un moyen efficace de gérer les conflits afin que la Proposition 1 soit vérifiée. En particulier, il n'est pas nécessaire que les conflits assertionnels  $\mathcal{C} \in \mathcal{C}(\mathcal{A})$  soient binaires (c-à-d. impliquant deux assertions), à condition qu'ils puissent être calculés en temps polynomial.

Dans le cas de conflits impliquant plus de deux assertions, nous redéfinissons la notion d'assertion élue comme étant celle qui est strictement préférée à au moins une assertion de chaque conflit où elle apparaît. Cela permettrait

également que notre caractérisation (sans calculer tous les préordres totaux) soit toujours maintenue. Par conséquent,  $Elect(\mathcal{A}, \triangleright)$  pourrait être généralisée dans des langages plus expressifs que DL-Lite à condition que le calcul des conflits puisse se faire de manière efficace.

## 7 Conclusion

Nous avons étudié le problème de la restauration de la cohérence d'une base ABox partiellement préordonnée et incohérente par rapport à la base TBox pour des ontologies DL-Lite. Nous avons proposé une méthode appelée Elect qui généralise la sémantique IAR (ABox non-ordonnée) et la sémantique dite non-contestée (ABox totalement préordonnée). Dans la méthode Elect, un préordre partiel est interprété comme une famille de préordres totaux auxquels l'inférence non-contestée est appliquée, produisant des réparations non-contestées. Nous avons introduit le concept d'assertions élues et proposé une caractérisation équivalente de la méthode Elect. En particulier, nous avons montré que la complexité de la méthode Elect est polynomiale. Dans des travaux futurs, nous envisageons d'investiguer les extensions de la méthode Elect introduites en Section 6.

**Remerciements** : Ce travail a été financé par le projet européen H2020-MSCA-RISE AniAge : ([www.cril.univ-artois.fr/aniage/](http://www.cril.univ-artois.fr/aniage/)).

## Références

- [1] F. Baader, D. Calvanese, D. McGuinness, D. Nardi, and P. Patel-Schneider. *The Description Logic Handbook : Theory, Implementation, and Applications*. 2007.
- [2] J. Baget, S. Benferhat, Z. Bouraoui, M. Croitoru, M. Mugnier, O. Papini, S. Rocher, and K. Tabia. A general modifier-based framework for inconsistency-tolerant query answering. In *KR, Le Cap, Afrique du Sud*, pages 513–516, 2016.
- [3] S. Benferhat, Z. Bouraoui, H. Chadhry, M. S. B. M. R. Fc, K. Tabia, and A. Telli. Characterizing non-defeated repairs in inconsistent lightweight ontologies. In *SITIS, Naples, Italie*, pages 282–287, 2016.
- [4] S. Benferhat, Z. Bouraoui, and K. Tabia. How to select one preferred assertional-based repair from inconsistent and prioritized DL-Lite knowledge bases? In *IJCAI, Buenos Aires, Argentine*, pages 1450–1456, 2015.
- [5] S. Benferhat, D. Dubois, and H. Prade. Representing default rules in possibilistic logic. In *Knowledge Representation and Reasoning*, pages 673–684, 1992.
- [6] S. Benferhat, D. Dubois, and H. Prade. *Some syntactic approaches to the handling of inconsistent knowledge bases : A comparative study. Part 2 : the prioritized case*, volume 24, pages 473–511. 1998.
- [7] S. Benferhat, S. Lagrue, and O. Papini. Reasoning with partially ordered information in a possibilistic logic framework. *Fuzzy Sets and Systems*, 144(1) :25–41, 2004.
- [8] M. Bienvenu, C. Bourgaux, and F. Goasdoué. Querying inconsistent description logic knowledge bases under preferred repair semantics. In *AAAI, Québec, Canada*, pages 996–1002, 2014.
- [9] M. Bienvenu and R. Rosati. Tractable approximations of consistent query answering for robust ontology-based data access. In *IJCAI, Pékin, Chine*, pages 775–781, 2013.
- [10] G. Brewka. Preferred subtheories : An extended logical framework for default reasoning. In *IJCAI, Detroit, USA*, pages 1043–1048, 1989.
- [11] D. Calvanese, G. D. Giacomo, D. Lembo, M. Lenzerini, and R. Rosati. Tractable reasoning and efficient query answering in description logics : The DL-Lite family. *Journal of Automated Reasoning*, 39(3) :385–429, 2007.
- [12] D. Calvanese, E. Kharlamov, W. Nutt, and D. Zheleznyakov. Evolution of DL-Lite knowledge bases. In *International Semantic Web Conference (1), Shanghai, Chine*, pages 112–128, 2010.
- [13] F. Cozman. Credal networks. *Artificial Intelligence Journal*, 120 :199–233, 2000.
- [14] J. Du, G. Qi, and Y. Shen. Weight-based consistent query answering over inconsistent SHIQ knowledge bases. *Knowledge and Information Systems*, 34(2) :335–371, 2013.
- [15] D. Dubois, H. Fargier, and H. Prade. Ordinal and probabilistic representations of acceptance. *Journal of Artificial Intelligence Research*, 22 :23–56, 2004.
- [16] D. Lembo, M. Lenzerini, R. Rosati, M. Ruzzi, and D. F. Savo. Inconsistency-tolerant semantics for description logics. In *Web Reasoning and Rule Systems*, volume 6333 of *LNCS*, pages 103–117, 2010.
- [17] M. V. Martinez, F. Parisi, A. Pugliese, G. I. Simari, and V. S. Subrahmanian. Inconsistency management policies. In *Knowledge Representation and Reasoning*, pages 367–377. AAAI Press, 2008.
- [18] S. Staworko, J. Chomicki, and J. Marcinkowski. Prioritized repairing and consistent query answering in relational databases. *Annals of Mathematics and Artificial Intelligence*, 64(2-3) :209–246, 2012.
- [19] A. Telli, S. Benferhat, M. Bourahla, Z. Bouraoui, and K. Tabia. Polynomial algorithms for computing a single preferred assertional-based repair. *Künstliche Intelligenz*, 31(1) :15–30, 2017.
- [20] F. Touazi, C. Cayrol, and D. Dubois. Possibilistic reasoning with partially ordered beliefs. *Journal of Applied Logic*, 13(4) :770–798, 2015.
- [21] D. Trivela, G. Stoilos, and V. Vassalos. Querying expressive DL ontologies under the ICAR semantics. In *DL workshop. Tempe, USA*, 2018.

# Apprendre à jouer aux jeux à deux joueurs à information parfaite sans connaissance

Quentin Cohen-Solal  
CRIL, Univ. Artois et CNRS  
62300 Lens, France  
cohen-solal@cril.fr

## Résumé

Dans cet article, plusieurs techniques pour l'apprentissage par renforcement de fonctions d'évaluation d'états de jeu sont proposées. La première consiste à apprendre les valeurs de l'arbre de jeu au lieu de se restreindre à la valeur de la racine. La seconde consiste à remplacer le gain classique d'un jeu (+1 / -1) par une heuristique favorisant les victoires rapides et les défaites lentes. La troisième permet de corriger certaines fonctions d'évaluation en tenant compte de la résolution des états. La quatrième est une nouvelle distribution de sélection d'actions. Enfin, la cinquième est une modification du minimax à profondeur non bornée étendant les meilleures séquences d'actions jusqu'aux états terminaux. En outre, nous proposons une autre variante du minimax non borné, qui joue l'action la plus sûre au lieu de jouer la meilleure action. Les expériences menées suggèrent que cela améliore le niveau de jeux lors des confrontations. Enfin, nous appliquons ces différentes techniques pour concevoir un programme-joueur au jeu de Hex (taille 11) atteignant le niveau de Mohex 2.0 à la suite d'un apprentissage par renforcement contre soi-même sans utilisation de connaissance.

## Mots Clef

Décision séquentielle, Jeux, Planification, Apprentissage, Renforcement, Minimax non borné.

## Abstract

In this paper, several techniques for learning game states evaluation functions by reinforcement are proposed. The first is to learn the values of the game tree instead of restricting oneself to the value of the root. The second is to replace the classic gain of a game (+1 / -1) with a heuristic favoring quick wins and slow defeats. The third corrects some evaluation functions taking into account the resolution of states. The fourth is a new action selection distribution. Finally, the fifth is a modification of the minimax with unbounded depth extending the best sequences of actions to the terminal states. In addition, we propose another variant of the unbounded minimax, which plays the safest action instead of playing the best action. The experiments conducted suggest that this improves the level

of play during confrontations. Finally, we apply these different techniques to design a program-player to the Hex game (size 11) reaching the level of Mohex 2.0 with reinforcement learning from self-play without knowledge.

## Keywords

Sequential Decision, Games, Planning, Learning, Reinforcement, Unbound Minimax.

## 1 Introduction

Une des tâches les plus difficiles en intelligence artificielle est la prise de décision séquentielle [15], dont les applications incluent la robotique et les jeux. Concernant les jeux, les succès sont nombreux. La machine dépasse l'homme pour plusieurs jeux, tels que le backgammon, les dames, les échecs et le go [27]. Une classe importante de jeux constitue les jeux à deux joueurs à information parfaite, c'est-à-dire les jeux où les joueurs jouent à tour de rôle, sans hasard ni information cachée. Il reste encore de nombreux défis pour ces jeux. Par exemple, pour le jeu de Hex, l'homme reste toujours supérieur à la machine. C'est également le cas du *general game playing* [27] (même restreint aux jeux à information parfaite) : l'homme est toujours supérieur à la machine sur un jeu inconnu (lorsque l'homme et la machine ont un temps d'apprentissage relativement court pour maîtriser les règles du jeu). Dans cet article, nous nous concentrons sur les jeux à deux joueurs à information parfaite et à somme nulle, bien que la plupart des contributions de cet article devraient pouvoir s'appliquer ou s'adapter aisément à un cadre plus général.

Les premières approches utilisées pour concevoir un programme capable de jouer à un jeu se basent sur un algorithme de recherche dans les arbres de jeu, tel que le minimax, combiné à une fonction d'évaluation des états de jeu, conçue à partir de connaissances expertes. L'avènement de cette technique est le programme Deep Blue [27] au jeu d'échecs. Cependant, le succès de Deep Blue tient en grande partie à la puissance brute de sa machine, lui permettant d'analyser deux cents millions d'états de jeu par seconde. De plus, cette approche est limitée par le fait de devoir concevoir une fonction d'évaluation manuellement (au moins partiellement). Cette conception est une tâche très complexe, qui doit, en plus, être réalisée pour chaque

jeu. Plusieurs travaux se sont donc focalisés sur l'apprentissage automatique de fonctions d'évaluation [16]. L'un des premiers succès de l'apprentissage de fonctions d'évaluation est sur le jeu Backgammon [27]. Cependant, pour de nombreux jeux, comme le Hex ou le Go, les approches basées sur le minimax, avec ou sans apprentissage automatique, n'ont pas permis de dépasser l'homme. Deux causes ont été identifiées [2]. Premièrement, le très grand nombre d'actions possibles à chaque état de jeu empêche d'effectuer une recherche exhaustive à une profondeur significative (le cours de la partie ne peut être anticipé qu'à un petit nombre de tours à l'avance). Deuxièmement, pour ces jeux, aucune fonction d'évaluation suffisamment performante n'a pu être identifiée. Une approche alternative permettant de résoudre ces deux problèmes a été proposée, donnant notamment de bons résultats au Hex et au Go, nommée recherche arborescente Monte Carlo et notée MCTS (pour Monte Carlo Tree Search) [5]. Cet algorithme explore l'arbre de jeu non uniformément, ce qui constitue une solution au problème du très grand nombre d'actions. De plus, il évalue les états de jeu à partir de statistiques de victoire provenant d'un grand nombre de simulations aléatoires de fin de partie. Il n'a ainsi pas besoin de fonction d'évaluation. Cela n'était cependant pas suffisant pour dépasser le niveau des joueurs humains. Plusieurs variantes du MCTS ont alors été proposées, utilisant notamment des connaissances pour orienter l'exploration de l'arbre de jeu et/ou les simulations aléatoires de fin de partie [5]. Les récentes améliorations du MCTS ont porté sur l'apprentissage automatique des connaissances du MCTS et leurs utilisations. Ces connaissances ont d'abord été générées par *apprentissage supervisé* [7, 8, 6] puis par *apprentissage supervisé* suivi d'un *apprentissage par renforcement* [21], et enfin par *apprentissage par renforcement* seulement [23, 22, 1]. Cela a permis d'atteindre et de dépasser le niveau de champion du monde au jeu de Go avec les dernières versions du programme Alphago [21, 23]. En particulier, Alphago zero [23], qui n'utilise que l'apprentissage par renforcement, n'a donc eut besoin d'aucune connaissance initiale, humaine, pour atteindre son niveau de jeu. Ce dernier succès a toutefois nécessité 29 millions de parties, ce qui est très largement supérieur à ce qu'il a fallu au champion humain du Go pour atteindre son niveau de jeu. Cette approche a également été appliquée aux échecs [22]. Le programme résultant a battu le meilleur programme des échecs, qui était toujours basé sur le minimax.

On peut dès lors se demander si les approches de type minimax sont totalement dépassées ou si les succès spectaculaires des récents programmes tiennent plus de l'apprentissage par renforcement que du MCTS. En particulier, on peut se demander si l'apprentissage par renforcement permettrait à une approche de type minimax de rivaliser avec le MCTS sur les jeux où celle-ci domine le minimax jusqu'à présent, tels que le Go ou le Hex.

Dans cet article, nous nous focalisons par conséquent sur l'apprentissage par renforcement dans le cadre du

minimax. Nous proposons de nouvelles techniques pour l'apprentissage par renforcement de fonctions d'évaluation. Nous les appliquons pour concevoir un nouveau programme-joueur au jeu de Hex (sans utiliser d'autres connaissances que les règles du jeu). Nous comparons en particulier ce programme-joueur à Mohex 2.0 [11], le champion au Hex (taille 11) lors des Olympiades informatiques de 2013 à 2017 [9], qui est également le programme-joueur le plus fort publiquement disponible.

Dans la section 2, nous présentons succinctement les algorithmes de jeux et en particulier le minimax à profondeur non bornée sur lequel nous basons plusieurs de nos expériences. Nous présentons également l'apprentissage par renforcement dans les jeux, le jeu de Hex et l'état de l'art des programmes-joueurs sur ce jeu. Dans la section suivante, nous proposons différentes techniques ayant pour objectif d'améliorer l'apprentissage. Ensuite, nous exposons dans la section 4 les expériences réalisées utilisant ces techniques.

## 2 Contexte et travaux connexes

Dans cette section, nous présentons succinctement les algorithmes de recherche dans les arbres de jeux, l'apprentissage par renforcement dans le contexte des jeux ainsi que leurs applications dans le cadre du jeu de Hex (pour plus de détails sur les algorithmes de jeux, voir [27]).

Les jeux peuvent être représentés de manière arborescente par leur *arbre de jeu* (un nœud correspond à un état de jeu et les fils d'un nœud sont les états atteignables par une action). À partir de cette représentation, on peut déterminer l'action à jouer en utilisant un algorithme de recherche dans l'arbre de jeu. Afin de gagner, chaque joueur cherche à maximiser son score (i.e. la valeur de l'état du jeu pour ce joueur à la fin de la partie). Comme nous nous plaçons dans le cadre des jeux à deux joueurs à somme nulle, maximiser son score revient à minimiser le score de son adversaire.

### 2.1 Recherche dans les arbres de jeu

L'algorithme central est le *minimax* qui détermine récursivement la valeur d'un nœud à partir de la valeur de ses fils et des fonctions min et max, jusqu'à une profondeur limite de récursion. Avec cet algorithme, l'arbre de jeu est exploré uniformément. Une implémentation plus performante du minimax utilise l'*élagage alpha-bêta* [27] qui permet de ne pas explorer les sections de l'arbre de jeux qui sont moins intéressantes étant données les valeurs des nœuds déjà rencontrés et les propriétés du min et du max. De nombreuses variantes et améliorations du minimax ont été proposées [18]. Certaines de ces variantes effectuent une recherche à profondeur non bornée (c'est-à-dire que la profondeur de leur recherche n'est pas fixée) [26]. Contrairement au minimax avec ou sans élagage alpha-bêta, l'exploration de ces algorithmes est non uniforme. L'un de ces algorithmes est la *recherche minimax en meilleur d'abord* [13]. Pour éviter toute confusion avec certaines approches en meilleur d'abord à profondeur fixée, nous appelons cet algorithme le

*minimax en meilleur d'abord non borné*, ou plus succinctement UBFM (Unbound Best-First Minimax). UBFM étend itérativement l'arbre de jeu en ajoutant les fils de l'une des feuilles de l'arbre de jeu ayant la même valeur que celle de la racine (valeur minimax). Ces feuilles sont les états obtenus après avoir joué une des meilleures séquences d'actions possibles étant donnée la connaissance actuelle partielle de l'arbre de jeu. Ainsi, cet algorithme étend itérativement les meilleures séquences d'actions *a priori*. Ces meilleures séquences changent généralement à chaque extension. Cela permet d'explorer non uniformément l'arbre de jeu en se focalisant sur les actions les plus intéressantes *a priori* sans pour autant n'explorer qu'une seule séquence d'actions. Dans cet article, nous utilisons la version anytime d'UBFM [13], c'est-à-dire que nous laissons un temps de réflexion fixé à UBFM pour déterminer l'action à jouer. Nous utilisons en plus les tables de transposition [18] avec UBFM, ce qui permet de ne pas avoir à construire explicitement l'arbre de jeu et de fusionner les nœuds correspondant au même état.

## 2.2 Apprentissage de fonctions d'évaluation

L'apprentissage par renforcement de fonctions d'évaluation peut être effectué par différentes techniques [16, 23, 1, 28]. L'idée générale de l'apprentissage par renforcement de fonctions d'évaluation d'états est d'utiliser un algorithme de recherche dans les arbres de jeu et une fonction d'évaluation adaptative  $f$  (par exemple un réseau de neurones) pour jouer une séquence de parties (par exemple contre soi-même, ce qui est le cas dans cet article). Chaque partie va générer des couples  $(e, v)$  où  $e$  est un état et  $v$  la valeur de  $e$  calculée par l'algorithme de recherche choisi utilisant la fonction d'évaluation  $f$ . L'apprentissage consiste alors à modifier  $f$  de façon à ce que pour tous les couples  $(e, v)$  générés  $f(e)$  se rapproche de  $v$  suffisamment pour en constituer une bonne approximation. Au lieu de restreindre l'apprentissage aux données générées lors de la dernière partie, les données des dernières parties peuvent être gardées en mémoire et utilisées lors de l'apprentissage (cette technique est appelée *experience replay* [19]). Cependant, dans cet article, les données utilisées lors d'un apprentissage sont celles générées lors de la dernière partie.

*Remarque 1.* Les fonctions d'évaluation adaptatives ne servent qu'à évaluer les états non terminaux puisque l'on connaît la vraie valeur des états terminaux.

## 2.3 Distribution de sélection d'actions

Un des problèmes lié à l'apprentissage par renforcement est le *dilemme exploration-exploitation* [16]. Celui-ci consiste à choisir entre explorer de nouveaux états pour apprendre de nouvelles connaissances et exploiter les connaissances acquises. De nombreuses techniques ont été proposées pour gérer ce dilemme [17]. Cependant la plupart de ces techniques ne passent pas à l'échelle car leur application nécessite de mémoriser tous les états rencontrés. Pour cette raison, dans le cadre des jeux à grand nombre

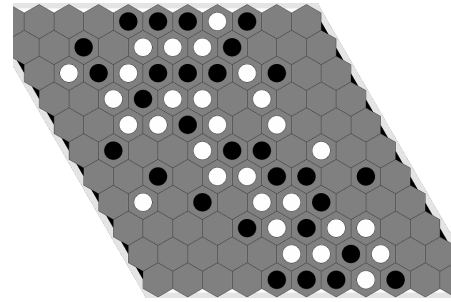


FIGURE 1 – Fin de partie au Hex, taille 11 (blanc gagne)

d'états, les travaux se basent sur une exploration probabiliste [28, 23, 16]. Avec cette approche, exploiter, c'est jouer la meilleure action et explorer, c'est jouer au hasard uniformément. Une distribution de probabilité paramétrique est alors utilisée pour associer à chaque action sa probabilité d'être jouée. Le paramètre associé à la distribution correspond au taux d'exploration compris entre 0 et 1, que nous notons  $\epsilon$  (le taux d'exploitation est donc  $1 - \epsilon$ , que nous notons  $\epsilon'$ ). Celui est souvent fixé expérimentalement. Un *recuit simulé* [12] peut cependant être appliqué afin de ne pas avoir besoin de choisir une valeur pour ce paramètre. Dans ce cas, au début de l'apprentissage par renforcement, le paramètre vaut 1 (on ne fait qu'explorer). Celui-ci diminue au fur et à mesure jusqu'à atteindre 0 à la fin de l'apprentissage. La distribution de sélection d'actions la plus simple est  $\epsilon$ -greedy [28] (de paramètre  $\epsilon$ ). Avec cette distribution, l'action est choisie uniformément avec probabilité  $\epsilon$  et la meilleure action est choisie avec probabilité  $1 - \epsilon$ . La distribution  $\epsilon$ -greedy a le désavantage de ne pas différencier les actions (hormis la meilleure action) en termes de probabilités. Une autre distribution est souvent utilisée, corrigeant cette inconvenance. Il s'agit de la fonction *softmax*, également appelée *distribution de Boltzmann* [16].

## 2.4 Jeu de Hex

Le jeu de Hex est un jeu de stratégie combinatoire pour deux joueurs. Il se joue sur un plateau hexagonal  $n \times n$  vide. On dira qu'un plateau  $n \times n$  est de taille  $n$ . Le plateau peut être de taille quelconque, bien que les tailles classiques soient 11, 13 et 19. A son tour, chaque joueur pose une pièce de sa couleur sur une case vide (chaque pièce est identique). Le but du jeu est d'être le premier à relier les deux bords opposés du plateau correspondant à sa couleur. La figure 1 illustre une fin de partie. Bien que ces règles soient simplistes, les tactiques et stratégies au jeu de Hex sont complexes. Le nombre d'états et le nombre d'actions par état sont très importants, similaires au jeu de Go. Le nombre d'états est par exemple supérieur à celui des échecs, au moins dès le plateau de taille 11 (table 6 de [25]). Quelle que soit la taille du plateau, le premier joueur a une stratégie gagnante [3] qui est inconnue, sauf pour les plateaux de taille inférieure ou égale à 9 [20]. En fait, résoudre un état particulier est PSPACE-complet [4].

## 2.5 Programmes-joueurs au Hex

De nombreux programmes-joueurs ont été développés au Hex. Par exemple, Mohex 1.0 [11] est un programme basé sur le MCTS. Il utilise en plus de nombreuses techniques dédiées au Hex, basées sur des résultats théoriques spécifiques au jeu de Hex. Il est en particulier capable de déterminer une stratégie gagnante pour certains états et d'élaguer à chaque état de nombreuses actions qu'il sait être *inférieures*. Il utilise également des connaissances ad-hoc pour biaiser les simulations du MCTS.

Mohex 2.0 [11] est une amélioration de Mohex 1.0 qui utilise des connaissances apprises par apprentissage supervisé pour orienter à la fois l'exploration de l'arbre et les simulations du MCTS. Cet apprentissage supervisé a permis d'apprendre les corrélations de victoire de motifs du plateau.

D'autres travaux se sont ensuite focalisés sur la prédiction des meilleures actions, par apprentissage supervisé d'une base de données de parties, à l'aide d'un réseau de neurones. Le réseau de neurones est utilisé pour apprendre une *politique*, c'est-à-dire une distribution de probabilité a priori sur les actions à jouer. Ces probabilités a priori sont utilisées pour guider l'exploration du MCTS. Il y a d'abord Mohex-CNN [7] qui est une amélioration de Mohex 2.0 utilisant un réseau de neurones convolutionnel [14]. Une nouvelle version de Mohex a ensuite été proposée : Mohex-3HNN [8]. Contrairement à Mohex-CNN, il est basé sur un réseau de neurones résiduel [10]. Il calcule, en plus de la politique, une valeur pour les états *et* les actions. La valeur des états remplace l'évaluation des états à base de simulations du MCTS. L'ajout d'une valeur pour les actions permet à Mohex-HNN de diminuer le nombre d'appel du réseau de neurones, améliorant ainsi les performances.

Des programmes apprenant la fonction d'évaluation par renforcement ont également été conçus. Ces programmes sont NeuroHex [28], EZO-CNN [24] et ExIt [1]. Ils apprennent en jouant contre eux-mêmes. Contrairement aux deux autres, NeuroHex effectue au préalable un apprentissage supervisé afin de diminuer le temps d'apprentissage par renforcement (en apprenant une heuristique commune du jeu de hex). NeuroHex démarre en plus ses parties d'un état provenant d'une base de données de parties. EZO-CNN, lui, utilise durant l'apprentissage des connaissances lui permettant de jouer (et donc d'apprendre) une stratégie gagnante dans certains états. ExIt, quant à lui, apprend une politique en plus de la valeur des états de jeu. Il est le seul programme à avoir appris à jouer au Hex sans utiliser de connaissances. Ce résultat est cependant limité aux plateaux de taille 9. La table 1 compare les caractéristiques principales de ces différents programmes.

## 3 Techniques proposées

Nous proposons maintenant plusieurs techniques visant à améliorer et/ou accélérer l'apprentissage de fonctions d'évaluation d'états de jeu. Nous terminons cette section en proposant une légère modification d'UBFM que nous appelons UBFM<sub>s</sub> afin d'améliorer le niveau de jeu.

## 3.1 Apprentissage de l'arbre de jeu

Avec les différents travaux de la littérature sur l'apprentissage par renforcement de fonctions d'évaluation, les états  $e$  des couples  $(e, v)$  utilisés lors de l'apprentissage sont les différents états du jeu obtenus au cours de la partie (il ne s'agit pas des états générés lors de la recherche dans l'arbre de jeu). Chaque valeur  $v$  est déterminée par l'algorithme de recherche utilisé. Nous appelons cette technique le *root learning*. Cette approche perd ainsi une partie des informations contenues dans l'arbre de jeu générées lors de la recherche de l'action à jouer. Plus précisément, les valeurs des états de l'arbre de jeu sont perdues puisqu'elles ne sont pas utilisées lors de l'apprentissage de la fonction d'évaluation. Nous proposons donc d'apprendre l'intégralité de l'arbre partiel de jeu (à l'exception des feuilles non terminales puisque l'on a déjà  $f(e) = v$ ) au lieu de se restreindre à apprendre la valeur de la racine. Nous appelons cette technique le *tree learning*.

Le *tree learning* peut sembler moins intéressant que le *root learning* puisqu'il ajoute à l'apprentissage des valeurs de précision inférieure à celle de la racine (puisque moins d'états sont générés pour les déterminer). Il est de plus possible qu'un certain nombre d'états de l'arbre de jeu n'aient jamais été explorés depuis le début de l'apprentissage par renforcement. Des valeurs aberrantes seraient alors apprises continuellement. Ces valeurs supplémentaires ont un gros risque de biaiser significativement l'apprentissage par renforcement. Il y a donc un dilemme entre le gain en information qui peut être apporté par l'apprentissage de ces valeurs supplémentaires et le risque de biaiser l'apprentissage dû à l'imprécision de ces valeurs.

Afin de réduire le biais dû à l'apprentissage de ces valeurs, nous proposons une technique supplémentaire que nous appelons *coefficiented tree learning*. Cette technique affecte à chaque état un coefficient  $c$  indiquant à quel point sa valeur est précise. Les valeurs avec un coefficient plus grand doivent alors être préférentiellement mieux apprises que celles avec un faible coefficient. Pour implémenter cette idée, lors des expériences de cet article, le coefficient est calculé par la formule suivante :  $c = 1 + \lfloor \log_2(n + 1) \rfloor$  avec  $n$  le nombre de fois que la valeur de cet état a été mise à jour au cours de la recherche ( $n$  est donc le nombre de nœuds descendants du nœud de cet état). Pour tenir compte des coefficients lors de l'apprentissage, nous dupliquons  $c$  fois chaque couple  $(e, v)$  avec  $c$  son coefficient.

**Exemple 2.** Suite à une partie, nous obtenons les couples suivant  $(e_1, 1)$ ,  $(e_2, -2)$  et  $(e_3, 5)$ . Les états  $e_1$ ,  $e_2$  et  $e_3$  ont respectivement 0, 1 et 3 états descendants dans l'arbre de jeu. Le coefficient de  $e_1$  (resp.  $e_2$ ; resp.  $e_3$ ) est donc 1 (resp. 2; resp. 3). Les données utilisées lors de l'apprentissage de la fonction d'évaluation adaptative  $f$  sont les couples suivants  $(e_1, 1)$ ,  $(e_2, -2)$ ,  $(e_2, -2)$ ,  $(e_3, 5)$ ,  $(e_3, 5)$ ,  $(e_3, 5)$ .

## 3.2 Heuristique de la profondeur

Dans cette section, nous proposons une fonction d'évaluation des états terminaux alternative à la fonction d'éva-

Programmes	Tailles	Recherche	Apprentissage	Réseau	Utilisation
Mohex-CNN	13	MCTS	supervisé	convolutionnel	politique
Mohex-3HNN	13	MCTS	supervisé	résiduel	politique, état, action
NeuroHex	13	aucun	supervisé et renforcement	convolutionnel	état
EZO-CNN	7, 9, 11	Minimax	renforcement	convolutionnel	état
ExIt	9	MCTS	renforcement	convolutionnel	politique et état

TABLE 1 – Comparatif des caractéristiques principales des derniers programmes au Hex. Ces caractéristiques sont respectivement les tailles de plateau sur lesquelles ils peuvent jouer, l’algorithme de recherche dans les arbres utilisé, le type d’apprentissage effectué, le type de réseau de neurones utilisé et l’utilisation du réseau de neurones qui en est fait, c’est-à-dire s’il est utilisé pour approximer les valeurs des états, des actions et/ou de la politique.

luation classique des états terminaux, qui retourne 1 si le joueur gagne et  $-1$  si son adversaire gagne [28, 23, 8]. Nous l’appelons *heuristique de la profondeur* et nous la notons  $p_t$ . Elle donne une meilleure valeur aux états gagnants proches du début de partie qu’aux états gagnants loin du début de partie. Elle donne également une meilleure valeur aux états perdants loin du début de partie qu’aux états perdants proches du début de partie. Avec cette fonction d’évaluation, on cherche à gagner le plus tôt possible et perdre le plus tard possible. En apprenant par renforcement avec cette fonction d’évaluation terminale, on préférera ainsi un état que l’on pense gagnant et proche de la fin de partie à un autre état que l’on pense gagnant et loin de la fin de partie. On peut espérer que cela améliore le niveau de jeu. Un état que l’on pense gagnant et proche de la fin de partie a moins de chances d’être un état perdant qu’un état que l’on pense gagnant et loin de la fin de partie. En effet, plus un état est proche de la fin de partie, plus il y a de chances que sa valeur soit précise. De plus, avec une partie longue, un joueur en difficulté aura plus d’occasions de reprendre le dessus. Inversement, un état que l’on pense perdant et loin de la fin de partie sera préféré à un état que l’on pense perdant et proche de la fin de la partie, puisque le premier état a hypothétiquement une valeur moins précise et a donc *a priori* plus de chances d’être en réalité un état gagnant.

Nous proposons comme réalisation de l’heuristique de la profondeur la fonction  $p_t$  suivante. La fonction d’évaluation  $p_t$  retourne la valeur  $l$  si le joueur gagne et la valeur  $-l$  si son adversaire gagne, avec  $l$  le nombre maximum d’actions réalisables au cours d’une partie plus 1 moins le nombre d’actions jouées depuis le début de la partie. Pour le Hex,  $l$  est le nombre de cases vides du plateau plus 1.

### 3.3 Complétion de fonctions d’évaluation

Nous introduisons dans cette section la complétion  $C$  d’une fonction d’évaluation d’états de jeu  $f$ . La complétion a pour objectif d’améliorer la fonction d’évaluation en tenant compte de la résolution des états. Un état est dit *résolu* si l’algorithme de recherche a permis de déterminer la vraie valeur de cet état. Sans la complétion, on peut avoir un état résolu et gagnant de valeur plus faible qu’un état non résolu que l’on pense gagnant mais qui est peut-être perdant. L’objectif de la complétion est donc de toujours choisir une action menant à un état résolu gagnant et de ne

jamais choisir une action menant à un état résolu perdant lorsque l’on a le choix (lors de l’exploration de l’arbre ou de la décision de l’action à jouer). La fonction de complétion peut se définir par  $C(f(e)) = (r(e), f(e))$  avec  $e$  un état et  $r$  la fonction suivante qui retourne 1 si l’état est résolu et si le joueur gagne,  $-1$  si l’état est résolu et si son adversaire gagne et 0 si l’état n’est pas résolu. On utilisera alors l’ordre lexicographique pour comparer les états.

*Remarque 3.* Bien que la complétion remplace la fonction d’évaluation lors de l’exploration et lors du choix de l’action à jouer, les couples utilisés lors de l’apprentissage restent les couples de la forme  $(e, f(e))$ .

Il n’y a pas besoin de compléter une fonction d’évaluation d’états non terminaux si elle est à valeurs dans  $]a, b[$  (avec  $a < b$ ) et si les états terminaux sont évalués par  $a$  s’ils sont perdants et par  $b$  s’ils sont gagnants.

### 3.4 Descente : générer de meilleures données

Nous introduisons dans cette section une modification d’UBFM, que nous nommons *descente*, visant à être utilisée lors de l’apprentissage par renforcement afin de générer des couples  $(e, v)$  différents mais supposément de meilleure qualité. L’idée de *descente* est de combiner UBFM avec des simulations déterministes de fin de partie apportant des valeurs intéressantes du point de vue de l’apprentissage. L’algorithme *descente* (algorithme 1) sélectionne récursivement le meilleur fils du nœud actuel, qui devient le nouveau nœud actuel. Il ajoute les fils du nœud actuel s’ils ne sont pas dans l’arbre. Il effectue cette récursion partant de la racine (l’état actuel du jeu) jusqu’à atteindre un nœud terminal (une fin de partie). Il met alors à jour la valeur des nœuds sélectionnés (valeur minimale). L’algorithme *descente* recommence cette opération récursive repartant de la racine tant qu’il reste du temps de recherche. *Descente* est quasiment identique à UBFM. La seule différence est que *descente* effectue une itération jusqu’à atteindre un état terminal alors qu’UBFM effectue cette itération jusqu’à atteindre une feuille de l’arbre (UBFM arrête donc l’itération beaucoup plus tôt). Autrement dit, lors d’une itération, UBFM étend juste une des feuilles de l’arbre de jeu alors que *descente* étend récursivement le meilleur des fils en partant de cette feuille jusqu’à atteindre une fin de partie. L’algorithme *descente* a l’avan-

tage d'UBFM, c'est-à-dire d'effectuer une recherche plus longue permettant de déterminer une meilleure action à jouer. Grâce au tree learning, il a également l'avantage d'une recherche minimax à profondeur 1, c'est-à-dire de faire remonter les valeurs des nœuds terminaux vers les autres nœuds plus rapidement. En outre, les états ainsi générés sont plus proches des états terminaux. Leurs valeurs sont donc de meilleures approximations.

```

Fonction descente( $e$ )
  if terminal( $e$ ) then
    | retourner  $f(e)$ 
  else
    if  $e \notin E$  then
      |  $E \leftarrow E \cup \{e\}$ 
      | foreach  $a \in \text{actions}(e)$  do
      | |  $v(e, a) \leftarrow f(a(e))$ 
     $a_b \leftarrow \text{meilleure\_action}(e)$ 
     $v(e, a_b) \leftarrow \text{descente}(a_b(e))$ 
     $a_b \leftarrow \text{meilleure\_action}(e)$ 
    retourner  $v(e, a_b)$ 

```

```

Fonction meilleure_action( $e$ )
  if premier_joueur( $e$ ) then
    | retourner  $\arg \max_{a \in \text{actions}(e)} v(e, a)$ 
  else
    | retourner  $\arg \min_{a \in \text{actions}(e)} v(e, a)$ 

```

```

Fonction apprentissage_descente()
   $E \leftarrow \{\}$ 
   $e \leftarrow \text{état initial}$ 
  while  $\neg \text{terminal}(e)$  do
    |  $t = \text{time}()$ 
    | while  $\text{time}() - t < \tau$  do descente( $e$ )
    |  $a_b \leftarrow \text{selection\_action}(e)$ 
    |  $e \leftarrow a_b(e)$ 
   $D \leftarrow \{\}$ 
  foreach  $e \in E$  do
    |  $a_b \leftarrow \text{meilleure\_action}(e)$ 
    |  $D \leftarrow D \cup \{(e, v(e, a_b))\}$ 
  adapter  $f$  par rapport à  $D$ 

```

**Algorithme 1** : Apprentissage d'une partie avec *descente* et *tree learning* ( $a(e)$  : état obtenu après avoir joué l'action  $a$  dans l'état  $e$ ;  $v(e, a)$  : valeur obtenue après avoir joué  $a$  dans  $e$ ;  $f$  est du point de vue du premier joueur;  $E$  : clefs de la table de transposition;  $\tau$  : temps de réflexion par action; *selection\_action* : *meilleure\_action*() ou une autre distribution d'actions).

### 3.5 Distribution ordinale d'actions

Dans cette section, nous proposons une distribution de probabilité alternative à la fonction softmax et à  $\epsilon$ -greedy (voir la section 2.3). Cette distribution ne dépend pas de la va-

leur des états. Elle dépend par contre de l'ordre de leurs valeurs. Sa formule est la suivante :

$$P(f_i) = \left( \epsilon' + \frac{1 - \epsilon'}{n - i} \right) \cdot \left( 1 - \sum_{j=0}^{j < i} P(f_j) \right)$$

avec  $n$  le nombre de fils de la racine,  $i \in \{0, \dots, n - 1\}$ ,  $f_i$  le  $i^{\text{ème}}$  meilleur fils de la racine,  $P(f_i)$  la probabilité de jouer l'action menant au fils  $f_i$  et  $\epsilon'$  le paramètre d'exploitation ( $\epsilon' = 1 - \epsilon$ ).

Nous proposons en plus d'utiliser la règle suivante lors de la sélection de l'action à jouer, pour réduire la durée des parties et donc *a priori* la durée de l'apprentissage : toujours jouer une action menant à un état résolu gagnant s'il existe et ne jamais jouer une action menant à un état résolu perdant si cela est possible. Ainsi, si parmi les actions disponibles on sait qu'une des actions est gagnante, on la joue. S'il n'y en a pas, on joue aléatoirement suivant la loi choisie parmi les actions ne menant pas à un état résolu perdant (si les états ne sont pas tous perdants).

### 3.6 UBFM<sub>s</sub> : jouer la sécurité

Nous proposons maintenant une légère modification d'UBFM, notée UBFM<sub>s</sub>, qui vise à fournir un jeu plus sûr. L'action qu'UBFM choisit de jouer est celle qui mène à l'état de meilleure valeur. Dans certains cas, la meilleure action (*a priori*) peut mener à un état qui n'a pas été *suffisamment* visité (comme une feuille non terminale). Choisir cette action est donc une décision risquée. Nous proposons, pour éviter ce problème, une décision différente qui vise à jouer l'action la plus sûre, à l'instar du MCTS (max child selection [5]). Si aucune action ne mène à un état résolu gagnant, l'action choisie par UBFM<sub>s</sub> est celle qui a été le plus de fois sélectionnée (depuis l'état actuel du jeu) au cours de l'exploration de l'arbre. En cas d'égalité, UBFM<sub>s</sub> départage en choisissant celle qui mène à l'état de meilleure valeur. Cette décision est plus sûre du fait que le nombre de fois qu'une action est sélectionnée correspond au nombre de fois que cette action est plus intéressante que les autres (la recherche est effectuée en meilleur d'abord).

## 4 Expériences

Nous décrivons maintenant les différentes expériences réalisées. Elles se basent sur les techniques présentées dans la section précédente. Nous commençons par évaluer le résultat d'un apprentissage en temps long utilisant nos techniques contre la version 1.0 et la version 2.0 de Mohex. Enfin, nous comparons expérimentalement différents algorithmes pour jouer, tels que UBFM<sub>s</sub> et la complétion.

### 4.1 Résultats contre Mohex 1.0 et Mohex 2.0

Un long apprentissage contre soi-même pour les plateaux de taille 11 a été effectué avec l'algorithme descente utilisant l'heuristique de la profondeur, la complétion, le coefficiented tree learning ainsi que la distribution et la règle associée de la section 3.5. La fonction d'évaluation a été



pré-initialisée en apprenant les valeurs d'états terminaux aléatoires (au nombre de 8.836.000). L'apprentissage par renforcement de la fonction d'évaluation a duré 58.674 parties, à raison d'une seconde de réflexion par action.

UBFM<sub>s</sub> doté de la fonction d'évaluation générée par cet apprentissage atteint le score de 60% de victoire (84% en premier joueur; 37% en second joueur) contre Mohex 1.0 à la suite de 300 parties en premier joueur et 300 autres parties en second joueur, avec une seconde de réflexion par action. Son score contre Mohex 2.0 est de 53% victoire (81% en premier joueur; 25% en second joueur) à la suite de 1000 parties en premier joueur et 1000 autres parties en second joueur, avec une seconde et demi par action.

## 4.2 Comparaison d'algorithmes de jeu

Nous comparons maintenant les taux de victoire de différents algorithmes pour jouer en les évaluant contre Mohex 2.0. Chacun d'eux se base sur la fonction d'évaluation des plateaux de taille 11 générée au cours de l'expérience précédente. Nous comparons UBFM<sub>s</sub> sans complétion et les algorithmes suivants avec complétion : UBFM, UBFM<sub>s</sub>, alpha-bêta à profondeur 1, alpha-bêta à profondeur 2 ainsi qu'une modification d'UBFM<sub>s</sub>, notée UBFM<sub>s</sub> + UCT. Cette variante effectue une exploration en meilleur d'abord tenant compte du nombre de visites des nœuds à la manière de la version la plus populaire du MCTS, UCT [5]. Pour cela, le fils sélectionné est le meilleur selon la somme de la valeur d'UBFM<sub>s</sub> (normalisée pour être dans [0, 1]) et du terme d'exploration  $c \cdot \sqrt{\frac{\log N}{n}}$  avec  $n$  le nombre de visites du fils considéré,  $N$  le nombre de visites du nœud père et  $c$  une constante d'exploration (en théorie  $\sqrt{2}$ ).

Les taux de victoire contre Mohex 2.0 sont décrits dans la table 2. C'est UBFM<sub>s</sub> qui obtient le meilleur taux de victoire. L'utilisation de la complétion apporte un très léger gain de l'ordre de 2%. Jouer l'action la plus sûre a apporté un gain de 6%. UBFM<sub>s</sub> est supérieur de 24% à l'alpha-bêta à profondeur 1. Il est supérieur de 18% à l'alpha-bêta à profondeur 2, bien que ce dernier ait un temps de réflexion par action  $\frac{1}{4}$  de temps supérieur en moyenne. Enfin, ajouter de l'exploration à UBFM<sub>s</sub> a diminué le taux de victoire d'au moins 17% (pour chaque valeur  $c$  choisie).

## 5 Conclusion

Nous avons proposé plusieurs nouvelles techniques pour l'apprentissage par renforcement de fonctions d'évaluation. La première technique consiste à apprendre les valeurs de l'arbre de jeu au lieu de se restreindre à la valeur de la racine. Pour limiter le biais induit sur l'apprentissage par l'ajout de ces valeurs, nous avons proposé d'associer un coefficient à chaque état, dépendant du nombre de visites. Nous avons également proposé de remplacer le gain classique d'un jeu (+1/-1) par l'heuristique de la profondeur. Cela permet de tenir compte de la durée de la partie afin de favoriser les victoires rapides et les défaites lentes. Nous suggérons l'utilisation de la complétion pour améliorer le comportement de certaines fonctions d'évaluation, comme

Algorithmes	1 <sup>er</sup>	2 <sup>ème</sup>	moyenne
UBFM <sub>s</sub>	81%	25%	53%
UBFM <sub>s</sub> sans complétion	77%	25%	51%
UBFM	61%	34%	47%
Alpha-bêta à profondeur 2	56%	13%	35%
Alpha-bêta à profondeur 1	37%	20%	29%
UBFM <sub>s</sub> + UCT ( $c = 2 \cdot \sqrt{2}$ )	55%	12%	33%
UBFM <sub>s</sub> + UCT ( $c = \sqrt{2}$ )	55%	15%	35%
UBFM <sub>s</sub> + UCT ( $c = \frac{\sqrt{2}}{2}$ )	50%	12%	31%
UBFM <sub>s</sub> + UCT ( $c = \frac{\sqrt{2}}{10}$ )	56%	15%	36%

TABLE 2 – Statistiques de victoire contre Mohex 2.0 de différents algorithmes de recherche pour 600 parties en premier joueur et 600 autres parties en second joueurs. UBFM, UBFM<sub>s</sub>, chaque UBFM<sub>s</sub> + UCT et Mohex 2.0 ont une seconde et demi par action. Alpha-bêta à profondeur 2 prend en moyenne environ deux secondes par action.

celles basées sur l'heuristique de la profondeur. Enfin, nous avons proposé l'algorithme descente qui explore à la manière du minimax à profondeur non bornée. Contrairement à ce dernier, descente explore itérativement les séquences de meilleures actions *jusqu'aux états terminaux*. Son objectif est d'améliorer la qualité des données utilisées lors de l'apprentissage, tout en gardant les avantages du minimax à profondeur non bornée.

En outre, nous avons proposé une autre variante du minimax à profondeur non bornée, qui joue l'action la plus sûre au lieu de jouer la meilleure action. Nos expériences suggèrent que cela améliore le niveau de jeux lors des confrontations. Nos expériences suggèrent également qu'une fonction d'évaluation apprise par renforcement donne un meilleur niveau de jeux lorsqu'elle est utilisée par un algorithme de recherche en meilleur d'abord. Enfin, l'utilisation des techniques proposées a permis de concevoir, pour la première fois, un programme-joueur au jeu de Hex (taille 11) atteignant le niveau de Mohex 2.0 par un apprentissage par renforcement contre soi-même sans utilisation de connaissance (ni de techniques Monte Carlo). Une étude comparant la vitesse d'apprentissage en fonction des techniques utilisées est en cours. Les résultats préliminaires suggèrent d'une part que l'utilisation de l'heuristique de la profondeur améliore l'apprentissage. Ils suggèrent d'autre part que l'apprentissage avec descente est meilleur qu'avec UBFM.

## Remerciements

Je remercie le GREYC de m'avoir donné accès à son serveur de calcul, ce qui m'a permis d'effectuer l'expérience de la section 4.1.

## Références

- [1] Thomas Anthony, Zheng Tian, and David Barber. Thinking fast and slow with deep learning and tree

- search. In *Advances in Neural Information Processing Systems*, pages 5360–5370, 2017.
- [2] Hendrik Baier and Mark HM Winands. Mcts-minimax hybrids with state evaluations. *Journal of Artificial Intelligence Research*, 62 :193–231, 2018.
- [3] Elwyn R Berlekamp, John H Conway, and Richard K Guy. *Winning Ways for Your Mathematical Plays, Volume 3*. AK Peters/CRC Press, 2018.
- [4] Édouard Bonnet, Florian Jamain, and Abdallah Saffidine. On the complexity of connection games. *Theoretical Computer Science*, 644 :2–28, 2016.
- [5] Cameron B Browne, Edward Powley, Daniel Whitehouse, Simon M Lucas, Peter I Cowling, Philipp Rohlfshagen, Stephen Tavener, Diego Perez, Spyridon Samothrakis, and Simon Colton. A survey of monte carlo tree search methods. *Transactions on Computational Intelligence and AI in games*, 4(1) :1–43, 2012.
- [6] Tristan Cazenave. Residual networks for computer go. *Transactions on Games*, 10(1) :107–110, 2018.
- [7] Chao Gao, Ryan B Hayward, and Martin Müller. Move prediction using deep convolutional neural networks in hex. *Transactions on Games*, 2017.
- [8] Chao Gao, Martin Müller, and Ryan Hayward. Three-head neural network architecture for monte carlo tree search. In *International Joint Conference on Artificial Intelligence*, pages 3762–3768, 2018.
- [9] Ryan Hayward and Noah Weninger. Hex 2017 : Mo-hex wins the  $11 \times 11$  and  $13 \times 13$  tournaments. *ICGA Journal*, 39(3-4) :222–227, 2017.
- [10] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. In *Conference on computer vision and pattern recognition*, pages 770–778, 2016.
- [11] Shih-Chieh Huang, Broderick Arneson, Ryan B Hayward, Martin Müller, and Jakub Pawlewicz. Mohex 2.0 : a pattern-based mcts hex player. In *International Conference on Computers and Games*, pages 60–71. Springer, 2013.
- [12] Scott Kirkpatrick, C Daniel Gelatt, and Mario P Vecchi. Optimization by simulated annealing. *Science*, 220(4598) :671–680, 1983.
- [13] Richard E Korf and David Maxwell Chickering. Best-first minimax search. *Artificial intelligence*, 84(1-2) :299–337, 1996.
- [14] Alex Krizhevsky, Ilya Sutskever, and Geoffrey E Hinton. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, pages 1097–1105, 2012.
- [15] Michael Lederman Littman. *Algorithms for sequential decision making*. PhD thesis, 1996.
- [16] Jacek Mandziuk. *Knowledge-free and learning-based methods in intelligent game playing*, volume 276. Springer, 2010.
- [17] Joseph Mellor. *Decision Making Using Thompson Sampling*. PhD thesis, 2014.
- [18] Ian Millington and John Funge. *Artificial intelligence for games*. CRC Press, 2009.
- [19] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A Rusu, Joel Veness, Marc G Bellemare, Alex Graves, Martin Riedmiller, Andreas K Fidjeland, Georg Ostrovski, et al. Human-level control through deep reinforcement learning. *Nature*, 518(7540) :529, 2015.
- [20] Jakub Pawlewicz and Ryan B Hayward. Scalable parallel dfpn search. In *International Conference on Computers and Games*, pages 138–150. Springer, 2013.
- [21] David Silver, Aja Huang, Chris J Maddison, Arthur Guez, Laurent Sifre, George Van Den Driessche, Julian Schrittwieser, Ioannis Antonoglou, Veda Panneershelvam, Marc Lanctot, et al. Mastering the game of go with deep neural networks and tree search. *Nature*, 529(7587) :484, 2016.
- [22] David Silver, Thomas Hubert, Julian Schrittwieser, Ioannis Antonoglou, Matthew Lai, Arthur Guez, Marc Lanctot, Laurent Sifre, Dhharshan Kumaran, Thore Graepel, et al. Mastering chess and shogi by self-play with a general reinforcement learning algorithm. *arXiv preprint arXiv :1712.01815*, 2017.
- [23] David Silver, Julian Schrittwieser, Karen Simonyan, Ioannis Antonoglou, Aja Huang, Arthur Guez, Thomas Hubert, Lucas Baker, Matthew Lai, Adrian Bolton, et al. Mastering the game of go without human knowledge. *Nature*, 550(7676) :354, 2017.
- [24] Kei Takada, Hiroyuki Iizuka, and Masahito Yamamoto. Reinforcement learning for creating evaluation function using convolutional neural network in hex. In *2017 Conference on Technologies and Applications of Artificial Intelligence*, pages 196–201. IEEE, 2017.
- [25] H Jaap Van Den Herik, Jos WHM Uiterwijk, and Jack Van Rijswijk. Games solved : Now and in the future. *Artificial Intelligence*, 134(1-2) :277–311, 2002.
- [26] H Jaap Van Den Herik and Mark HM Winands. Proof-number search and its variants. In *Oppositional Concepts in Computational Intelligence*, pages 91–118. Springer, 2008.
- [27] Georgios N Yannakakis and Julian Togelius. *Artificial Intelligence and Games*. Springer, 2018.
- [28] Kenny Young, Gautham Vasan, and Ryan Hayward. Neurohex : A deep q-learning hex agent. In *Computer Games*, pages 3–18. Springer, 2016.

# Argumentation-based Negotiation with Incomplete Opponent Profiles

Yannis Dimopoulos<sup>1</sup>Jean-Guy Mailly<sup>2</sup>Pavlos Moraitis<sup>2</sup><sup>1</sup> Department of Computer Science, University of Cyprus, Cyprus<sup>2</sup> LIPADE, Université Paris Descartes, France

yannis@cs.ucy.ac.cy, jean-guy.mailly@parisdescartes.fr, pavlos@mi.parisdescartes.fr

## Abstract

*Computational argumentation has taken a predominant place in the modeling of negotiation dialogues over the last years. A competent agent participating in a negotiation process is expected to decide its next move taking into account an, often incomplete, model of its opponent. This work provides a complete computational account of argumentation-based negotiation under incomplete opponent profiles. After the agent identifies its best option, in any state of a negotiation, it looks for suitable arguments that support this option in the theory of its opponent. As the knowledge on the opponent is uncertain, the challenge is to find arguments that, ideally, support the selected option despite the uncertainty. We present a negotiation framework based on these ideas, along with experimental evidence that highlights the advantages of our approach.*

## Keywords

Abstract argumentation, automated negotiation.

## 1 Introduction

During the last years computational argumentation has taken a predominant place in the modeling of negotiation dialogues (for a survey see [11], [24]). The goal of a negotiation dialogue is to allow interacting agents to resolve conflicts and reach a mutually accepted agreement, which in this work is a mutually accepted offer (e.g. the price of a product, the mode of payment). In an argumentation-based negotiation (ABN), agents choose offers that are likely to be accepted by the opponent and exchange arguments that support these offers, either based on their own theories (see e.g. [1], [3], [18],[13], [22], [14]), or based on the opponent's profile (e.g. [15], [23], [8]).

The modeling of the opponent profile is an important issue in negotiation dialogues (and more generally other types of dialogue such as persuasion). As explained in [5], although there are important differences between opponent models, there are strong reasons justifying their use, such as the *minimization of negotiation cost*, the *adaptation to the opponent* and the *capacity to reach win-win agreements*, especially in cooperative environments. Learning

the opponent profile means learning its acceptance and bidding strategies, the deadlines and its preference profile [5]. In most of the proposed works, the (online) opponent modeling is based on learning techniques (see e.g. [4] for a survey). Apart from the fact that learning the opponent profile with traditional learning techniques is not an easy task, as pointed out in [28], those techniques seem better suited to game-theoretic (or utility-based) negotiations, rather than argumentation-based negotiations. Other works (although they concern persuasion dialogues and legal disputes), have proposed a probabilistic approach for dealing with the uncertainty about the opponent profile. In these works (e.g. [16], [27], [17]), probabilities are used in different ways for finding the arguments that are most likely to be accepted by the opponent. Finally, some works (e.g. [26], [21], [7]) investigate other approaches to modeling the opponent profile in argumentation-based dialogues.

This work advances the state of the art in argumentation-based negotiation by making original contributions to the *opponent modeling*, and the associated *acceptance strategy* (i.e. what offers are most likely to be accepted) as well as *bidding strategy* (i.e. the strategy that an agent applies for choosing the next offer). For opponent modeling, it builds on the work of [9] on control argumentation frameworks (CAFs), a formalism for modeling the uncertainty about the *opponent profile*. More specifically, it borrows the concepts of "on/off" arguments (i.e. arguments we don't know whether they are present or not in a theory), and the three different categories of attacks (i.e. attacks we know their existence and direction, attacks we know the existence but not the direction, attacks we don't know the existence but we know the direction). This allows generating different profiles modeled as completions of the known part of the opponent's theory, and seeking offers that satisfy all possible profiles (or as many as possible). Regarding the *bidding* and *acceptance strategies*, the originality of this work lies in the assumption that in argumentation-based negotiation, a central challenge for an agent is to lead, by means of appropriate arguments, its counter party to change its theory, and eventually accept the offer it proposes, hence influencing its *acceptance strategy*. Thus, in our approach, we propose a *bidding strategy* that relies on

the previous assumption. More precisely, the idea is that a proponent agent uses first its own theory for choosing the best offer to propose, but next, it uses the incomplete theory of its opponent to find the arguments to support it. Then, it seeks and puts forward a set of arguments called *control configuration*, that could reinstate the supporting arguments, if these are rejected in the current state of the argumentative negotiation theories of all (or most) of the generated opponent profiles. Once the arguments of the control configuration are inserted in the opponent theory, they would, ideally, allow it to reach an agreement with the proponent, thus they alter its *acceptance decision*.

## 2 Background

We assume that the reader is familiar with abstract argumentation frameworks (AFs) as introduced in [12], presented as a pair  $\langle A, R \rangle$ , where  $A$  is a set of *arguments*, and  $R \subseteq A \times A$  is an *attack relation*.  $a$  attacks  $b$  is denoted by  $(a, b) \in R$ . Different acceptability semantics were also introduced in this work. Based on the semantics, we can define the status of any argument, namely *skeptically accepted*, *credulously accepted* and *rejected arguments*.

Now we introduce briefly the control argumentation frameworks (CAF) proposed in [9], and discuss how they capture the knowledge of an agent on its opponents. On a high level, a CAF is an AF where arguments are divided in three parts, *fixed*, *uncertain* and *control*.

The *fixed* part of the theory concerns the certain knowledge that an agent holds about its opponent. This includes arguments as well as attacks that undoubtedly belong to the argumentation theory of the opponent. For instance, a seller agent knows that the customer agent prefers European cars, that safety is an important issue for it and that it prefers electric or gasoline-powered cars than diesel cars. The *uncertain* part captures the uncertainty about the presence of arguments in a theory (expressed by the “on/off” arguments as shown below), as well as the presence and the direction of attacks between arguments in this theory. It reflects the uncertainty that arises due to lack of complete information on the current state of the world that determines the decisions of the opponent, but also its beliefs and preferences. For example, the seller agent may not know the income of the customer agent, whether a car is a social status symbol for it, the highest price that it is ready to pay, or whether it is willing to pay more if some extras are included, and payment by installments is accepted. Finally, the *control* part contains arguments that can be used against arguments of the fixed or uncertain parts that attack arguments that are in favour of some offer of the proponent. Therefore, the control part serves to ensure that arguments in the fixed part that support some offer of the seller that is not adequate with some certain (i.e. European car) or uncertain (e.g. max price, preferred mode of payment) preferences of the customer, can be accepted under some circumstances. For instance, a control argument could allow a seller agent to propose a car from abroad Europe (which is

against the known preference of the customer agent and represented in the fixed part) by proposing some interesting options (e.g. five airbags knowing that safety is an important issue for the customer and also represented in the fixed part) and in a price that is probably higher than the highest price the customer is intended to pay (this is part of the uncertain knowledge) but which allows the seller to accept a payment by installments, if this is the preferred payment mode for the customer (this is also part of the uncertain knowledge). Formally, a CAF is defined as follows :

**Definition 1** *Let  $\mathcal{L}$  be a language from which we can build arguments, and let  $\text{Args}(\mathcal{L})$  be the set which contains all those arguments. A Control Argumentation Framework (CAF) is a triple  $\text{CAF} = \langle F, C, U \rangle$  where  $F$  is the fixed part,  $U$  is the uncertain part and  $C$  is the control part of  $\text{CAF}$  with :*

1.  $F = \langle A_F, \rightarrow \rangle$  where  $A_F$  is a set of arguments that we know they belong to the system and  $\rightarrow \subseteq (A_F \cup A_U) \times (A_F \cup A_U)$  is an attack relation representing a set of attacks for which we are aware both of their existence and their direction.

2.  $U = \langle A_U, (\rightleftharpoons \cup \dashrightarrow) \rangle$  where  $A_U$  is a set of arguments for which we are not sure that they belong to the system,  $\rightleftharpoons \subseteq (((A_U \cup A_F) \times (A_U \cup A_F)) \setminus \rightarrow)$  is an attack relation representing a set of attacks for which we are aware of their existence but not of their direction, and  $\dashrightarrow \subseteq (((A_U \cup A_F) \times (A_U \cup A_F)) \setminus \rightarrow)$  is an attack relation representing a set of attacks for which we are not aware of their existence but we are aware of their direction, with  $\rightleftharpoons \cap \dashrightarrow = \emptyset$ .

3.  $C = \langle A_C, \Rightarrow \rangle$  where  $A_C$  is a set of arguments, called control arguments, that the agent can choose to use or not, and  $\Rightarrow \subseteq \{(a_i, a_j) \mid a_i \in A_C, a_j \in A_F \cup A_C \cup A_U\}$  is an attack relation.

$A_F, A_U$  and  $A_C$  are disjoint subsets of  $\text{Args}(\mathcal{L})$ .

A CAF features a set of distinct attack relations that capture different sorts of information. Its simplest part is  $\langle A_F, \rightarrow \cap (A_F \times A_F) \rangle$ , which is a classical AF that contains the indisputable knowledge of the agent on its opponent. The idea of CAFs essentially extends this basic argumentation framework with additional attack relations defined on arguments from the sets  $A_U$  and  $A_C$ . For instance, there is an attack  $(a_i, a_j) \in \rightleftharpoons$ , with  $a_i, a_j \in A_F$  when it is certain that the two arguments exist and are in conflict (e.g. because they make mutually exclusive claims), but the direction of the attack(s) is unknown (e.g. because of lack of information on the intrinsic strength of arguments, or on the preference relation between arguments). An attack  $(a_i, a_j) \in \dashrightarrow$ , with  $a_i \in A_U$  and  $a_j \in A_F$ , represents a situation where it is unknown whether  $a_i$  is present in the system (e.g. some of its premises could be false at the current time), but if  $a_i$  is in the system, then  $a_i$  definitely attacks  $a_j$ .

Central to controllability is the notion of *completion* of a CAF, i.e. a classical AF which is built by choosing one of the possible options for each uncertain argument or attack.

**Definition 2** [9] Given a CAF  $\mathcal{CAF} = \langle F, C, U \rangle$ , a completion of  $\mathcal{CAF}$  is an AF  $\mathcal{AF} = \langle A, R \rangle$ , s.t.

1.  $A = A_F \cup A_C \cup A_{comp}$  where  $A_{comp} \subseteq A_U$ ;
2. if  $(a, b) \in R$ , then  $(a, b) \in \rightarrow \cup \rightleftharpoons \cup \dashrightarrow \cup \Rightarrow$ ;
3. if  $(a, b) \in \rightarrow$ , then  $(a, b) \in R$ ;
4. if  $(a, b) \in \rightleftharpoons$  and  $a, b \in A$ , then  $(a, b) \in R$  or  $(b, a) \in R$ ;
5. if  $(a, b) \in \Rightarrow$  and  $a, b \in A$ , then  $(a, b) \in R$ .

Note that the definition of a completion leaves the attacks from  $\dashrightarrow$  unspecified, as these attacks may not appear in the theory. For some examples of completions, see [9].

*Controllability* means that we can select a subset  $A_{conf} \subseteq A_C$  and the corresponding attacks  $\{(a_i, a_j) \in \Rightarrow \mid a_i \in A_C, a_j \in (A_F \cup A_C \cup A_U)\}$  such that whatever the completion of  $\mathcal{CAF}$ , a given target is always reached. We focus on two kinds of targets : credulous acceptance of a set of arguments (this is reminiscent of extension enforcement [6]), and skeptical acceptance of a set of arguments.

**Definition 3** [9] A control configuration of a CAF  $\mathcal{CAF} = \langle F, C, U \rangle$  is a subset  $A_{conf} \subseteq A_C$ . Given a set of arguments  $T \subseteq A_F$  and a semantics  $\sigma$ , we say that  $T$  is skeptically (resp. credulously) reached by the configuration  $A_{conf}$  under  $\sigma$  if  $T$  is included in every (resp. at least one)  $\sigma$ -extension of every completion of  $\mathcal{CAF}' = \langle F, C', U \rangle$ , with  $C' = \langle A_{conf}, \{(a_i, a_j) \in \Rightarrow \mid a_i \in A_C, a_j \in (A_F \cup A_C \cup A_U)\} \rangle$ . We say that  $\mathcal{CAF}$  is skeptically (resp. credulously) controllable w.r.t.  $T$  and  $\sigma$ .

In a nutshell, CAFs are a powerful enabler of advanced negotiation techniques, that blend together a number of desirable features such as the qualitative representation of uncertainty, simultaneous reasoning with different profiles through completions, simultaneous consideration of both certain and uncertain knowledge of the opponent, the use of control arguments (corresponding to a persuasion phase embedded in negotiation, allowing for the reinstatement of rejected arguments), along with a computational model based on QBFs.

### 3 The Negotiation Framework

This section presents a new argumentation-based negotiation framework that relies on CAFs [9] for representing the incomplete information that agents have about their opponents. Agents communicate through the exchange of messages (or dialogue moves, see e.g. [11]). We assume that agents play the roles of the proponent and opponent in a turn-taking round-based protocol (e.g. similar to the alternating offers protocol of [14]), where a proponent initiates a round and passes the token to its opponent when it is unable to defend an offer rejected by the opponent. The opponent may accept an offer when one of the supporting arguments is an acceptable argument for it, or reject an offer if it cannot accept any of the different supporting arguments sent by the proponent. We build on the works of [1, 14], and in the following,  $\mathcal{L}$  denotes a logical language,

and  $\equiv$  an equivalence relation associated with it. From  $\mathcal{L}$ , a set  $\mathcal{O} = \{o_1, \dots, o_n\}$  of  $n$  offers is identified, such that  $\nexists o_i, o_j \in \mathcal{O}$  such that  $o_i \equiv o_j$ . This means that the offers are different. Offers correspond to the different alternatives (e.g. prices for a product) that can be exchanged during the negotiation dialogue. We assume that agents share the same set of offers  $\mathcal{O}$  but those offers can be supported by different practical arguments (although not necessarily) in the theories of the negotiating agents. By argument, we mean a *reason* in believing (called epistemic arguments) or doing something (called practical arguments). The set  $Args(\mathcal{L})$  is then divided into two subsets : a subset  $Args_p(\mathcal{L})$  of practical arguments supporting offers, and a subset  $Args_e(\mathcal{L})$  of epistemic arguments supporting beliefs. Thus,  $Args(\mathcal{L}) = Args_p(\mathcal{L}) \cup Args_e(\mathcal{L})$ . A negotiation theory is therefore represented as follows :

**Definition 4 (Negotiating agent theory)** Let  $\mathcal{O}$  be a set of  $n$  offers. A negotiating theory of an agent  $\alpha$  is a tuple  $\mathcal{T} = \langle \mathcal{O}, \mathcal{T}^\alpha, \mathcal{CAF}^{\alpha, \beta}, \mathcal{F}^\alpha \rangle$  with  $\mathcal{T}^\alpha = \langle A^\alpha, \rightarrow_\alpha \rangle$  and  $\mathcal{CAF}^{\alpha, \beta} = \langle F^{\alpha, \beta}, U^{\alpha, \beta}, C^{\alpha, \beta} \rangle$  and where : 1.  $A^\alpha \subseteq Args(\mathcal{L})$  is a set of arguments s.t.  $A^\alpha = A_p^\alpha \cup A_e^\alpha$  where  $A_p^\alpha$  is a set of practical arguments,  $A_e^\alpha$  a set of epistemic arguments, and  $A_e^\alpha \subseteq A_e^\alpha$  is the set of control arguments. For the attack relation it holds  $\rightarrow_\alpha = \rightarrow_p \cup \rightarrow_e \cup \rightarrow_m$ , with  $\rightarrow_p \subseteq A_p^\alpha \times A_p^\alpha$ , representing an attack relation for practical arguments,  $\rightarrow_e \subseteq A_e^\alpha \times A_e^\alpha$  representing an attack relation for epistemic arguments and  $\rightarrow_m \subseteq A_e^\alpha \times A_p^\alpha$  representing an attack relation between epistemic and practical arguments i.e.  $(a, \delta) \in \rightarrow_m$ , if  $a \in A_e^\alpha$  and  $\delta \in A_p^\alpha$  (see [2], [14]).

2.  $\mathcal{CAF}^{\alpha, \beta}$  is defined by :

- $F^{\alpha, \beta} = \langle A_F^{\alpha, \beta}, \rightarrow_{\alpha, \beta} \rangle$  with  $A_F^{\alpha, \beta} = A_{F_e}^{\alpha, \beta} \cup A_{F_p}^{\alpha, \beta}$ ,  $\rightarrow_{\alpha, \beta} = \rightarrow_e^{\alpha, \beta} \cup \rightarrow_p^{\alpha, \beta}$  and  $\langle A_{F_e}^{\alpha, \beta}, \rightarrow_e^{\alpha, \beta} \rangle$  defining the epistemic arguments subpart s.t.  $\rightarrow_e^{\alpha, \beta} \subseteq (A_{F_e}^{\alpha, \beta} \cup A_{U_e}^{\alpha, \beta}) \times (A_{F_e}^{\alpha, \beta} \cup A_{U_e}^{\alpha, \beta})$ . The above hold also for the practical arguments subpart. It also holds  $A_U^{\alpha, \beta} = A_{U_e}^{\alpha, \beta} \cup A_{U_p}^{\alpha, \beta}$ .
- $U^{\alpha, \beta} = \langle A_U^{\alpha, \beta}, \rightleftharpoons_{\alpha, \beta} \cup \dashrightarrow_{\alpha, \beta} \rangle$  with  $\rightleftharpoons_{\alpha, \beta} = \rightleftharpoons_e \cup \rightleftharpoons_p$ ,  $\dashrightarrow_{\alpha, \beta} = \dashrightarrow_e \cup \dashrightarrow_p$  and  $\langle A_{U_e}^{\alpha, \beta}, \rightleftharpoons_e \cup \dashrightarrow_e \rangle$ ,  $\rightleftharpoons_e \subseteq ((A_{U_e}^{\alpha, \beta} \cup A_{F_e}^{\alpha, \beta}) \times (A_{U_e}^{\alpha, \beta} \cup A_{F_e}^{\alpha, \beta})) \setminus \rightarrow_e^{\alpha, \beta}$ ,  $\dashrightarrow_e \subseteq ((A_{U_e}^{\alpha, \beta} \cup A_{F_e}^{\alpha, \beta}) \times (A_{U_e}^{\alpha, \beta} \cup A_{F_e}^{\alpha, \beta})) \setminus \rightarrow_e^{\alpha, \beta}$ , defining the epistemic arguments subpart. The same hold for the practical arguments subpart.  $\rightleftharpoons_e \cap \dashrightarrow_e = \emptyset$ .
- $C^{\alpha, \beta} = \langle A_C^\alpha, \Rightarrow \rangle$  where  $\Rightarrow \subseteq \{(a_i, a_j) \mid a_i \in A_C^\alpha \text{ and } a_j \in A_C^\alpha \cup A_{F_e}^{\alpha, \beta} \cup A_{U_e}^{\alpha, \beta}\} \setminus (\rightarrow_e^{\alpha, \beta} \cup \rightleftharpoons_e \cup \dashrightarrow_e)$ .

3.  $\mathcal{F}^\alpha : \mathcal{O} \rightarrow 2^{A_p^\alpha}$  s.t.  $\forall i, j$  with  $i \neq j$ ,  $\mathcal{F}^\alpha(o_i) \cap \mathcal{F}^\alpha(o_j) = \emptyset$ . Let  $A_p^\alpha = \bigcup \mathcal{F}^\alpha(o_i)$  with  $i = 1, \dots, n$ . This function returns the practical arguments supporting offers in  $\mathcal{O}$ .

In the following we present the different procedures that implement the new negotiation framework of this paper.

### 3.1 Best Offers Selection

Algorithm 1 is the procedure invoked by the proponent agent  $\alpha$  in order to compute, first, its best offer, based on its own theory, and it is implemented through function *comp\_next\_offer*. This function looks for the best offer supported by an acceptable practical argument by using a ranking on the supporting arguments based on a partial preorder (other methods can be also applied here). Then, based on its  $\mathcal{CAF}^{\alpha,\beta}$ , it computes the practical arguments that support this offer in its opponent theory and calls a procedure, implemented by algorithm 2, that selects the supporting argument to be sent. If the proponent agent has no (other) offer to propose, the opponent of the agent is informed by a suitable message (i.e. nothing).

---

**Algorithm 1** : choose-best-offer( $\mathcal{O}, \mathcal{T}^\alpha, \mathcal{CAF}^{\alpha,\beta}, \mathcal{F}^{\alpha,\beta}(o)$ )

---

```

1  $o \leftarrow \text{comp\_next\_offer}(\mathcal{O}, \mathcal{T}^\alpha)$ ;
2 if  $o \neq \emptyset$  then
3    $\mathcal{F}^{\alpha,\beta}(o) \leftarrow \text{compute\_sup\_arg}(o, A_{F_p}^{\alpha,\beta} \cup A_{U_p}^{\alpha,\beta})$ ;
4   call choose-support-arg( $o, \mathcal{F}^{\alpha,\beta}(o), \mathcal{CAF}^{\alpha,\beta}$ );
5 else
6   message( $\alpha, \beta$ )=nothing; send(message( $\alpha, \beta$ ));
```

---

### 3.2 Supporting Argument Selection

The algorithm described below, selects (through function *choose - arg*, where the choice can be random, as herein, or based on other methods) the argument that the proponent agent  $\alpha$  sends to its opponent agent to support its offer. Moreover, another procedure finds the arguments that defend this supporting argument whenever this argument is currently rejected by the opponent. This task is carried out by the procedure implemented by algorithm 3. If there is no other available argument that supports the current offer, the agent abandons this offer and passes the negotiation token to the opponent agent.

---

**Algorithm 2** : choose-support-arg( $o, \mathcal{F}^{\alpha,\beta}(o), \mathcal{CAF}^{\alpha,\beta}$ )

---

```

1 if  $\mathcal{F}^{\alpha,\beta}(o) \neq \emptyset$  then
2    $\theta \leftarrow \text{choose\_arg}(\mathcal{F}^{\alpha,\beta}(o))$ ;
3   call defend-offer( $o, \theta, \mathcal{F}^{\alpha,\beta}(o), \mathcal{CAF}^{\alpha,\beta}$ )
4 else
5    $\mathcal{O} = \mathcal{O} - \{o\}$ ;  $A_p^\alpha = A_p^\alpha - \mathcal{F}^\alpha(o)$ ;
6   message( $\alpha, \beta$ )=give_token; send(message( $\alpha, \beta$ ));
```

---

### 3.3 The Bidding Strategy

The *bidding strategy* of the proponent agent is implemented by algorithm 3. The main task here is to defend the proposed offer by an argument that (as said before) supports the offer in the opponent's theory. Consider for instance a

car seller agent who proposes an expensive luxury SUV of a prestigious brand to a customer who, as the agent understands, seems to afford it. The reason (argument) that the seller agent has chosen this particular car is probably the high sales commission that it brings. However, this is not an argument it can use to convince its customer. The pool of appropriate arguments could include the smooth ride, fast acceleration, high top speed, off-road capabilities, safety features, or even the high social status associated with the brand. In fact, the discovery of those arguments takes place inside algorithms 1 and 2. The role of the bidding strategy algorithm is to determine whether such a supporting argument is already acceptable in the opponent's theory, or to search for a *control configuration* that can defend the selected supporting argument under all possible opponent profiles.

---

**Algorithm 3** : defend-offer( $o, \theta, \mathcal{F}^{\alpha,\beta}(o), \mathcal{CAF}^{\alpha,\beta}$ )

---

```

1 if  $\theta$  is credulously accepted in all completions of the theory
    $A_F^{\alpha,\beta} \cup A_U^{\alpha,\beta}$ 
2 then
3   offer( $\alpha, \beta$ ) =  $\langle o, \theta, \langle \emptyset, \emptyset \rangle \rangle$ ;
4    $\mathcal{F}^{\alpha,\beta}(o) = \mathcal{F}^{\alpha,\beta}(o) - \{\theta\}$ ;
5   message( $\alpha, \beta$ )=offer( $\alpha, \beta$ ); send(message( $\alpha, \beta$ ))
6 else
7    $S \leftarrow \text{comp\_contr\_conf}(\mathcal{CAF}^{\alpha,\beta}, \theta)$ ;
8   if  $S \neq \emptyset$  then
9      $\mathcal{R} = \{(a_i, a_j) | a_i \in S, a_j \in A_F^{\alpha,\beta} \cup A_U^{\alpha,\beta}\}$ ;
10    offer( $\alpha, \beta$ ) =  $\langle o, \langle \theta, \langle S, \mathcal{R} \rangle \rangle \rangle$ ;
11    message( $\alpha, \beta$ )=offer( $\alpha, \beta$ );
12    send(message( $\alpha, \beta$ ))
13  else
14     $\mathcal{F}^{\alpha,\beta}(o) = \mathcal{F}^{\alpha,\beta}(o) - \{\theta\}$ ;
15    call choose-support-arg( $o, \mathcal{F}^{\alpha,\beta}(o), \mathcal{CAF}^{\alpha,\beta}$ );
```

---

More precisely, acceptance in the context of incomplete theories is based on the notion of *completion* which represents a possible profile (see definition 2). The computation in line 1 of the algorithm relies on reasoning with Quantified Boolean Formulas (QBFs), as described in [9], that is carried out by the *quantom* solver [25]. For details about the logical encoding, we refer the reader to [10].

Now, in case this computation succeeds,  $\theta$  is acceptable in all possible opponent profiles (completions), and agent  $\alpha$  sends to agent  $\beta$  the offer  $o$ , along with  $\theta$ .

In case  $\theta$  is not acceptable wrt the above theory, agent  $\alpha$  reacts as depicted in lines 7-13 of algorithm 3. First, it uses its *CAF* to seek a *control configuration*  $S$ , that defends  $\theta$ . This is again a problem on QBFs that is solved by a call to *quantom* solver (line 7 of the algorithm).

Now if in this second case the call succeeds, agent  $\alpha$  sends offer  $o$  to agent  $\beta$ , along with the supporting argument  $\theta$ , the set of arguments  $S$ , and the associated attacks  $R$ . Otherwise, the agent abandons this argument and picks another from  $\mathcal{F}^{\alpha,\beta}(o)$  in order to continue defending  $o$ . This is

done by function *choose-support-arg*. Recall that our approach looks for sets of arguments that are control configurations, i.e. work for all possible profiles of agent  $\beta$ . However, if there is no such solution, the QBF based techniques of *quantom* [25], can find sets of arguments that work for most of these profiles.

In the following we define an operator  $\oplus$  that is used in algorithms 4 and 5.

**Definition 5** *Let  $A_1, A_2, A_3$  be sets. We define  $(A_1, A_2) \oplus A_3$  as the pair  $(A'_1, A'_2)$  such that  $A'_1 = A_1 \setminus (A_1 \cap A_3)$  and  $A'_2 = A_2 \cup (A_1 \cap A_3)$ .*

At the beginning of the negotiation each agent has in its theory (i.e.  $A^\alpha$  and  $A^\beta$  respectively) only a part of the possible epistemic arguments (wrt a specific application). That means that some arguments are in  $A^\alpha$  and not in  $A^\beta$  (and vice-versa). However, when an agent will use arguments (and the associated attacks) that do not belong to the opponent's theory, the opponent agent will add them (as well as the associated attacks) in its own theory, and it will be able to use them from that point onwards in the negotiation. This situation may take place in the algorithms 4 and 5.

### 3.4 The Acceptance Strategy

This section discusses Algorithm 4, that implements the *acceptance strategy* of an agent. Upon receiving an offer and its supporting arguments (and the associated attacks) sent by a proponent agent, the algorithm updates the theory as well as the  $\mathcal{CAF}$  of the receiving agent by integrating the supporting arguments, the defending arguments (i.e. the control configuration), and the associated attacks into both theories (i.e. the receiving agent own theory and its  $\mathcal{CAF}$ ). Then, the receiver agent either accepts the offer (i.e. if the supporting arguments are acceptable) and informs the proponent accordingly, or sends to the proponent the reasons for rejecting its offer.

### 3.5 The Negotiation Protocol

The algorithm 5 described below implements the core procedure that drives the overall negotiation between the two negotiating agents through the necessary updates of their negotiation theories and calls to appropriate functions. The first part of algorithm (lines 1-2) implements the behavior of an agent when it is the proposer of the first offer, whereas the second part (lines 3-24) is concerned with its reaction when it receives an answer from another agent (i.e. the opponent). While the first part is straightforward as it concerns the selection of the best offer to propose, the second part is more involved and breaks down to several sub-cases. Those cases concern different situations that may arise during a negotiation, such as the rejection of an offer by the opponent, the acceptance of an offer (that terminates the negotiation with an agreement), the situation where the opponent informs that it has no other offer to propose, the situation where the opponent responds that it has no offer to propose too in a received similar message by the

---

**Algorithm 4** : decide-upon-offer( $\mathcal{T}^\alpha, \mathcal{CAF}^{\alpha,\beta}$ , offer( $\beta, \alpha$ ))

---

```

1  $\langle o, \theta, \langle S, \mathcal{R} \rangle \rangle = \text{offer}(\beta, \alpha);$ 
2 if  $S \neq \emptyset$  then
3    $\mathcal{T}^\alpha = (A^\alpha \cup S, \rightarrow_\alpha \cup \mathcal{R});$ 
4    $(A_U^{\alpha,\beta}, A_F^{\alpha,\beta}) = (A_U^{\alpha,\beta}, A_F^{\alpha,\beta}) \oplus S;$ 
5    $(\neg\rightarrow_{\alpha,\beta}, \rightarrow_{\alpha,\beta}) = (\neg\rightarrow_{\alpha,\beta}, \rightarrow_{\alpha,\beta}) \oplus \mathcal{R};$ 
6    $(\rightleftharpoons_{\alpha,\beta}, \rightarrow_{\alpha,\beta}) = (\rightleftharpoons_{\alpha,\beta}, \rightarrow_{\alpha,\beta}) \oplus \mathcal{R}$ 
7 if  $\theta$  is a credulous conclusion of theory  $\mathcal{T}^\alpha$  then
8   message( $\alpha, \beta$ ) = Accept( $o$ );
9   send(message( $\alpha, \beta$ ))
10 else
11   Compute  $Q \subseteq \mathcal{E}$  where  $\mathcal{E}$  is an extension of  $\mathcal{T}^\alpha$  and  $Q$  is
the set of arguments from which  $\theta$  is reachable in the
attack graph;
12   Reasons =  $\{(p, \theta) \mid (p, \theta) \in \rightarrow_\alpha \text{ and } p \in Q\};$ 
13   message( $\alpha, \beta$ ) = Reject( $o, \theta, \langle Q, \text{Reasons} \rangle$ );
14   send(message( $\alpha, \beta$ ));
```

---

(proponent) agent (this ends the negotiation without agreement), the situation where an agent informs that it gives the token, and the situation where an offer is received and the receiver agent has to decide upon its acceptance or rejection. A detailed example of a negotiation is given in [10].

## 4 Experimental evaluation

Our framework has been implemented by using the JADE (<http://jade.tilab.com/>) platform and evaluated on negotiations with random argumentation theories.

### 4.1 Random Theory Generation

The experimental evaluation of the proposed framework is based on a system, implemented in Java, that generates pairs of random negotiation theories and associated CAFs, with different user specified characteristics.

Each negotiation experiment involves a pair of random theories  $\mathcal{T}^\alpha = \langle A^\alpha, \rightarrow_\alpha \rangle$  and  $\mathcal{T}^\beta = \langle A^\beta, \rightarrow_\beta \rangle$  that share a common part, i.e. there exists  $N_{\alpha,\beta} = \langle A^{N_{\alpha,\beta}}, \rightarrow_{N_{\alpha,\beta}} \rangle$ , such that  $A^{N_{\alpha,\beta}} = A^\alpha \cap A^\beta$  and  $(a, b) \in \rightarrow_{N_{\alpha,\beta}}$  iff  $(a, b) \in \rightarrow_\alpha \cap \rightarrow_\beta$ . Moreover, control arguments are only attacked by other control arguments, i.e.  $((A^\alpha \setminus A_C^\alpha) \times A_C^\alpha) \cap \rightarrow_\alpha = \emptyset$ .

The structure of the generated theories depends on a number of user supplied parameter values that are explained briefly below.

The user inputs the number of epistemic, practical and control arguments of theories  $\mathcal{T}^\alpha$  and  $\mathcal{T}^\beta$ , as well as their density, defined as the ratio of attacks present in the theory to the number of all possible attacks between the arguments of the theory. Moreover, the instance generation system receives as input the number of epistemic, practical and control arguments of the shared part  $N_{\alpha,\beta}$ .

From theory  $\mathcal{T}^\beta$ , the CAF  $\mathcal{CAF}^{\alpha,\beta} = \langle \langle A_F^{\alpha,\beta}, \rightarrow_{\alpha,\beta} \rangle, \langle A_U^{\alpha,\beta}, \rightleftharpoons_{\alpha,\beta} \cup \neg\rightarrow_{\alpha,\beta} \rangle, \langle A_C^\alpha, \Rightarrow \rangle \rangle$  is built (similarly for

---

**Algorithm 5** : Procedure  $\text{negotiate}(\langle \mathcal{O}, \mathcal{T}^\alpha, \mathcal{CAF}^{\alpha,\beta}, \mathcal{F}^\alpha \rangle)$

---

```

1 if agent  $\alpha$  proposes first then
2   call choose-best-offer( $\mathcal{O}, \mathcal{T}^\alpha, \mathcal{CAF}^{\alpha,\beta}, \mathcal{F}^{\alpha,\beta}(o)$ );
3 while true do
4   get message( $\beta, \alpha$ );
5   switch message( $\beta, \alpha$ ) do
6     case  $\text{Reject}(o, \theta, \langle Q, \text{Reasons} \rangle)$  do
7        $(A_U^{\alpha,\beta}, A_F^{\alpha,\beta}) = (A_U^{\alpha,\beta}, A_F^{\alpha,\beta}) \oplus Q$ ;
8        $(\rightarrow_{\alpha,\beta}, \rightarrow_{\alpha,\beta}) = (\rightarrow_{\alpha,\beta}, \rightarrow_{\alpha,\beta}) \oplus \text{Reasons}$ ;
9        $(\rightleftarrows_{\alpha,\beta}, \rightarrow_{\alpha,\beta}) = (\rightleftarrows_{\alpha,\beta}, \rightarrow_{\alpha,\beta}) \oplus \text{Reasons}$ ;
10      call defend-offer( $o, \theta, \mathcal{F}^{\alpha,\beta}(o), \mathcal{CAF}^{\alpha,\beta}$ );
11     case  $\text{Accept}(o)$  do
12       End of negotiation with agreement on offer
13        $o$ 
14     case nothing do
15       if  $\mathcal{O} \neq \emptyset$  then
16         call choose-best-
17         offer( $\mathcal{O}, \mathcal{T}^\alpha, \mathcal{CAF}^{\alpha,\beta}, \mathcal{F}^{\alpha,\beta}(o)$ );
18       else
19         answer( $\alpha, \beta$ )=nothing_too;
20         send(answer( $\alpha, \beta$ ))
21     case nothing_too do
22       End of negotiation without agreement
23     case give_token do
24       call choose-best-
25       offer( $\mathcal{O}, \mathcal{T}^\alpha, \mathcal{CAF}^{\alpha,\beta}, \mathcal{F}^{\alpha,\beta}(o)$ );
26     case offer( $\beta, \alpha$ )= $\langle o, \langle \theta, \langle S, \mathcal{R} \rangle \rangle$  do
27       call decide-upon-offer( $\mathcal{T}^\alpha, \mathcal{CAF}^{\alpha,\beta},$ 
28       offer( $\beta, \alpha$ ));

```

---

$\mathcal{T}^\alpha$  and  $\mathcal{CAF}^{\beta,\alpha}$ ), which is the theory that agent  $\alpha$  holds about agent  $\beta$ .  $\mathcal{CAF}^{\alpha,\beta}$  satisfies the following conditions (a)  $A_F^{\alpha,\beta} \cup A_U^{\alpha,\beta} = A^\beta \cup A_p^\alpha$ , (b)  $A_p^\beta \subseteq A_F^{\alpha,\beta}$ .

The attack relation  $\rightarrow_{\alpha,\beta} \cup \rightleftarrows_{\alpha,\beta} \cup \dashrightarrow_{\alpha,\beta}$  of  $\mathcal{CAF}^{\alpha,\beta}$ , is generated so that it satisfies the following conditions : a)  $\rightarrow_{\alpha,\beta} \subseteq \rightarrow_\beta$ , b)  $\rightarrow_\beta \cap (A_F^{\alpha,\beta} \times A_F^{\alpha,\beta}) \subseteq \rightarrow_{\alpha,\beta}$ , c)  $(\rightleftarrows_{\alpha,\beta} \cup \dashrightarrow_{\alpha,\beta}) \subseteq (\rightarrow_\beta \setminus \rightarrow_{\alpha,\beta})$ , and d)  $\rightleftarrows_{\alpha,\beta} \cap \dashrightarrow_{\alpha,\beta} = \emptyset$ .

The main consequence of the above requirements is that the attack relation of  $\mathcal{CAF}^{\alpha,\beta}$  is a subset of the attack relation of  $\mathcal{T}^\beta$ . The rationale for this restriction, in this initial experimental evaluation, is to focus on negotiation experiments where agents possess an "accurate" model of their opponent. One way to formalize the model accuracy is via the above relation between individual theories and CAFs. Moreover, it is interesting to study how the framework behaves when this restriction is removed. Indeed, the next section provides initial evidence that the method of this paper can cope with the relaxation of this restriction.

As with the individual agent theories  $\mathcal{T}^\alpha$  and  $\mathcal{T}^\beta$ , the

random instance generation software accepts as input a number of parameter values that determine various features of the CAFs of the agents. Most of them concern the uncertainty of an agent profile on its opponent, as captured by the corresponding CAF. The first is parameter  $\text{rateUncertArgs}$  that defines the ratio of uncertain arguments to all (fixed and uncertain) arguments of the theory. That is,  $\text{rateUncertArgs} = |A_U^{\alpha,\beta}| / |A_F^{\alpha,\beta} \cup A_U^{\alpha,\beta}|$  for agent  $\alpha$ , and similarly for agent  $\beta$ . Other parameters of the system include  $\text{rateUncertAtt}$ , that defines the ratio of uncertain attacks over all attacks, as well as  $\text{rateUnidirAtt}$  that defines the ratio of undirected attacks to all attacks. That is,  $\text{rateUncertAtt} = | \dashrightarrow_{\alpha,\beta} | / | \rightarrow_{\alpha,\beta} \cup \rightleftarrows_{\alpha,\beta} \cup \dashrightarrow_{\alpha,\beta} |$ , and  $\text{rateUnidirAtt} = | \rightleftarrows_{\alpha,\beta} | / | \rightarrow_{\alpha,\beta} \cup \rightleftarrows_{\alpha,\beta} \cup \dashrightarrow_{\alpha,\beta} |$ . Moreover, parameter  $\text{densContrAtt}$  defines the ratio of attacks from the control arguments of the agent to the arguments of its opponent that are included in its CAF to all possible such attacks from control arguments. For instance,  $\text{densContrAtt} = 0.1$  for  $\mathcal{CAF}^{\alpha,\beta}$ , means that 10% of all possible attacks from arguments of  $A_c^\alpha$  to arguments in  $A_F^{\alpha,\beta} \cup A_U^{\alpha,\beta}$  are included in the particular  $\mathcal{CAF}^{\alpha,\beta}$ . Finally, the instance generation system receives as input the number of offers, i.e.  $|O^\alpha|$  and  $|O^\beta|$ , and the number of practical arguments that support each offer.

## 4.2 Experimental Results

This section reports on selected results of the experimental evaluation of the framework. As the negotiation theory generation system accepts several parameter values, it is outside the scope of this work to provide exhaustive experimental results for all possible value combinations. Instead, we present results for selected runs that reveal important factors that influence the working of the negotiation algorithm, and highlight its merits and limitations. In all experiments we fix  $|A^\alpha| = |A^\beta| = 40$ ,  $|A_p^\alpha| = |A_p^\beta| = 6$ ,  $|O^\alpha| = |O^\beta| = 4$  and  $A_c^\alpha \cap A^{N_{\alpha,\beta}} = A_c^\beta \cap A^{N_{\alpha,\beta}} = \emptyset$ .

The experimental evaluation is centered around 12 sets of agent theories, and associated CAFs, that differ in the uncertainty of these CAFs and the size of the shared part of agent theories. More specifically, four (4) combinations of parameter values concerning the CAFs are considered, same for both agents. The first combination, abbreviated as *comb1*, is determined by the values  $\text{rateUncertArgs} = 0$ ,  $\text{rateUnidirAtt} = 0$ ,  $\text{rateUncertAtt} = 0$  which correspond to the case where both agents have complete knowledge of their opponent. Then, *comb2* is defined by the values  $\text{rateUncertArgs} = 0.10$ ,  $\text{rateUnidirAtt} = 0.5$ ,  $\text{rateUncertAtt} = 0.5$ . Moreover, the third combination *comb3* is  $\text{rateUncertArgs} = 0.25$ ,  $\text{rateUnidirAtt} = 0.125$ ,  $\text{rateUncertAtt} = 0.125$ . Finally, the last combination *comb4* is  $\text{rateUncertArgs} = 0.50$ ,  $\text{rateUnidirAtt} = 0.25$ ,  $\text{rateUncertAtt} = 0.25$  and is the case where the agents have the highest uncer-



tainty about their opponents among all the experiments. Each of the above set of values for the 3 CAF parameters is combined with one of the three possible values  $\{0.25, 0.5, 0.75\}$  for the ratio  $|A^{N_{\alpha,\beta}}|/|A^\alpha|$  that capture different degrees of similarity between agent theories. Each row of Tables 1 and 2 presents the *agreement rate*, (i.e. ratio of the number of negotiations terminated with agreement over the total number of negotiations) of 600 negotiations consisting of 50 randomly generated experiments for each of the 12 parameter values combinations described above. Therefore, each experiment is an amalgamation of negotiation theories of various types as far as the values of the 12 value parameters is concerned. Each row of Table 1 corresponds to an experiment (600 negotiations) where the number of control arguments is shown in the `numContrArg` column, whereas the value of parameter `densContrAtt` in the corresponding column. The last two columns refer to the agreement rates achieved when the *density* of the individual theories of the agents participating in the negotiations is fixed to 0.15 (column "Agr 0.15") and 0.2 (column "Agr 0.2") respectively. The first row corresponds to the case where none of the agents has any control arguments.

The first conclusion that can be readily drawn from Table 1 is that control arguments increases significantly the number of negotiations that terminate with agreement. Indeed, for theories with density 0.15 (column "Agr 0.15"), the agreement rate almost doubles from 0.23 to 0.44 for cases where there are relatively few control arguments and attacks from those arguments, and triples to 0.65 in the experiments with the highest number of control arguments and attacks. Similar are the results when the density of the individual theories of the participating agents is set to 0.2 (column "Agr 0.2"). Observe that the slight increase of the density leads to a decrease in the rate of agreements in all cases. However, again the presence of control arguments increases the agreement rate from 0.16 to as high as 0.56.

numContrArg	densContrAtt	Agr 0.15	Agr 0.2
0	0	0.23	0.16
3	0.03	0.44	0.39
3	0.05	0.46	0.44
3	0.1	0.57	0.49
3	0.2	0.60	0.52
6	0.03	0.58	0.52
6	0.05	0.59	0.50
6	0.1	0.65	0.56
6	0.2	0.58	0.54

TABLE 1 – Agreement rate for negotiations with individual theories of density 0.15 and 0.20

Recall that the negotiation experiments are generated so that  $A_F^{\alpha,\beta} \cup A_U^{\alpha,\beta} = A^\beta \cup A_p^\alpha$  i.e. agent  $\alpha$  CAF about  $\beta$  contains all the arguments of its opponents. In the experiments of Table 2 this assumption is removed by allo-

wing agent  $\beta$  to possess arguments that are not part of the CAF of agent  $\alpha$ . The number of these arguments is determined by the value of parameter `unknown` defined as  $|A^\beta - (A_F^{\alpha,\beta} \cup A_U^{\alpha,\beta})|/|(A_F^{\alpha,\beta} \cup A_U^{\alpha,\beta})|$ . In the experiments of Table 2 this value is set to 0.25 with the effect of a decrease in the agreement rate when compared to the case with no unknown arguments. This decrease was less significant for theories with more control attacks. The ex-

numContrArg	densContrAtt	Agreement
3	0.03	0.32
3	0.05	0.37
3	0.1	0.42
3	0.2	0.43
6	0.03	0.45
6	0.05	0.43
6	0.1	0.55
6	0.2	0.57

TABLE 2 – Agreement rate for negotiations with individual theories of density 0.15 and `unknown=0.25`

perimental evaluation leads to a number of general conclusions. The first is that, not surprisingly, the effectiveness of the approach wrt the rate of agreements depends on a number of parameters including the density of the individual theories, the number of attacks from control arguments, etc. Moreover, other experiments not reported here, have shown that the agreement rate also depends on the size of the shared part  $N_{\alpha,\beta}$ . In all cases it seems that, for "reasonably good" opponent profiles, the method leads to a significant increase in the number of negotiations that terminate with agreement.

## 5 Related work and Conclusions

In this paper we presented an original argumentation-based negotiation framework that exploits a recent work proposed in [9] on control argumentation frameworks for modeling the uncertainty about the opponent profile and also the acceptance and bidding strategies of the negotiating agents. Compared to previous works proposed in the literature on argumentation-based negotiation (see e.g. [1],[3],[18],[13],[22],[14],[19],[20]) this new framework introduces and combines together a number of original ideas, with most notable a qualitative representation of uncertainty that enables simultaneous consideration of several different profiles, the bidding strategy that allows an agent to use arguments that do not belong to its theory, along with the notion of control arguments that facilitates persuasion and utilizes arguments that defend against all the possible attacks at once, hence minimizing the number of exchanged messages. We consider that our work generalizes the works proposed in [15],[8]. Our work is also different from the work proposed in [23] where the agents have an incomplete theory on the opponent which evolves based on

the information contained in the exchanged offers during the negotiation through classical belief revision. The bidding strategy also used in this work is different to ours. Our experimental results have shown that the outcome of an argumentation-based negotiation dialogue depends on different parameters of the argumentation theories of the agents but in all cases the use of control arguments seems to have a positive impact on the number of agreements.

## Références

- [1] L. Amgoud, Y. Dimopoulos, and P. Moraitis. A unified and general framework for argumentation-based negotiation. In *Proc. AAMAS 2007*, page 158, 2007.
- [2] L. Amgoud, Y. Dimopoulos, and P. Moraitis. Making decisions through preference-based argumentation. In *Proc. KR 2008*, pages 113–123, 2008.
- [3] L. Amgoud and S. Kaci. On the study of negotiation strategies. In *Proc. Agent Communication II*, pages 150–163, 2006.
- [4] T. Baarslag, M. J. C. Hendriks, K. V. Hindriks, and C. M. Jonker. Learning about the opponent in automated bilateral negotiation : a comprehensive survey of opponent modeling techniques. *Autonomous Agents and Multi-Agent Systems*, 30(5) :849–898, 2016.
- [5] T. Baarslag, M. J. C. Hendriks, K. V. Hindriks, and C. M. Jonker. A survey of opponent modeling techniques in automated negotiation. In *Proc. AAMAS 2016*, pages 575–576. ACM, 2016.
- [6] R. Baumann and G. Brewka. Expanding argumentation frameworks : Enforcing and monotonicity results. In *Proc. COMMA 2010*, pages 75–86, 2010.
- [7] E. Black and K. Atkinson. Choosing persuasive arguments for action. In *Proc. AAMAS 2011*, pages 905–912, 2011.
- [8] E. Bonzon, Y. Dimopoulos, and P. Moraitis. Knowing each other in argumentation-based negotiation. In *Proc. AAMAS 2012*, pages 1413–1414, 2012.
- [9] Y. Dimopoulos, J.-G. Mailly, and P. Moraitis. Control argumentation frameworks. In *Proc. AAI 2018*, pages 4678–4685, 2018.
- [10] Y. Dimopoulos, J.-G. Mailly, and P. Moraitis. Argumentation-based negotiation with incomplete opponent profiles. In *Proc. AAMAS 2019*, pages 1252–1260, 2019.
- [11] Y. Dimopoulos and P. Moraitis. Advances in argumentation-based negotiation. In *Negotiation and Argumentation in Multi-Agent Systems : Fundamentals, Theories, Systems and Applications*, pages 82–125, 2014.
- [12] P. M. Dung. On the acceptability of arguments and its fundamental role in nonmonotonic reasoning, logic programming and n-person games. *Artificial Intelligence*, 77 :321–357, 1995.
- [13] P. M. Dung, P. M. Thang, and F. Toni. Towards argumentation-based contract negotiation. In *Proc. COMMA 2008*, pages 134–146, 2008.
- [14] N. Hadidi, Y. Dimopoulos, and P. Moraitis. Argumentative alternating offers. In *Proc. AAMAS 2010*, pages 441–448, 2010.
- [15] N. Hadidi, Y. Dimopoulos, and P. Moraitis. Tactics and concessions for argumentation-based negotiation. In *COMMA12*, pages 285–296, 2012.
- [16] C. Hadjinikolis, Y. Siantos, S. Modgil, E. Black, and P. McBurney. Opponent modelling in persuasion dialogues. In *Proc. IJCAI 2013*, pages 164–170, 2013.
- [17] A. Hunter. Modelling the persuadee in asymmetric argumentation dialogues for persuasion. In *Proc. IJCAI 2015*, pages 3055–3061, 2015.
- [18] A. Kakas and P. Moraitis. Adaptive agent negotiation via argumentation. In *Proc. AAMAS 2006*, pages 384–391, 2006.
- [19] O. Marey, J. Bentahar, E. K. Asl, M. Mbarki, and R. Dssouli. Agents’ uncertainty in argumentation-based negotiation : Classification and implementation. In *Proc. ANT 2014*, pages 61–68, 2014.
- [20] A. Monteserin and A. Amandi. A reinforcement learning approach to improve the argument selection effectiveness in argumentation-based negotiation. *Expert Syst. Appl.*, 40(6) :2182–2188, 2013.
- [21] N. Oren and T. J. Norman. Arguing using opponent models. In *ArgMAS09*, pages 160–174, 2009.
- [22] S. Parsons, C. Sierra, and N. R. Jennings. Agents that reason and negotiate by arguing. *J. Log. Comput.*, 8(3) :261–292, 1998.
- [23] P. Pilotti, A. Casali, and C. I. Chesñevar. A belief revision approach for argumentation-based negotiation agents. *Applied Mathematics and Computer Science*, 25(3) :455–470, 2015.
- [24] I. Rahwan, S. D. Ramchurn, N. R. Jennings, P. McBurney, S. Parsons, and L. Sonenberg. Argumentation-based negotiation. *Knowledge Eng. Review*, 18(4) :343–375, 2003.
- [25] S. Reimer, M. Sauer, P. Marin, and B. Becker. QBF with soft variables. In *Proc. AVOCS 2014*, 2014.
- [26] T. Rienstra, M. Thimm, and N. Oren. Opponent models with uncertainty for strategic argumentation. In *Proc. IJCAI 2013*, pages 332–338, 2013.
- [27] R. Riveret, A. Rotolo, G. Sartor, H. Prakken, and B. Roth. Success chances in argument games : a probabilistic approach to legal disputes. In A. R. Lodder and L. Mommers, editors, *Proc. JURIX 2007*, volume 165, pages 99–108. IOS Press, 2007.
- [28] F. Zafari and F. N. Mofakham. POPPONENT : highly accurate, individually and socially efficient opponent preference model in bilateral multi issue negotiations (extended abstract). In *Proc. IJCAI 2017*, pages 5100–5104, 2017.

# Exploration et couverture par stigmergie d'un environnement inconnu avec une flotte de robots autonomes réactifs

Nicolas Gauville<sup>1,2</sup>

François Charpillet<sup>1</sup>

<sup>1</sup> Université de Lorraine, CNRS, Inria, LORIA, F-54000 Nancy

<sup>2</sup> Safran Electronics & Defense

nicolas.gauville@loria.fr

## Résumé

*L'exploration autonome d'un environnement inconnu peut être envisagée de différentes manières. On peut notamment citer les approches par frontières, où des robots sont affectés à des zones inexplorées de la carte. Ces dernières méthodes sont efficaces mais nécessitent de partager une carte, globaliser les décisions d'affectation. Les approches Brick and Mortar, quant à elles, utilisent un marquage au sol avec une prise de décision locale, mais donnent des performances beaucoup moins intéressantes. L'algorithme présenté ici est un compromis entre ces deux approches, permettant une prise de décision locale et, de façon surprenante, des performances proche des approches par frontières globales. Nous proposons également une étude comparative de la performance des trois différentes approches : Brick & Mortar, frontières globales et frontières locales. Notre algorithme local est également complet pour le problème d'exploration et peut être facilement distribué sur des robots avec une perte de performance mineure.*

## Mots Clef

Multi-robot, stigmergie, exploration

## Abstract

*Different approaches exist for multi-robot autonomous exploration. These include frontier approaches, where robots are assigned to unexplored areas of the map, which provide good performance but require sharing the map and centralizing decision-making. The Brick and Mortar approaches, on the other hand, use a ground marking with local decision-making, but give much lower performance. The algorithm presented here is a trade-off between these two approaches, allowing local decision-making and, surprisingly, performances are closed to centralized frontier approaches. We also propose a comparative study of the performance of the three different approaches : Brick & Mortar, Global Frontiers and Local Frontiers. Our local algorithm is also complete for the exploration problem and can be easily distributed on robots with a minor loss of performance.*



FIGURE 1 – Robot Minirex (utilisé pour le projet Cart-O-Matic) explorant une zone test.

## 1 Introduction

L'exploration autonome d'un environnement inconnu est un défi important en robotique mobile. L'exploration peut être une fin en soi : des robots de nettoyage (aspirateur, tondeuse à gazon, etc.) explorent systématiquement chaque zone à couvrir. Dans les applications plus sophistiquées, il s'agit principalement d'acquérir des informations : cartographier l'environnement [2], rechercher un intrus, localiser une source sonore ou olfactive [14, 17], ou encore rechercher d'éventuelles personnes à secourir dans un bâtiment en feu, ...

L'objectif de cet article est de proposer une approche par frontières *locales* pour aborder le problème de l'exploration multi-robot. L'objectif principal de ce travail est de proposer une approche évolutive et efficace, même dans un environnement où la communication est restreinte, pour résoudre le problème de l'exploration avec un grand nombre de robots. Nous avons également essayé de fournir un algorithme aussi simple que possible, exécutable avec des ressources limitées.

Notre approche, bien qu'inspirée des approches par frontières globales, aborde différemment le problème de l'ex-

ploration : dans le cas de l'approche globale, c'est avant tout un problème de planification, où l'algorithme doit décider de l'affectation des robots aux frontières connues. Dans notre version locale, les robots ne connaissent pas le nombre ou la position des autres robots et ne s'en soucient pas. Ils réagissent uniquement aux informations locales. La collaboration entre robots se fait par stigmergie, dans notre cas simulée à l'aide d'une carte qui est partagée entre robots à intervalles réguliers, mais d'autres méthodes pourraient être envisagées pour limiter au mieux la quantité et la fréquence des informations partagées [11].

## 2 État de l'art

L'exploration multi-robot vise à explorer un environnement inconnu le plus rapidement possible pour une équipe homogène de robots mobiles. C'est une question proche du problème de couverture [10], mais avec une carte inconnue.

Différentes approches ont été développées en fonction des contraintes. On y retrouve notamment les approches de type *Brick & Mortar* [8, 9, 1] qui peuvent être classées dans les « algorithmes fourmis », où les robots partagent une carte commune pour *marquer* l'environnement avec des traces similaires aux phéromones des insectes sociaux (à la différence que ces traces ne changent pas avec le temps). Les décisions sont alors prises localement, mais le partage d'une carte globale entre robots est nécessaire pour partager les traces virtuelles laissées sur la carte.

Des approches « purement fourmis » existent aussi, utilisant des phéromones dispersées par l'environnement pour répartir les robots [6]. Ces approches sont souvent plus difficiles à mettre en place que les approches *Brick & Mortar*, car il faut également faire évoluer les phéromones déposées (évaporation, dispersion, ...). Garnier et al. ont par exemple proposé un système de projection de points avec des robots équipés de capteurs de lumière pour simuler ces phéromones partagées [11, 15], mais cela nécessite d'avoir un appareil pouvant projeter des points sur l'ensemble de la surface à explorer, ce qui n'est pas envisageable dans le cas de l'exploration d'un bâtiment inconnu. Nous préférons donc ici les approches *Brick & Mortar* où il suffit de partager les traces laissées.

L'exploration peut également être considérée comme un problème d'ordonnement; l'approche par frontières [18, 7, 5] consiste à attribuer à chaque robot une frontière, c'est-à-dire une limite entre les parties explorées et inexplorées de la carte. Dans ce cas, un consensus global doit également être assuré.

Si les approches locales sont efficaces, elles restent beaucoup plus lentes que les approches par frontières en termes de temps d'exploration : les robots doivent effectuer de nombreux marquages pour ne pas se bloquer mutuellement ou « oublier » une zone à explorer. D'autre part, les approches par frontières sont plus efficaces et basées sur des algorithmes plus simples, mais nécessitent une coordination des décisions entre robots qui peut poser différents

problèmes dans des situations réelles où une communication constante ne peut être assurée [12]. De plus, l'algorithme d'assignation des frontières utilisé nécessite souvent un temps de calcul non linéaire en fonction du nombre de robots et de la taille de la grille, ce qui peut la rendre inapplicable dans une situation réelle si la carte est trop grande ou s'il y a trop de robots. Cependant, il existe différentes méthodes pour optimiser l'affectation et les données échangées, par exemple en partageant uniquement les frontières et non la carte entière [13]. Enfin, notre approche locale peut facilement être distribuée, et fonctionne également avec des communications limitées (voir section 5).

L'algorithme proposé ici est un compromis entre l'approche *Brick & Mortar* et l'assignation de frontières : c'est une approche de frontières locales, où les robots partagent une carte, mais avec une prise de décision locale pour chaque robot. Notre approche est donc moins sensible à l'asynchronisme et réduit significativement les calculs effectués par chaque robot, car le nombre de cellules prises en compte à un moment donné est beaucoup moins important.

Ce travail fait suite au projet *Cart-O-Matic* [3]. *Cart-O-Matic* était l'un des cinq projets fondés par l'Agence Nationale de la Recherche (ANR) pour sa participation au concours de robotique organisé par la *Délégation générale pour l'armement* (DGA). Ce concours intitulé « *Defi CAROTTE* » avait pour objectif de définir un système robotique capable d'explorer un environnement intérieur inconnu et identifier des objets localisés dans cet environnement. Notre projet *Cart-O-Matic* a choisi de déployer un système multi-robot. *Cart-O-Matic* a remporté le concours final en 2012.

Toutes les approches considérées ici utilisent des grilles d'occupation pour représenter le monde, et considèrent que chaque robot est équipé d'un Lidar principalement utilisé pour la localisation et la cartographie simultanées (SLAM). Cependant, notre approche, comme celle par frontières globales, peut être adaptée à d'autres représentations du monde [16].

### 2.1 Frontières globales

Dans le cas de l'approche par frontières globales, chaque robot possède une grille d'occupation pouvant prendre trois états : *inexploré*, *exploré* et *obstacle*, initialisée à *inexploré*. À chaque itération, chaque robot marque l'ensemble des cellules dans son rayon de perception comme *exploré* ou *obstacle*. Les frontières entre les zones explorées et inexplorées de la carte peuvent alors être calculées en détectant toutes les cellules contiguës qui ont à la fois des voisines immédiates explorées et inexplorées.

Une fois toutes les frontières de la carte détectées, l'algorithme d'assignation (*MinPos* [4]) est utilisé toutes les 4 itérations pour assigner une frontière à chaque robot, et ceux-ci peuvent alors se diriger vers la frontière qui leur a été assigné. La stratégie d'assignation *MinPos* est décrite en 4.1.

## 2.2 Algorithme BMILRV

L'algorithme *BMILRV*[1] utilise de nombreux marquages sur la grille d'occupation : un état parmi un ensemble d'états  $\{ \textit{inexploré}, \textit{exploré}, \textit{fermé}, \textit{obstacle}, \textit{rendez-vous} \}$ , une valeur indiquant si un robot contrôle une cellule donnée, la direction prise lors du dernier passage sur une cellule, et une variable booléenne indiquant si un robot donné est déjà passé sur une cellule. Les robots peuvent également être dans différents modes parmi  $\{ \textit{détection de boucle}, \textit{contrôle de boucle}, \textit{fermeture de boucle}, \textit{nettoyage}, \textit{pause}, \textit{arrêt} \}$ .

De manière informelle, l'algorithme *BMILRV* consiste à maintenir une liste de cellules ouvertes permettant d'accéder à toutes les zones inexplorées de la carte, tout en fermant progressivement toutes les cellules qui peuvent l'être (qui ne bloqueront pas le passage vers une zone inexplorée), évitant ainsi de repasser dans les zones déjà visitées. Lorsque des boucles sont présentes dans le chemin constitué par les cellules explorées, les robots peuvent les fermer en plusieurs étapes permettant d'éviter de bloquer un robot. Le robot détectant la boucle (entrant dans une cellule explorée sur laquelle il est déjà passé en venant de la même direction), va alors passer en mode « *contrôle de boucle* » et parcourir cette boucle pour la marquer comme contrôlée avec son identifiant, vérifiant ainsi qu'il est le seul à chercher à la fermer à ce moment. S'il parvient à marquer la totalité de la boucle, il effectue alors un second passage en mode « *fermeture* » ou il va fermer les cellules de la boucle jusqu'à la première intersection, et enfin parcourir la boucle une troisième fois en sens inverse en mode « *nettoyage* » pour enlever les traces indiquant qu'il contrôle la boucle. S'il ne parvient pas à contrôler la boucle, alors il passe directement en mode « *nettoyage* » si le robot qui la contrôle a une priorité supérieure, ou en mode « *pause* » si il est prioritaire, en attendant que l'autre robot ait nettoyé ses traces sur la boucle.

Les différentes étapes de l'algorithme *BMILRV* ne sont pas triviales ; elles nécessitent de nombreuses vérifications pour fermer correctement les cellules, et un décalage ou asynchronisme dans la procédure de marquage des cellules pourra bloquer les robots. Pour des raisons de place, nous n'entrons pas dans les détails du fonctionnement de l'algorithme ici.

## 3 Suivi de frontières locales

### 3.1 Exploration exhaustive

De façon informelle, l'algorithme local consiste à marquer tout chemin parcouru comme visité, à se déplacer le plus loin possible vers des zones inconnues, et revenir sur ses pas lorsqu'il n'y a plus de zone inconnue visible, jusqu'à ce que l'on en retrouve une.

Lorsque le robot est revenu à son point de départ sans avoir vu aucune zone inexplorée, toute la carte a été explorée.

Dans le cas de plusieurs robots, selon les conditions initiales et la topologie de la carte, un ou plusieurs robots

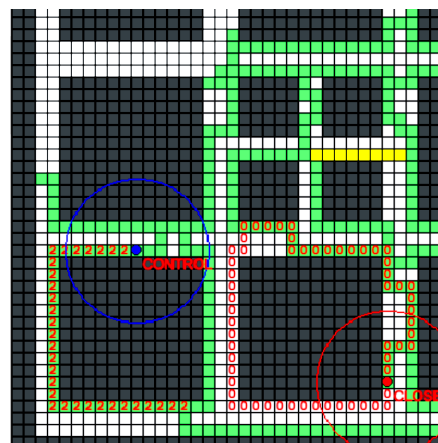


FIGURE 2 – Représentation de la grille lors d'une exploration par l'algorithme *BMILRV*. Les cellules fermées sont en gris, ouvertes en vert, point de rendez-vous en jaune et inexplorées en blanc. Les chiffres dans les cellules correspondent aux numéros des robots contrôlant ces cellules, l'état des robots est indiqué en rouge.

peuvent être revenus à leur point de départ avant la fin de l'exploration, ce qui entraîne une mauvaise distribution de l'exploration et réduit la performance de l'algorithme. Pour réduire cette perte de performance, on peut envisager d'utiliser dans ce cas les traces des autres robots pour reprendre l'exploration une fois qu'ils sont revenus au point de départ, si tous les autres robots ne se sont pas arrêtés. Cette piste d'amélioration est implémentée dans la partie 6.

### 3.2 Formalisation (frontières locales)

Nous considérons une flotte homogène de  $m$  robots mobiles  $R = \{r_1, \dots, r_m\}$  équipés de capteurs leur permettant de détecter des obstacles dans un rayon de perception  $\rho$ . Les robots utilisent une grille d'occupation  $\mathcal{O}$  représentant la carte, où chaque cellule peut être dans trois états : *inexploré*, *obstacle* ou *exploré*, ainsi qu'une grille  $\mathcal{P}$  pour assigner un booléen à chaque cellule, indiquant si une cellule est visitée ou non (c'est à dire si un robot a été sur cette cellule), initialisée à *faux*. Ces cartes sont échangées à intervalles réguliers entre les robots (voir section 5). Enfin, chaque robot peut définir une cellule objectif (représentant un point à atteindre sur la carte), et possède une trace  $\mathcal{G}_{r_i}$  correspondant à la distance maximale parcourue depuis le point de départ.

Pour chaque déplacement (d'une cellule), chaque robot connaît l'état des cellules dans son rayon de perception et sa position sur la grille  $(x, y)$ . On notera  $\mathcal{V}_{r_i}^t$  l'ensemble des cellules visibles perçues par le robot à un moment donné  $t$  qui ne sont pas des obstacles.

Les cartes locales ( $\mathcal{G}$  et  $\mathcal{P}$ ) sont partagées entre les robots toutes les  $d_{\text{sync}}$  itérations avec la procédure décrite dans la partie 5.

Le paramètre  $d_{\text{max}}$  est utilisé pour forcer les robots à réassi-



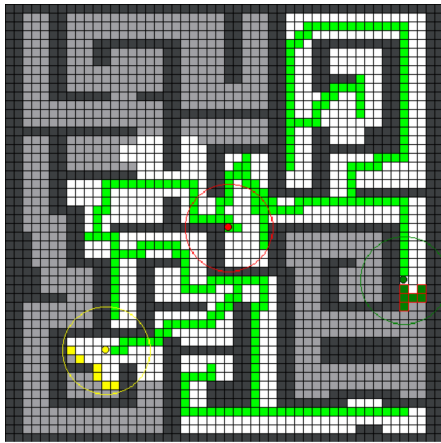


FIGURE 3 – Représentation de la grille lors d’une exploration par l’algorithme par frontières locales (cellules explorées en blanc, traces en vert; les cellules vert foncé et jaune correspondent aux frontières locales des robots vert foncé et jaune).

gner la cellule objectif à intervalles réguliers, évitant ainsi de choisir une cible maintenue trop longtemps si le rayon de perception est grand.

Chaque robot  $r_i$  applique l’algorithme suivant :

1. Marquer toutes les cellules visibles comme explorées :
 
$$\forall(x, y) \in \mathcal{V}_{r_i}^t, \mathcal{O}_{r_i}(x, y) = \text{exploré}$$
2. Si aucune cellule objectif n’est définie, déterminez les frontières locales  $\mathcal{F}$ , c’est-à-dire toutes les cellule de  $\mathcal{V}_{r_i}^t$  ayant des voisins inexplorés.
  - (a) Si  $|\mathcal{F}| > 0$ , choisir la cellule frontière la plus éloignée de toute cellule visitée (dans la grille  $\mathcal{P}_{r_i}$ ) à une distance maximale  $d_{max}$  de la cellule comme nouvel objectif.
  - (b) Si  $|\mathcal{F}| = 0$  et si on est de retour à la cellule de départ, s’arrêter.
  - (c) Sinon, retournez un pas en arrière sur la trace  $G_{r_i}$  : choisir la cellule, parmi les 4 cellules adjacentes, ayant la valeur  $\mathcal{G}_{r_i}(x, y)$  la plus basse.
3. Sinon, si une cellule objectif est définie
  - (a) Si  $\mathcal{G}_{r_i}(x, y) = 0$ , mettre à jour la trace :
 
$$\mathcal{G}_{r_i}(x, y) = 1 + \max(\mathcal{G}_{r_i}(x - 1, y), \mathcal{G}_{r_i}(x + 1, y), \mathcal{G}_{r_i}(x, y - 1), \mathcal{G}_{r_i}(x, y + 1)).$$
  - (b) Mettre à jour la grille :  $\mathcal{P}_{r_i}(x, y) = 1$
  - (c) Se rapprocher de l’objectif (à l’aide d’un algorithme de *pathfinding* parmi les cellules visibles), puis revenir à l’étape 1.
4. Si plus de  $d_{sync}$  itérations ont été effectuées depuis le dernier partage de carte, partagez la carte ( $\mathcal{O}_{r_i}, \mathcal{P}_{r_i}$ ) avec les autres robots (voir section 5).

Dans nos expériences, nous avons utilisé l’algorithme  $A^*$  pour rechercher le chemin optimal entre les robots et leurs cellules objectifs.

## 4 Comparaison des différentes approches

### 4.1 Algorithmes choisis

Puisque l’algorithme présenté ici est un compromis entre les approches *Brick & Mortar* et par *frontière*, nous avons choisi de le comparer à un algorithme par frontières globales et à un algorithme *Brick & Mortar*.

Les approches par frontières peuvent être mises en œuvre avec différentes stratégies d’affectation et différentes conditions de réaffectation. Nous avons choisi la stratégie d’affectation *MinPos* présentée dans [4], et réassignons les frontières toutes les  $d_{max}$  itérations ( $d_{max}$  est le paramètre utilisé dans l’algorithme local, et correspond au nombre maximum de déplacement entre chaque changement de cellule objectif, donnant ainsi une condition de réaffectation proche). La stratégie d’affectation *MinPos* consiste à calculer, pour chaque robot  $i$  et chaque frontière  $j$ , un rang  $r_{i,j}$ , correspondant au nombre de robots plus proches de la frontière  $j$  que le robot  $i$ , puis à attribuer à chaque robot la frontière la plus proche pour laquelle son rang est minimal. Il existe également différentes approches *Brick & Mortar*. Celle que nous avons implémenté ici, *BMILRV* (*Brick & Mortar Improved Long Range Vision*) [1], permet de considérer un ensemble de cellules dans le rayon de perception du robot, et pas seulement la cellule sur laquelle le robot se trouve, ce qui permet une comparaison plus directe avec notre approche locale, considérant à un instant donné pour un robot le même nombre de cellules voisines.

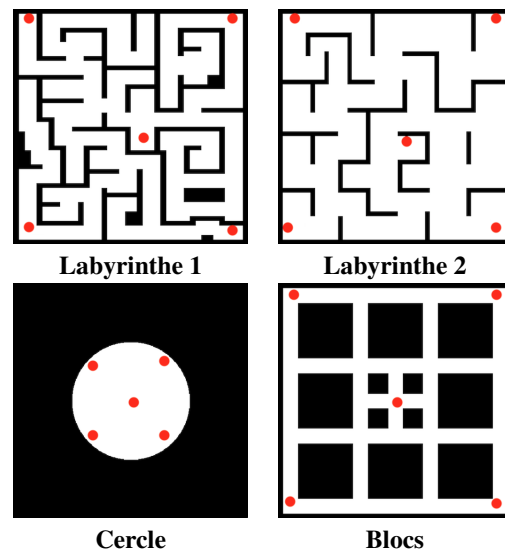


FIGURE 4 – Cartes testées (les points rouges correspondent aux points de départ des robots).

## 4.2 Conditions expérimentales

Nous allons comparer les trois approches afin d'explorer une carte inconnue. L'approche par frontières globales reproduit l'expérience décrite dans [7] avec réaffectation de l'objectif tous les  $d_{\max}$  changements de cellule, et la stratégie d'assignation *MinPos*. Comme cette approche ne fournit pas de point de rendez-vous, nous considérerons le nombre total d'itérations jusqu'à ce que la carte soit entièrement visitée. Nous observerons ainsi le pourcentage de cellules explorées par rapport au nombre d'itérations passées. L'approche *BMILRV* est une reproduction de l'algorithme décrit dans [1]. Enfin, l'approche par frontières locales correspond à l'algorithme décrit dans la partie 3.2.

Cette partie présente les résultats expérimentaux réalisés sur simulateur pour 4 cartes différentes présentées figure 4. Les points de départ des robots sont représentés par les points rouges. Dans ces expériences, la carte est partagée à chaque itération ( $d_{\text{sync}} = 1$ , voir section 5 pour la version distribuée), le rayon de perception  $\rho$  est de 5 cellules, et  $d_{\max} = 2$ .

Le *labyrinthe 2* est celui utilisé dans [1], le *labyrinthe 1* est une variante avec des couloirs plus étroits. La carte *Cercle* permet de tester un environnement avec peu d'obstacles (par rapport au rayon de perception des robots), et la carte *Blocs* correspond à un environnement avec de nombreux obstacles isolés, formant ainsi des boucles pour les robots.

## 4.3 Simulateur utilisé

Pour obtenir une première comparaison expérimentale des performances des différents algorithmes, nous avons implémenté un simulateur en Python permettant d'exécuter les algorithmes dans des conditions identiques sur les 4 cartes présentées figure 4.

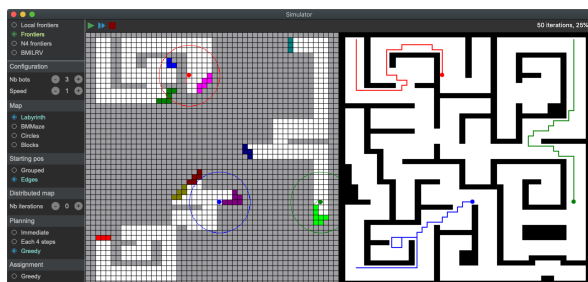


FIGURE 5 – Simulateur utilisé pour les expériences, ici avec l'algorithme par frontières globales. La partie gauche montre la carte des robots (cellules explorées en blanc, inexplorées en gris et frontières en couleurs), et la partie droite le monde « réel » avec les trajectoires des robots. Le menu de gauche permet de sélectionner les différents algorithmes (frontières globales, frontières locales, *BMILRV*) et les conditions de l'expérience (points de départ robot, cartes, synchronisation).

Notre simulateur (capture d'écran 5) permet une exécution décentralisée des algorithmes des robots. Dans nos expé-

riences, nous supposons que la localisation est parfaite, que les robots ont une taille plus petite que la taille d'une cellule dans la grille d'occupation et que la détection des obstacles est toujours correcte.

## 4.4 Durée de l'exploration

La figure 6 représente le nombre d'itérations nécessaires pour explorer la carte entière avec  $m = 3$  robots pour les différentes approches.

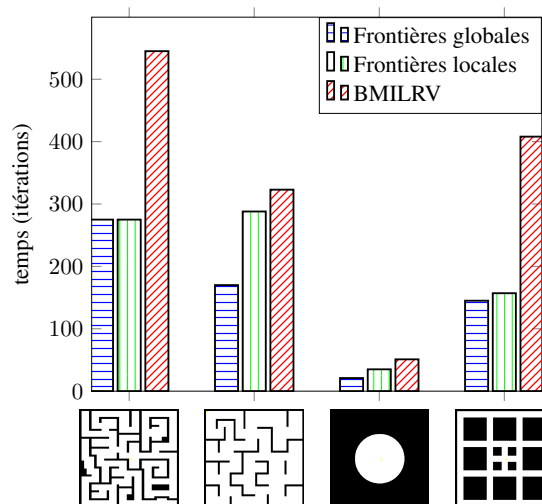


FIGURE 6 – Durée totale de l'exploration pour chaque approche sur les cartes testées avec  $m = 3$  robots.

L'approche *BMILRV* est beaucoup plus lente que les approches par frontières locales et globales, surtout lorsqu'il y a des obstacles isolés (forçant l'algorithme *BMILRV* à contrôler les cycles), ce qui est le cas des cartes *Labyrinthe 1* et *Blocs*. Les deux approches frontières donnent des résultats comparables, l'approche globale restant légèrement meilleure.

L'approche par frontières locales donne parfois une moins bonne dispersion des robots, et donc de moins bonnes performances, en particulier sur la carte *Labyrinthe 2*. Plusieurs améliorations peuvent être envisagées pour se rapprocher des performances de l'approche globale (voir 7.2).

## 4.5 Progression de l'exploration

La figure 7 présente l'évolution du pourcentage de cellules vues au cours du temps (en nombre d'itérations) pour les approches par frontières locales et globales, sur les deux premières cartes.

L'algorithme local perd facilement du temps lorsque les robots reviennent en arrière parce qu'ils prennent le même chemin que lors de leur premier passage, contrairement à l'algorithme global qui calcule le meilleur chemin pour atteindre une frontière. Cependant, les deux algorithmes montrent des résultats similaires, les deux courbes se croisent plusieurs fois pour la carte *Labyrinthe 1*.

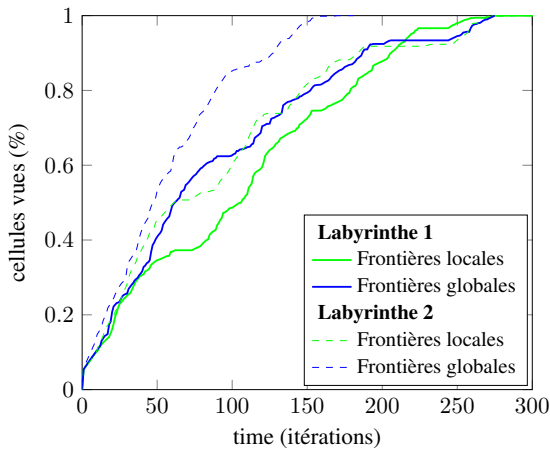


FIGURE 7 – Évolution du pourcentage de cellules vues au cours du temps ( $m = 3$  robots) sur les deux premières cartes.

#### 4.6 Nombre de robots

La figure 8 montre le nombre d’itérations nécessaires pour explorer les 4 cartes pour différents nombres de robots. Sur ces 4 cartes, les approches par frontières locales et globales donnent des résultats comparables.

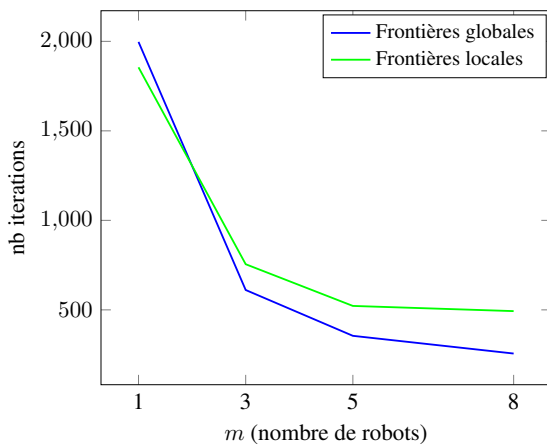


FIGURE 8 – Nombre d’itérations nécessaires pour explorer toutes les cartes pour différents nombres de robots.

Lorsque le nombre de robots augmente, l’algorithme par frontières globales prend l’avantage en offrant une meilleure répartition des robots sur la carte.

### 5 Version distribuée

Cette partie se concentre sur la version décentralisée de l’algorithme par frontières locales et tente d’évaluer l’impact de la décentralisation sur les performances.

#### 5.1 Procédure de partage de carte

Le partage de la carte est nécessaire pour deux raisons : la grille d’occupation permet aux robots de définir les frontières locales ; si elle n’est pas partagée, les robots ne peuvent pas savoir si autre un robot a déjà exploré une partie donnée de la carte. De plus, la trace est également partagée pour permettre une meilleure répartition des robots, car elle est utilisée dans le choix de la frontière à définir comme objectif.

Toutes les  $d_{sync}$  itérations, chaque robot peut envoyer sa grille d’occupation  $\mathcal{O}$  et ses cellules visitées  $\mathcal{P}$  à tous les autres robots. Lorsqu’un robot  $r_i$  reçoit des données d’un autre robot  $r_j$  (grille d’occupation  $\mathcal{O}_{r_j}$  et cellules visitées  $\mathcal{P}_{r_j}$ ), il met à jour sa grille d’occupation de la façon suivante :

$$\mathcal{P}_{r_i}(x, y) = \mathcal{P}_{r_i}(x, y) \vee \mathcal{P}_{r_j}(x, y)$$

Un *ou* logique est appliqué entre la valeur du robot et la valeur reçue : une cellule est visitée si le robot ou tout autre robot l’a déjà marquée comme visitée.

Il n’est pas nécessaire que les robots soient synchronisés pour partager leurs cartes en même temps, ils peuvent les envoyer ou les recevoir à tout moment.

#### 5.2 Résultats expérimentaux sur l’effet de la décentralisation sur la durée de l’exploration

La figure 9 montre le nombre d’itérations nécessaires pour explorer les 4 cartes avec 3 robots pour différentes valeurs  $d_{sync}$ .

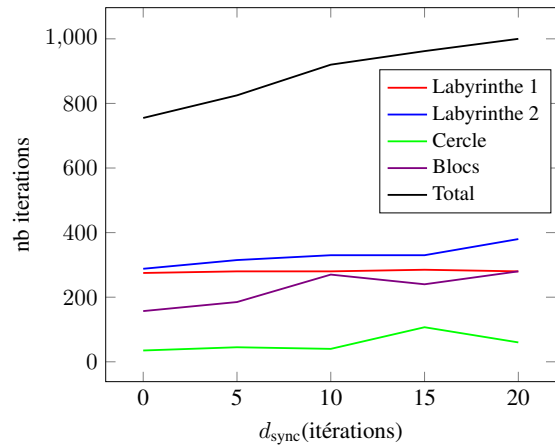


FIGURE 9 – Évolution de la durée de l’exploration des 4 cartes pour différentes valeurs de  $d_{sync}$

L’augmentation de la durée de l’exploration lorsque  $d_{sync}$  augmente est due à une moins bonne dispersion, notamment lorsque plusieurs robots sont proches les uns des autres : ils explorent alors la même zone simultanément. Concrètement, si l’on suppose que deux robots donnés partagent leur carte d’autant plus souvent qu’ils sont proches,



cette diminution sera alors limitée la plupart du temps (la présence de boucles peut cependant faire explorer une même zone à deux robots éloignés).

### 5.3 Perte de communication

Bien qu'augmenter  $d_{sync}$  diminue les performances de l'exploration, une perte de communication, ou une communication éparsée ne causera pas de problème dans l'exploration (« oubli » d'une zone, robot bloqué,...) et ne peut, dans le pire des cas, que prolonger la durée de l'exploration. Dans des applications concrètes, le partage de cartes entre robots peut être adapté en fonction de leur vitesse et de leur rayon de perception, éventuellement en fonction de la distance entre les robots.

## 6 Amélioration : « seconde chance »

Il peut arriver lors de l'exploration qu'un ou plusieurs robots retournent à leur point de départ et ne trouvent plus de frontières visibles avant que l'exploration soit terminée, laissant les robots encore actifs terminer seuls. Lorsque cela se produit, l'exploration reste complète, mais les robots revenus à leur point de départ ne participent plus, ce qui peut augmenter la durée totale de l'exploration.

Pour éviter cette perte de performances, nous pouvons envisager un système de « seconde chance », autorisant ces robots à repartir explorer les zones restantes en suivant les traces des autres robots. Nous allons donc modifier l'algorithme décrit dans la partie 3.2.

### 6.1 Algorithme avec seconde chance

Pour permettre aux robots de repartir, nous allons ajouter une variable *mode*, assigné à « normal » par défaut, et qui vaudra « *secondeChance* » lorsqu'un robot repart en suivant les traces des autres robots.

Pour activer le mode seconde chance, nous modifions le point 2. (b) de l'algorithme :

2. (b) Si  $|\mathcal{F}| = 0$  et si on est de retour à la cellule de départ :
  - i. Si il existe au moins un autre robot qui ne s'est pas arrêté et qui n'est pas en mode seconde chance, passer en mode seconde chance :  $mode \leftarrow secondeChance$ . Se déplacer vers une cellule voisine déjà visitée par un autre robot, si possible parmi celles que l'on a pas soit-même visité. Si on voit à nouveau une frontière, repasser en mode normal.
  - ii. Sinon, s'arrêter.

### 6.2 Évolution de l'exploration

Nous allons maintenant comparer trois algorithmes : frontières globales (avec *MinPos*), frontières locales sans seconde chance, et frontières locales avec seconde chance. Cette fois, nous choisissons  $m = 10$  robots qui partiront tous du centre du labyrinthe.

La figure 10 présente l'évolution de l'exploration pour les trois algorithmes sur le premier labyrinthe. Sur cette carte,

où les approches globales et locales donnaient des résultats très proches pour  $m = 3$  robots, l'algorithme local prend l'avantage avec la seconde chance, finissant l'exploration avant l'algorithme global.

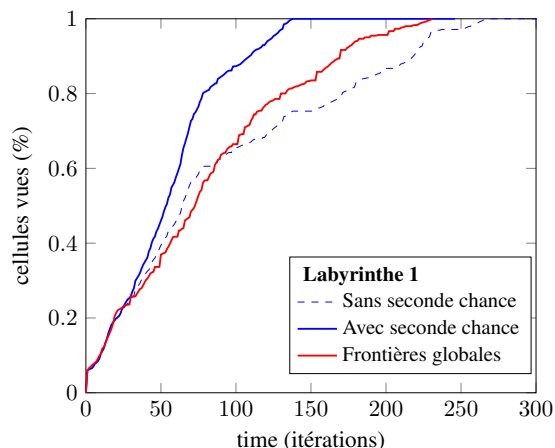


FIGURE 10 – Évolution du pourcentage de cellules vues au cours du temps ( $m = 10$  robots) avec et sans seconde chance.

### 6.3 Durée de l'exploration

Nous allons maintenant nous intéresser à la durée totale de l'exploration des quatre cartes par les trois algorithmes, pour  $m \in \{3, 5, 8\}$ .

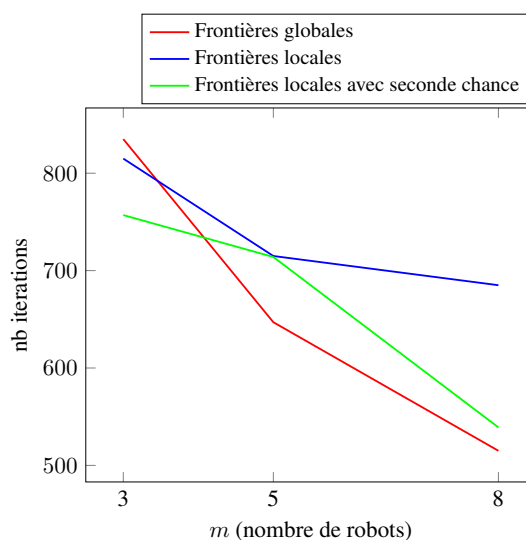


FIGURE 11 – Nombre d'itérations nécessaires pour explorer toutes les cartes pour différents nombres de robots. Les robots partent ici groupés au centre des cartes à explorer.

La figure 11 présente le nombre d'itérations nécessaires pour explorer les quatre cartes avec 3, 5 ou 8 robots. L'approche locale s'avère nettement meilleure avec la seconde

chance pour 3 ou 8 robots et s'approche des performances de l'approche globale.

Notons que nous mesurons ici la durée nécessaire pour parcourir tout le labyrinthe. Si nous prenons en compte le temps de retour au point de départ des robots (leur permettant notamment de savoir que l'exploration est terminée), la seconde chance peut dans certains cas allonger l'exploration.

## 7 Discussion

### 7.1 Forces et faiblesses des différents algorithmes

L'approche *BMILRV* est locale, mais a un fonctionnement complexe (de nombreux états pour les robots, la fermeture des cellules nécessite de nombreux calculs [1]) et donne de mauvais résultats lorsque des obstacles sont isolés ou qu'il y a des boucles sur la carte, forçant les robots à faire plusieurs passages pour marquer les cellules successivement comme contrôlées, fermées et finalement nettoyées (comme le confirment les expériences sur la figure 6). A notre connaissance, cette approche n'a pas été adaptée pour une mise en œuvre décentralisée.

L'approche par frontières (globales) est à la fois simple et efficace. Elle peut être distribuée facilement, mais nécessite une synchronisation fréquente de la carte et des positions du robot. Avec la stratégie d'affectation *MinPos*[4], chaque robot prend en compte la position de tous les autres robots pendant chaque affectation, ce qui implique que la carte des frontières globales et la position du robot sont fréquemment partagés. De plus, le temps de calcul de cette stratégie d'affectation dépend du nombre de robots et du nombre de frontières (les distances entre toutes les paires frontière/robot sont calculées à chaque itération), ce qui peut rendre cette méthode inutilisable si le nombre de robots est trop grand, ou si la carte est trop grande.

Enfin, l'approche par frontières locales présente des résultats proches de l'approche globale dans la plupart des cas, tout en ayant une complexité constante, ne dépendant ni du nombre total de robots ni de la taille de la carte. Il est également facilement distribuable sans contrainte sur la fréquence de partage des cartes, et les robots n'ont pas besoin de connaître le nombre ou la position des autres robots. Sans l'ajout de la « seconde chance », cependant, certains robots peuvent s'arrêter avant la fin de l'exploration, réduisant ainsi la performance de l'exploration.

### 7.2 Possibilités d'amélioration

La version de l'algorithme décrite ici ramène au point de départ des robots qui ne trouvent plus de frontières à explorer. Selon la carte et les conditions initiales, un robot peut revenir à son point de départ et s'arrêter avant que la carte entière ne soit explorée (laissant les autres robots terminer l'exploration). La distribution entre les robots est alors moins équilibrée et les performances de l'algorithme sont réduites. Pour éviter cela, nous pouvons envisager un système de « seconde chance » (voir partie 6), permettant

aux robots qui sont revenus à leur point de départ avant la fin de l'exploration de recommencer en suivant la trace de l'un des autres robots jusqu'à ce qu'il trouve une frontière. La dispersion des robots est également rudimentaire et ne repose que sur le choix d'une cible aussi loin que possible de toute trace laissée. Lorsque plusieurs robots sont proches (dans leurs rayons de perception respectifs), et en particulier si  $d_{sync}$  est élevé, l'ajout d'un système de dispersion supplémentaire pourrait également augmenter de manière significative les performances de l'algorithme sans augmenter considérablement sa complexité.

### 7.3 Conclusion

Dans la version proposée ici, l'algorithme donne de bonnes performances (nombre d'étapes pour compléter la couverture) par rapport aux deux autres approches, et au bénéfice d'une complexité constante (en temps et en mémoire) pour un robot donné, ce qui permet d'utiliser un très grand nombre de robots. En effet, dans le cas de l'approche par frontières globales, les stratégies les plus efficaces, telles que *MinPos*, nécessitent de calculer autant de distances que de paires robot/frontière, ce qui peut poser problème s'il y a un trop grand nombre de robots ou si la carte est trop grande. De plus, notre approche locale est plus facilement distribuable au prix d'une perte de performance possible lorsque les robots sont proches, et ne nécessite pas de synchronisation temporelle entre les robots.

Ces résultats montrent également qu'il existe des approches intermédiaires entre les algorithmes locaux et les algorithmes globaux, pouvant prendre en compte les cellules de la grille dans un rayon de perception déterminé et garantissant une exploration complète.

## Références

- [1] M. ANDRIES et F. CHARPILLET : Multi-robot taboo-list exploration of unknown structured environments. *In 2015 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 5195–5201, 09 2015.
- [2] M. ARAYA-LOPEZ, V. THOMAS, O. BUFFET et F. CHARPILLET : A closer look at momdps. *In 2010 22nd IEEE International Conference on Tools with Artificial Intelligence*, volume 2, pages 197–204, Oct 2010.
- [3] Antoine BAUTIN, Philippe LUCIDARME, Rémy GUYONNEAU, Olivier SIMONIN, Sébastien LAGRANGE, Nicolas DELANOUE et François CHARPILLET : Cart-O-matic project : autonomous and collaborative multi-robot localization, exploration and mapping. 5th Workshop on Planning, Perception and Navigation for Intelligent Vehicles. *In IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, page x, Tokyo, Japan, novembre 2013.
- [4] Antoine BAUTIN, Olivier SIMONIN et François CHARPILLET : Minpos : A novel frontier allocation algorithm for multi-robot exploration. *In International conference on intelligent robotics and applications*, pages 496–508. Springer, 2012.
- [5] W. BURGARD, M. MOORS, D. FOX, R. SIMMONS et S. THRUN : Collaborative multi-robot exploration. *In Pro-*

- ceedings 2000 ICRA. Millennium Conference. IEEE International Conference on Robotics and Automation. Symposia Proceedings (Cat. No.00CH37065)*, volume 1, pages 476–481 vol.1, 04 2000.
- [6] João Paulo Lima Silva de ALMEIDA, Renan Taizo NAKASHIMA, Flávio NEVES-JR et Lúcia Valéria Ramos de ARRUDA : Bio-inspired on-line path planner for cooperative exploration of unknown environment by a multi-robot system. *Robotics and Autonomous Systems*, 112:32 – 48, 2019.
- [7] Jan FAIGL, Olivier SIMONIN et Francois CHARPILLET : Comparison of task-allocation algorithms in frontier-based multi-robot exploration. In Nils BULLING, éditeur : *Multi-Agent Systems*, pages 101–110, Cham, 2015. Springer International Publishing.
- [8] Ettore FERRANTI, Niki TRIGONI et Mark LEVENE : Brick & mortar : an on-line multi-agent exploration algorithm. In *ICRA*, pages 761–767, 2007.
- [9] Ettore FERRANTI, Niki TRIGONI et Mark LEVENE : Rapid exploration of unknown areas through dynamic deployment of mobile and stationary sensor nodes. *Autonomous Agents and Multi-Agent Systems*, 19(2):210–243, 10 2009.
- [10] Enric GALCERAN et Marc CARRERAS : A survey on coverage path planning for robotics. *Robotics and Autonomous Systems*, 61(12):1258 – 1276, 2013.
- [11] S. GARNIER, F. TACHE, M. COMBE, A. GRIMAL et G. THERAULAZ : Alice in pheromone land : An experimental setup for the study of ant-like robots. In *2007 IEEE Swarm Intelligence Symposium*, pages 37–44, April 2007.
- [12] Elizabeth A. JENSEN et Maria GINI : Effects of communication restriction on online multi-robot exploration in bounded environments. In Nikolaus CORRELL, Mac SCHWAGER et Michael OTTE, éditeurs : *Distributed Autonomous Robotic Systems*, pages 469–483, Cham, 2019. Springer International Publishing.
- [13] N. MAHDoui, V. FRÉMONT et E. NATALIZIO : Cooperative frontier-based exploration strategy for multi-robot system. In *2018 13th Annual Conference on System of Systems Engineering (SoSE)*, pages 203–210, 06 2018.
- [14] Q. V. NGUYEN, F. COLAS, E. VINCENT et F. CHARPILLET : Localizing an intermittent and moving sound source using a mobile robot. In *2016 IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, pages 1986–1991, Oct 2016.
- [15] O. SIMONIN, T. HURAUx et F. CHARPILLET : Interactive surface for bio-inspired robotics, re-examining foraging models. In *2011 IEEE 23rd International Conference on Tools with Artificial Intelligence*, pages 361–368, Nov 2011.
- [16] Sebastian THRUN et Arno BÜCKEN : Integrating grid-based and topological maps for mobile robot navigation. In *Proceedings of the National Conference on Artificial Intelligence*, pages 944–951, 1996.
- [17] E. VINCENT, A. SINI et F. CHARPILLET : Audio source localization by optimal control of a mobile robot. In *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*, pages 5630–5634, April 2015.
- [18] B. YAMAUCHI : A frontier-based approach for autonomous exploration. In *Proceedings 1997 IEEE International Symposium on Computational Intelligence in Robotics and Automation CIRA'97. 'Towards New Computational Principles for Robotics and Automation'*, pages 146–151, 1997.



# Apprentissage automatique sur des données de type graphe utilisant le plongement de Poincaré et les algorithmes stochastiques riemanniens

Hatem Hajri      Hadi Zaatiti      Georges Hébrail      Patrice Aknin  
 Institut de recherche technologique SystemX, 8 Avenue de la Vauve, 91120 Palaiseau, France.  
 Prénom.Nom@irt-systemx.fr

## Résumé

*Afin de mieux capter la complexité des relations qui existent entre les nœuds d'un graphe binaire, des travaux originaux ont montré l'intérêt de représenter ces données dans des espaces hyperboliques. Ces travaux sont issus d'une première communauté qui s'intéresse aux données graphiques et à leurs représentations et visualisations. D'autre part, une seconde communauté qui adresse plutôt des applications en traitement des données Radar et en vision par ordinateur, a développé ces dernières années des outils d'apprentissage sur certains espaces hyperboliques en exploitant leurs intéressantes propriétés géométriques. Dans cet article, nous présentons nos travaux récents [13, 27] qui visent à rapprocher ces deux approches. Plus précisément, nous combinons les plongements des graphes et les méthodes récentes de partitionnement sur les espaces hyperboliques dans le but de réaliser une classification par apprentissage sur les données initiales du graphe. Nous illustrons cette proposition par des applications en montrant le gain obtenu vis-à-vis de l'état de l'art.*

## Mots Clef

Plongement de Poincaré, Barycentre Riemannien, Algorithme des  $K$ -moyennes Riemannien, Algorithme espérance-maximisation Riemannien.

## 1 Introduction

L'idée de plonger des données dans un nouvel espace pour en faciliter le traitement a démontré son efficacité dans de multiples applications. Les techniques comme Word2vec [15], Glove [18], Node2vec [12], Graph2vec [16] et DeepWalk [19] sont devenues célèbres grâce à leurs capacité à représenter les données tout en réduisant la complexité de l'espace initial pour les traitements ultérieurs. Des travaux récents comme [17, 10, 23, 24] montrent que les représentations des graphes dans des espaces hyperboliques préservent plus fidèlement la structure sous-jacente ainsi que les relations latentes entre les noeuds d'un graphe que leurs homologues euclidiens.

Dans le but de partitionner les noeuds d'un graphe binaire donné, plusieurs équipes proposent de plonger d'abord le graphe dans un espace euclidien avec les techniques citées (Node2vec, Graph2vec, DeepWalk) et ensuite d'appliquer des méthodes classiques de partitionnement comme les al-

gorithmes des  $K$ -moyennes et d'espérance-maximisation (EM) [28, 25, 26, 9, 28]. Nos articles récents [13, 27] proposent des versions hyperboliques de ces travaux, motivés en grande partie par les bons résultats obtenus en traitement de données Radar [3] et en vision par ordinateur [7, 21, 22]. En particulier [7] utilise le barycentre riemannien sur l'espace des matrices de covariances pour classifier des signaux EEG. [3] utilise la médiane riemannienne sur le disque de Poincaré pour détecter les données aberrantes parmi des signaux Radar. [21] introduit l'algorithme EM sur l'espace des matrices de covariance et l'applique au problème de classification des images. Ces applications se basent sur les propriétés intéressantes des espaces hyperboliques qui seront rappelées dans ce papier.

L'article est organisé comme suit. La section 2 présente un rappel sur la géométrie riemannienne du disque de Poincaré comme un modèle typique de géométrie hyperbolique et introduit la version riemannienne des algorithmes des  $K$ -moyennes et d'espérance-maximisation. La section 3 rappelle la méthode récente du plongement des données de type graphe dans le disque de Poincaré [17]. Nous complétons ensuite cette section par deux propositions pour l'apprentissage supervisé et non supervisé sur les graphes. La section 4 présente des expérimentations démontrant l'intérêt des approches proposées tout en les comparant avec l'état de l'art récent. Enfin la section 5 conclut l'article en esquissant les perspectives de ce travail.

## 2 Apprentissage statistique riemannien sur le disque de Poincaré

Dans cette section, nous introduisons les outils mathématiques nécessaires qui seront utilisés pour la partie apprentissage automatique. Nous commençons par rappeler la géométrie du disque de Poincaré. Ensuite, nous présentons deux extensions des algorithmes  $K$ -moyennes et espérance maximisation sur cet espace.

Le disque de Poincaré est l'espace noté  $\mathbb{D} = \{z \in \mathbb{C} : |z| < 1\}$ , muni de la distance dite de Poincaré  $d(z_0, z_1)$  définie par :

$$\frac{1}{2} \log \left( \left( 1 + \left| \frac{z_1 - z_0}{1 - \bar{z}_0 z_1} \right| \right) \left( 1 - \left| \frac{z_1 - z_0}{1 - \bar{z}_0 z_1} \right| \right)^{-1} \right)$$

Par ailleurs, rappelons qu'une transformation de Möbius

est une application de la forme  $f_{a,b} : z \in \mathbb{D} \mapsto \frac{az+b}{\bar{b}z+\bar{a}}$  où  $a, b \in \mathbb{C}$  et  $|a|^2 - |b|^2 = 1$ .  $\mathcal{M}$ , l'ensemble de toutes les transformations de Möbius est un groupe pour la composition qui agit sur  $\mathbb{D}$  transitivement (i) et isométriquement (ii) :

- (i) pour tous  $z_1, z_2 \in \mathbb{D}$ , il existe  $f_{a,b} \in \mathcal{M}$  t.q  $f_{a,b}(z_1) = z_2$
- (ii) pour tous  $z_1, z_2 \in \mathbb{D}$ ,  $d(f_{a,b}(z_1), f_{a,b}(z_2)) = d(z_1, z_2)$

Muni du groupe  $\mathcal{M}$ ,  $\mathbb{D}$  devient un espace riemannien symétrique homogène de courbure négative [14].

Une propriété importante satisfaite par ces espaces est l'existence et l'unicité du barycentre riemannien (aussi appelé moyenne riemannienne ou moyenne de Fréchet). Plus précisément, pour tout ensemble  $\{z_i, 1 \leq i \leq n\}$  inclus dans  $\mathbb{D}$ , la solution du problème d'optimisation suivant (barycentre riemannien)

$$\hat{z}_n = \operatorname{argmin}_{z \in \mathbb{D}} \left( \sum_{i=1}^n d^2(z, z_i) \right) \quad (1)$$

existe, est unique et appartient à  $\mathbb{D}$  [2].  $\hat{z}_n$  peut être approché numériquement par une extension au cadre riemannien de la méthode de descente de gradient classique. Ceci a fait l'objet de plusieurs travaux comme [8, 4, 6, 5, 3]. Dans la suite, nous allons utiliser l'algorithme (1) proposé dans [3]. Les notations  $\exp_z$  et  $\log_z$  désignent respectivement les fonctions exponentielle et logarithme riemanniennes.

---

**Algorithm 1** Calcul approché du barycentre riemannien sur  $\mathbb{D}$  par descente de gradient

---

**Entrée :**  $Z = \{z_i, 1 \leq i \leq n\}$  : sous ensemble de  $\mathbb{D}$ ,  $z_{init}$  : initialisation du barycentre,  $\tau$  : pas de la descente de gradient,  $\varepsilon$  : précision d'arrêt

**Sortie :**  $\hat{z}$  : approximation du barycentre riemannien

- 1:  $\hat{z} \leftarrow z_{init}$
- 2:  $d = 1e^6$  ▷ un nombre assez large
- 3: **while**  $d > \varepsilon$  **do**
- 4:

$$\mu \leftarrow \frac{2}{n} \sum_{i=1}^n \log_{\hat{z}}(z_i), \hat{z} \leftarrow \exp_{\hat{z}}(\tau \mu), d \leftarrow \sqrt{\frac{|\mu|^2}{(1 - |\hat{z}|^2)^2}}$$

- 5: **end while** **return**  $\hat{z}$
- 

Dans les deux paragraphes suivants, nous présentons les algorithmes des  $K$ -moyennes et espérance-maximisation riemanniens qui exploitent les propriétés intéressantes de l'espace de Poincaré.

## 2.1 Algorithme des $K$ -moyennes riemannien

Par analogie avec le cas euclidien, nous pouvons étendre l'algorithme des  $K$ -moyennes au cadre riemannien. Cette extension est basée sur l'observation que pour un ensemble  $\{x_1, \dots, x_n\}$ , la moyenne  $\hat{x}_n = \frac{1}{n} \sum_{i=1}^n x_i$  correspond aussi au minimum global de la fonction  $\sum_{i=1}^n (x - x_i)^2$ . Dans la

version riemannienne, la distance euclidienne sera remplacée par son homologue riemannien.

---

**Algorithm 2**  $K$ -moyennes riemannien sur  $\mathbb{D}$

---

**Entrée :**  $K$  : nombre de classes,  $Z$  : un ensemble de  $n$  complexes appartenant à  $\mathbb{D}$

**Sortie :**  $N$  : ensemble de  $K$  barycentres,  $labels$  :  $n$  étiquettes des données d'entrée  $Z$

- 1: **Initialisons**  $K$  barycentres,  $N = \{\nu_1, \nu_2, \dots, \nu_K\}$  aléatoirement dans  $\mathbb{D}$
  - 2: **repeat**
  - 3: **for**  $z_i \in Z$  **do**  $c_i \leftarrow \operatorname{argmin}_{j \in \{1, \dots, K\}} d(z_i, \nu_j)$
  - 4: **for**  $j \in \{1, \dots, K\}$  **do**
  - 5:  $\nu_j \leftarrow \operatorname{Barycentre\_Riemannien}(\{z_i | c_i = j\})$
  - 6: **end for**
  - 7: **until** les barycentres deviennent stables
  - 8: **return**  $N, labels = \{c_i\}$
- 

## 2.2 Algorithme $EM$ riemannien

En exploitant les propriétés de la distance de Poincaré, il est possible d'étendre l'algorithme  $EM$  classique à cet espace [22]. Cette extension s'appuie sur une version riemannienne de la loi gaussienne.

Étant donné deux paramètres  $\bar{z} \in \mathbb{D}$  (une position) et  $\sigma > 0$  (un écart type), la loi gaussienne  $G(\bar{z}, \sigma)$  dépendant de  $(\bar{z}, \sigma)$  est définie dans [22] par sa densité

$$p(z | \bar{z}, \sigma) = \frac{1}{\zeta(\sigma)} \exp \left[ -\frac{d^2(z, \bar{z})}{2\sigma^2} \right]$$

relativement au volume riemannien  $dv(z)$ . Une propriété clé de cette distribution, découlant de la symétrie de  $\mathbb{D}$ , est que  $\zeta(\sigma)$  ne dépend pas de  $\bar{z}$  (tout comme la loi gaussienne classique). Ceci permet d'en déduire les estimateurs du maximum de vraisemblance (EMV) de  $\bar{z}$  et  $\sigma$  étant donné un échantillon  $z_1, \dots, z_N$  de  $G(\bar{z}, \sigma)$  comme suit.

**EMV de  $\bar{z}$ .** L'EMV de  $\bar{z}$  noté  $\hat{z}_N$  est le barycentre riemannien empirique de  $z_1, \dots, z_N$ .

**EMV de  $\sigma$ .** L'EMV de  $\sigma$  est  $\hat{\sigma}_N = \Phi \left( \frac{1}{N} \sum_{n=1}^N d^2(\hat{z}_N, z_n) \right)$  où  $\Phi$  est l'inverse de la fonction  $\sigma \mapsto \sigma^3 \times \frac{d}{d\sigma} \log \zeta(\sigma)$

En pratique, comme  $\hat{z}_N, \hat{\sigma}_N$  n'a pas d'expression explicite mais peut être approché numériquement par des méthodes de Newton. Considérons maintenant un modèle de mélange gaussien

$$p(z | (\varpi_\mu, \bar{z}_\mu, \sigma_\mu)_{1 \leq \mu \leq M}) = \sum_{\mu=1}^M \varpi_\mu \times p(z | \bar{z}_\mu, \sigma_\mu)$$

où les  $\varpi_\mu$  sont positifs de somme 1. Les paramètres du modèle  $\varpi_\mu, \bar{z}_\mu, \sigma_\mu$  peuvent être estimés par une extension de l'algorithme EM classique [22]. Pour ceci, introduisons pour tout  $\vartheta = \{(\varpi_\mu, \bar{z}_\mu, \sigma_\mu)\}$ ,

$$\omega_\mu(z_n, \vartheta) = \frac{\varpi_\mu \times p(z_n | \bar{z}_\mu, \sigma_\mu)}{\sum_{s=1}^M \varpi_s \times p(z_n | \bar{z}_s, \sigma_s)}$$

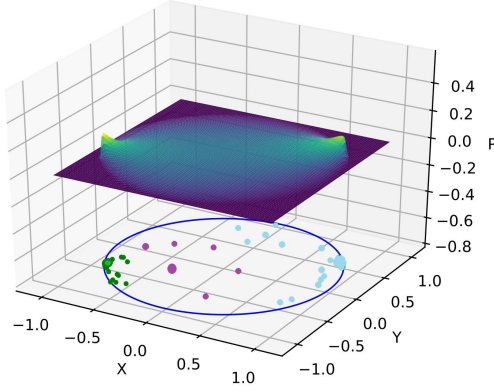


FIGURE 1 – Application de l’algorithme EM riemannien au dataset Polbooks à trois classes (présenté dans la section 4) avec une représentation 3D des densités correspondantes

et

$$N_\mu(\vartheta) = \sum_{n=1}^N \omega_\mu(z_n)$$

L’algorithme EM riemannien met à jour  $\hat{\vartheta} = \{(\hat{\varpi}_\mu, \hat{Y}_\mu, \hat{\sigma}_\mu)\}$  de façon itérative pour générer l’EMV de  $\vartheta = (\varpi_\mu, Y_\mu, \sigma_\mu)$  de la manière suivante.

► Mise à jour de  $\hat{\varpi}_\mu$  :  $\hat{\varpi}_\mu = N_\mu(\hat{\vartheta})/N$ .

► Mise à jour de  $\hat{z}_\mu$  (en utilisant une descente de gradient riemannienne) :

$$\hat{z}_\mu = \operatorname{argmin}_z \sum_{n=1}^N \omega_\mu(z_n, \hat{\vartheta}) d^2(z, z_n)$$

► Mise à jour de  $\hat{\sigma}_\mu$  (en utilisant une méthode de Newton) :

$$\hat{\sigma}_\mu = \Phi(N_\mu^{-1}(\hat{\vartheta}) \times \sum_{n=1}^N \omega_\mu(z_n, \hat{\vartheta}) d^2(\hat{z}_\mu, z_n))$$

La figure 1 présente une visualisation d’un partitionnement par l’algorithme EM riemannien du dataset Polbooks à trois classes (présenté dans la section 4).

### 3 Apprentissage sur les graphes en utilisant le plongement de Poincaré

Dans cette section, nous présentons une approche récente [17] qui propose le plongement d’un graphe binaire dans un disque de Poincaré. Ensuite, nous proposons deux algorithmes de partitionnement des noeuds d’un graphe dans les cas supervisé et non supervisé.

Considérons un graphe  $(\mathcal{V}, \mathcal{E})$  où  $\mathcal{V}$  est l’ensemble des noeuds et  $\mathcal{E}$  est l’ensemble des arêtes. Pour plonger  $\mathcal{V}$  dans  $\mathbb{D}$ , [17] propose d’apprendre une application  $u \in \mathcal{V} \mapsto \theta_u \in \mathbb{D}$  minimisant la fonction

$$\mathcal{L}(\Theta) = \log(\sigma(-d(\theta_u, \theta_v))) + \sum_{i=1}^M \mathbb{E}_{P_N} [\log(\sigma(d(\theta_{u_i}, \theta_v)))]$$

où  $u_1, \dots, u_M$  est un ensemble d’échantillons négatifs tirés suivant une distribution  $P_N$  et  $\sigma(x) = (1 + e^{-x})^{-1}$  est la fonction softmax [15]. En minimisant cette fonction,  $\theta_u$  et  $\theta_v$  se rapprochent l’un de l’autre et les noeuds des échantillons négatifs s’éloignent de  $\theta_v$ . En pratique  $\mathcal{L}(\Theta)$  est optimisé en générant des noeuds sur le graphe à l’aide de l’algorithme DeepWalk [19] puis en tirant des échantillons négatifs suivant la distribution unigram à la puissance 3/4 [15].

Étant donné un cluster  $\mathcal{C} = \{z_i, 1 \leq i \leq n\}$  dans  $\mathbb{D}$ , ayant comme barycentre  $\hat{z}$ , on définit sa variance empirique par

$$\operatorname{Var}(\mathcal{C}) = \frac{1}{n} \sum_{i=1}^n d^2(\hat{z}, z_i)$$

L’algorithme 3 ci-dessous présente notre schéma pour effectuer un partitionnement non supervisé sur un graphe. L’idée est de générer plusieurs plongements de Poincaré et de garder le plongement le plus concentré après déroulement d’un algorithme  $K$ -moyennes. Pour notre travail, nous allons considérer comme meilleur, le plongement ayant la variance minimale (d’autres métriques sont possibles). Dans ce qui suit, la fonction `Poincare_Plongement` plonge un graphe sur le disque de Poincaré en minimisant  $\mathcal{L}(\Theta)$ .

Cet algorithme peut être suivi aussi par un algorithme EM riemannien pour obtenir une classification plus robuste tenant compte de la dispersion des points au sein des clusters. L’algorithme 4 ci-dessous présente le schéma proposé pour l’apprentissage supervisé. Bien que cet algorithme puisse être alimenté par l’algorithme 3, une version simple et indépendante de celui-ci est présentée ici.

La seconde alternative proposée réalise le plongement de la vérité de terrain selon un mélange gaussien en utilisant l’algorithme EM dans un premier temps et associant par la suite une donnée de l’ensemble de test à une classe en utilisant la règle de Bayes (voir règle (69) dans [21]).

## 4 Expérimentations

Dans cette section, nous illustrons l’intérêt des approches proposées précédemment par quelques applications en apprentissage supervisé et non supervisé sur des graphes de référence. Nous montrons le gain apporté vis-à-vis de l’état de l’art en nous concentrant sur l’algorithme des  $K$ -moyennes. Les datasets utilisés sont publics [20] et leurs caractéristiques sont données dans le tableau suivant (le nombre de noeuds, arêtes et classes est désigné respectivement par Nœuds, Liens et  $K$ ).

### 4.1 Apprentissage non supervisé

Dans une première série d’expériences, nous avons utilisé les 5 premiers datasets donnés dans le tableau précédent. Pour chaque dataset, nous avons généré 10 marches aléatoires sur le graphe en suivant DeepWalk [19]. Sur chaque

**Algorithm 3** Apprentissage non supervisé

**Entrée :**  $\mathcal{G}$  : graphe binaire,  $K$  : nombre de clusters (entier),  $NE$  : nombre d'expérimentations

**Sortie :**  $Meilleur\_Plongement$  : Plongement des noeuds du graphe d'entrée ayant la variance totale minimale,  $N$  : barycentres des  $K$  clusters,  $labels$  : étiquette de chaque noeud

```

1: repeat
2:
    $Plongement_i \leftarrow Poincare\_Plongement(\mathcal{G})$ 
3:
    $N_i, labels_i \leftarrow K\_moyennes(K, Plongement_i)$ 
4:
    $Clusters \leftarrow \{c_1, \dots, c_K\}$ 
5: tel que  $c_j = \{z_q \in Plongement_i | labels_i[q] = j\}$ 
6:
    $var_i = max_{j \in \{1, \dots, K\}} (Var(c_j))$ 
7: until  $NE$  plongements sont effectuées
8:
    $Meilleur\_Plongement \leftarrow Plongement_{meilleur}$ 
9: tel que  $meilleur = argmin_{i \in \{1, \dots, NE\}} (var_i)$ 
10: return
11:  $Meilleur\_Plongement, N_{Meilleur}, labels_{Meilleur}$ 

```

marche aléatoire, nous générons un plongement de Poincaré et un plongement euclidien dans  $\mathbb{R}^{10}$  (en utilisant la version euclidienne de l'algorithme (3)). Ensuite, nous appliquons à chaque plongement hyperbolique (resp. euclidien), un algorithme des  $K$ -moyennes hyperbolique (resp. euclidien). Dans les deux cas, nous sélectionnons comme meilleur plongement celui ayant une variance totale minimale au sens de l'algorithme (3). Enfin, nous évaluons la performance des meilleurs plongements dans les deux cas ainsi que la performance moyenne sur les 10 tests effectués. Les résultats sont donnés dans le tableau 2 en utilisant les abréviations suivantes :

- MPP : meilleur plongement de Poincaré
- MPE : meilleur plongement euclidien
- 10MP : moyenne sur les 10 plongements de Poincaré
- 10ME : moyenne sur les 10 plongements euclidiens

Le tableau 2 montre que suivre un plongement par un algorithme des  $K$ -moyennes est plus avantageux dans le cas hyperbolique que dans le cas euclidien. La figure 2 montre l'évolution croissante de la performance MPP avec le nombre d'expérimentation effectués  $NE$  pour le dataset Football et illustre ainsi l'intérêt de générer plusieurs plongements.

En ce qui concerne l'exploitation des plongements de Poin-

**Algorithm 4** Apprentissage supervisé

**Entrée :**  $\mathcal{G}$  : graphe binaire,  $VT$  : une vérité terrain pour chaque noeud de  $\mathcal{G}$

**Sortie :**  $Cluster$  : cluster d'appartenance de chaque noeud des données d'apprentissage

```

1:  $Plongement \leftarrow Poincare\_Plongement(\mathcal{G})$ 
2:
    $DE, DT \leftarrow Diviser(Plongement)$ 
    $\triangleright DE$  sont les données d'entraînements
    $\triangleright DT$  sont les données de tests
3:
    $\{c_1, \dots, c_K\} \leftarrow Calculer\_Clusters(DE, VT)$ 
4: for  $q \in \{1, \dots, K\}$  do
5:    $\nu_q \leftarrow Barycentre\_Riemannien(c_q)$ 
6: end for
7: for  $u \in DT$  do
8:
    $Cluster(u) \leftarrow argmin_{p \in \{1, \dots, K\}} (d(\nu_p, u))$ 
    $\triangleright d$  est la distance riemannienne
9: end for
10: return  $Cluster$ 

```

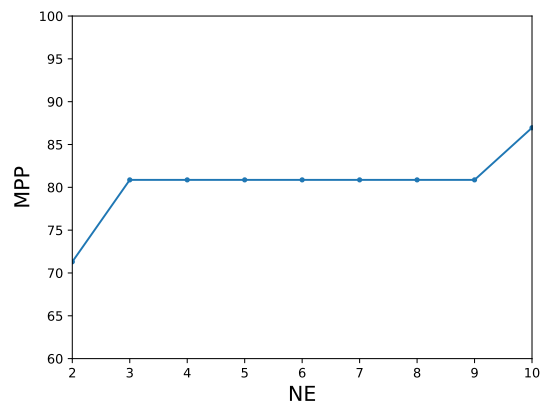


FIGURE 2 – Courbe représentant la variation du MPP en fonction de  $NE$  (nombre d'expérimentations effectuées)



Datasets	Noeuds	Liens	$K$
Karate	34	77	2
Polblogs	1224	16781	2
Polbooks	105	441	3
Football	115	613	12
Adjnoun	112	425	2
Mammifères	1179	6541	NA

TABLE 1 – Caractéristiques des graphes utilisés.

caré pour le partitionnement des noeuds, remarquons que le meilleur plongement de Poincaré (MPP) tel que défini précédemment, n’obtient pas nécessairement la meilleure performance. Pour un seul dataset, la performance moyenne (10MP) obtenue est légèrement supérieure à MPP. Ceci peut être expliqué par l’existence de données aberrantes (des personnes pour lesquelles il est difficile de prédire une classe). L’avantage du meilleur plongement hyperbolique est remarquable pour l’exemple Football (+17%) et pour l’exemple Polbooks (+4.8%). En comparaison avec l’état de l’art récent, nos résultats dépassent ceux obtenus dans [1] par un plongement sur une surface généralisée. Les datasets considérés dans [1] sont **Polbooks**, **Football**, **Adjnoun** avec des taux de réussite respectifs de 75%, 77% et 51% respectivement. Les améliorations pour ces exemples en utilisant notre méthode sont significatives : +9.8%, +10% et +0.8%.

La figure 3 montre des visualisations des clusters calculés en utilisant le meilleur plongement de Poincaré (MPP) pour chaque dataset. Chaque cluster est représenté par une couleur et le barycentre est symbolisé par un carré.

Comme deuxième application, nous avons considéré un exemple de données hiérarchiques (sous forme d’arbre) qui est le sous-arbre des mammifères extrait de Wordnet traité comme un graphe. La figure 4 montre les clusters obtenus quand  $K = 6$ . Les barycentres sont représentés par des carrés. Enfin, la figure 5 montre explicitement certaines étiquettes des noeuds choisies aléatoirement. Un focus est mis sur les noeuds proches des barycentres d’une part et sur les frontières entre les différents clusters d’autre part.

Remarquons que les clusters calculés permettent de discerner entre différents types de mammifères. Par exemple le cluster bleu contient en majorité la famille des canidés alors que le cluster orange contient des mammifères de taille plus importante (lion, tigre, éléphant,...).

## 4.2 Apprentissage supervisé

Dans cette partie, nous exploitons la notion du barycentre riemannien pour effectuer de l’apprentissage supervisé sur les graphes en suivant l’algorithme 4. Pour évaluer la méthode, nous avons subdivisé chaque dataset de la liste précédente (wordnet exclu) en 5 parties : 4/5 pour l’entraînement et 1/5 pour le test. Après plongement de tout le graphe, chaque élément de l’ensemble test est associé à un cluster suivant la règle du plus proche barycentre. Nous avons répété cette expérience 5 fois en

permutant les ensembles tests par validation croisée. La même expérience (5 permutations de l’ensemble test) a été refaite 10 fois. Finalement, nous avons évalué la performance moyenne en utilisant la vérité terrain des données tests. Le tableau 2 présente les résultats obtenus. On utilisera l’abréviation suivante :

10VC : Performance moyenne par validation croisée sur les 10 expériences effectuées.

Nous allons maintenant comparer nos résultats avec ceux de [11] qui utilise une généralisation de la méthode SVM au disque de Poincaré. [11] considère les datasets **Karate**, **Polbooks**, **Football**, **Polblogs** et obtient des taux de réussite respectifs de 86%, 73%, 24%, 93% sur 5 expérimentations par validation croisée sur 5 plongements différents selon [10]. Nous avons obtenu des améliorations significatives pour les datasets Karate, Polbooks et Football de +7.9%, +10.3%, de +53.9% avec une légère baisse de performance de -0.6% pour le dataset Polblogs.

## 5 Conclusion

Dans cet article, nous avons présenté des algorithmes d’apprentissage sur les graphes basés sur les plongements de Poincaré et des extensions au cadre riemannien des algorithmes des  $K$ -moyennes et d’espérance-maximisation. Nous avons illustré l’intérêt des algorithmes des  $K$ -moyennes par des expérimentations en montrant le gain apporté vis-à-vis de l’état de l’art.

D’autre part, nous avons présenté les algorithmes EM et esquissé quelques exemples de leurs applications en lien avec les plongements de Poincaré. Notons qu’avec la notion de variance totale minimale introduite dans ce papier, il est toujours possible d’augmenter le nombre de plongements (fixé à 10 dans l’article) et aussi la dimension de l’espace de Poincaré (fixé à 2 dans l’article). Ce raisonnement est possible grâce aux bonnes propriétés des variétés hyperboliques et est complètement non supervisé dans le sens où il ne nécessite aucune vérité terrain. Par conséquent, des améliorations des résultats présentés dans ce papier sont toujours possibles. Les résultats obtenus avec le plongement euclidien utilisant DeepWalk suggèrent que pour obtenir de bonnes performances avec cette approche (ou autres variétés comme Graph2vec [16], Node2vec [12]), il faudrait prendre des espaces euclidiens avec des dimensions très grandes.

En perspective, nous envisageons d’étudier la complexité des algorithmes riemanniens présentés dans cet article. Nous comptons aussi utiliser ces travaux pour étendre les résultats de [9, 28] qui propose d’apprendre les communautés sur les graphes en utilisant les lois de mélange gaussien (euclidien). Notre proposition sera d’étendre cela à une version géométrique. Plus précisément au lieu de générer plusieurs plongements et en choisir un à la fin, il sera possible d’optimiser le plongement et d’entraîner le partitionnement par le mélange gaussien simultanément. Enfin, nous souhaitons étendre les travaux à l’apprentissage

Datasets	Apprentissage non-supervisé					Apprentissage supervisé	
	MPP	MPE	10MP	10ME	[1]	10VC	[11]
Karate	<b>91.2%</b>	70.6%	<b>91.4%</b>	65.8%	—	<b>93.9%</b>	86%
Polblogs	<b>92.8%</b>	51.9%	<b>92.5%</b>	53.5%	—	<b>92.4</b>	<b>93%</b>
Polbooks	<b>84.8%</b>	77.1%	<b>80%</b>	62%	75%	<b>83.3%</b>	73%
Football	<b>87%</b>	67.8%	<b>69.4%</b>	56.8%	77%	<b>77.9 %</b>	24%
Adjnoun	51.8%	51.8 %	<b>52.5%</b>	51.6%	51%	<b>57.8 %</b>	-

TABLE 2 – Tableau comparatif des performances obtenues pour les algorithmes d’apprentissage supervisé et non-supervisé. Les meilleurs résultats sont en gras.

en ligne et adresser le problème des tests statistiques basés sur les lois gaussiennes et leurs applications sur les données graphiques et les données texte.

## Références

- [1] M. Aalto and N. Verma. Metric learning on manifolds. *CoRR*, <https://arxiv.org/abs/1902.01738>, 2018.
- [2] B. Afsari. Riemannian  $L^p$  center of mass : existence, uniqueness and convexity. *Proc. Amer. Math. Soc.*, 139(2) :655–6673, 2011.
- [3] M. Arnaudon, F. Barbaresco, and L. Yang. Riemannian medians and means with applications to radar signal processing. *J. Sel. Topics Signal Processing*, 7(4) :595–604, 2013.
- [4] M. Arnaudon, C. Dombry, A. Phan, and L. Yang. Stochastic algorithms for computing means of probability measures. *Stoch. Proc. Appl.*, 58(9) :1473–1455, 2012.
- [5] M. Arnaudon and L. Miclo. Means in complete manifolds : uniqueness and approximation. *ESAIM Probability and statistics*, 18 :185–206, 2014.
- [6] M. Arnaudon, L. Yang, and F. Barbaresco. Stochastic algorithms for computing p-means of probability measures, geometry of Radar Toeplitz covariance matrices and applications to HR Doppler processing. In *International Radar Symposium (IRS)*, pages 651–656, 2011.
- [7] A. Barachant, S. Bonnet, M. Congedo, and C. Jutten. Multiclass brain-computer interface classification by Riemannian geometry. *IEEE Trans. Biomed. Engineering*, 59(4) :920–928, 2012.
- [8] S. Bonnabel. Stochastic gradient descent on Riemannian manifolds. *IEEE Trans. Autom. Control.*, 122(4) :2217–2229, 2013.
- [9] S. Cavallari, V. W. Zheng, H. Cai, K. C. Chang, and E. Cambria. Learning community embedding with community detection and node embedding on graphs. In *Proceedings of the 2017 ACM on Conference on Information and Knowledge Management, CIKM 2017, Singapore, November 06 - 10, 2017*, pages 377–386, 2017.
- [10] B. P. Chamberlain, J. Clough, and M. P. Deisenroth. Neural embeddings of graphs in hyperbolic space. *13th international workshop on mining and learning with graphs. Available at <https://arxiv.org/abs/1705.10359>*, 2017.
- [11] H. Cho, B. Demeo, J. Peng, and B. Berger. Large-margin classification in hyperbolic space. *CoRR*, abs/1806.00437, 2018.
- [12] A. Grover and J. Leskovec. node2vec : Scalable feature learning for networks. *KDD : proceedings. International Conference on Knowledge Discovery & Data Mining*, 2016 :855–864, 2016.
- [13] H. Hajri, H. Zaatiti, and G. Hébrail. Learning graph-structured data using poincaré embeddings and riemannian k-means algorithms. *CoRR*, abs/1907.01662, 2019.
- [14] S. Helgason. *Differential geometry, Lie groups, and symmetric spaces*. American Mathematical Society, 2001.
- [15] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean. Distributed representations of words and phrases and their compositionality. In C. J. C. Burges, L. Bottou, M. Welling, Z. Ghahramani, and K. Q. Weinberger, editors, *Advances in Neural Information Processing Systems 26*, pages 3111–3119. Curran Associates, Inc., 2013.
- [16] A. Narayanan, M. Chandramohan, R. Venkatesan, L. Chen, Y. Liu, and S. Jaiswal. graph2vec : Learning distributed representations of graphs. *CoRR*, abs/1707.05005, 2017.
- [17] M. Nickel and D. Kiela. Poincaré embeddings for learning hierarchical representations. In *Advances in Neural Information Processing Systems 30*, pages 6338–6347. Curran Associates, Inc., 2017.
- [18] J. Pennington, R. Socher, and C. D. Manning. Glove : Global vectors for word representation. In *In EMNLP*, 2014.
- [19] B. Perozzi, R. Al-Rfou, and S. Skiena. Deepwalk : Online learning of social representations. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD ’14*, pages 701–710, New York, NY, USA, 2014. ACM.

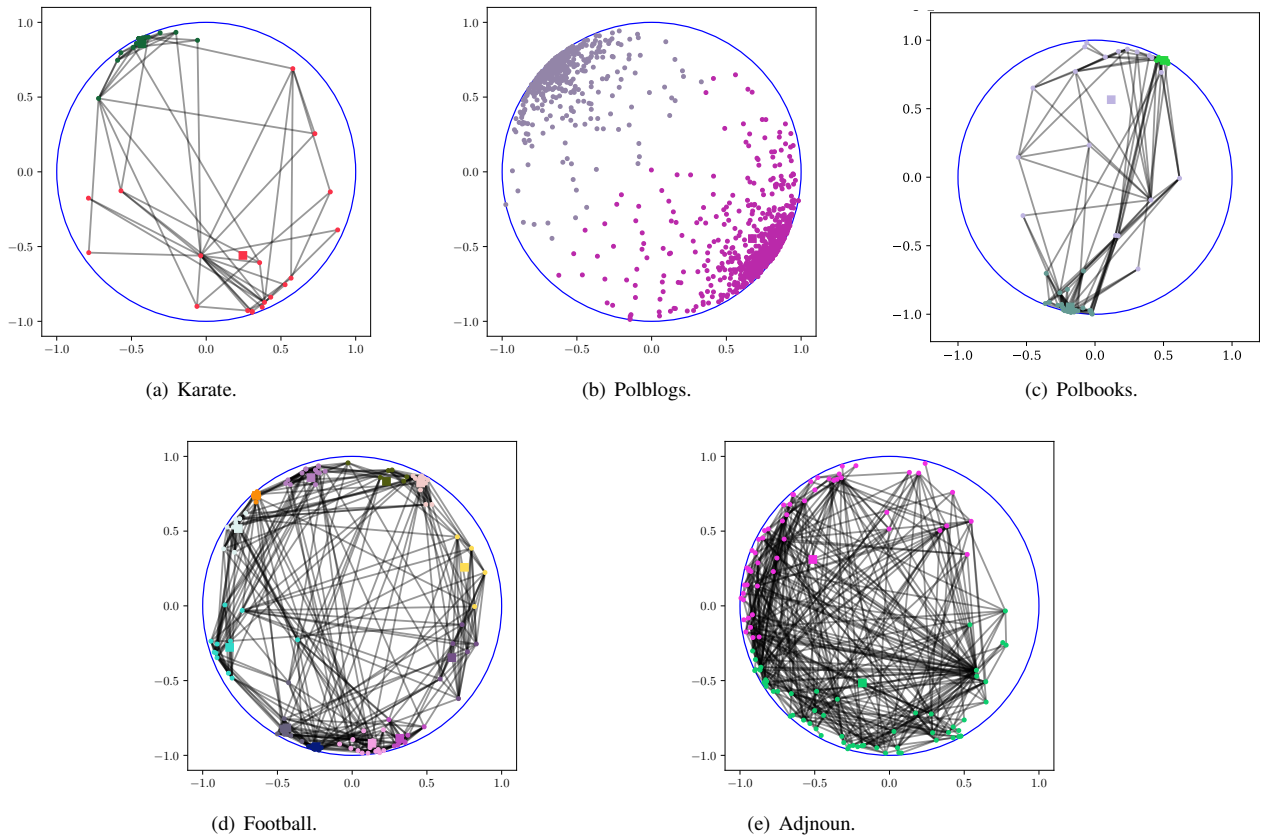


FIGURE 3 – Visualisation des clusters dans le meilleur plongement de Poincaré

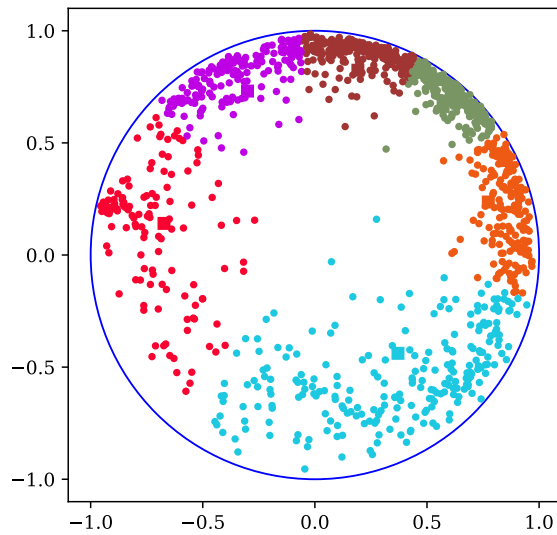


FIGURE 4 – Partitionnement du sous-arbre des mammifères en 6 clusters par l'algorithme des  $K$ -moyennes riemannien.

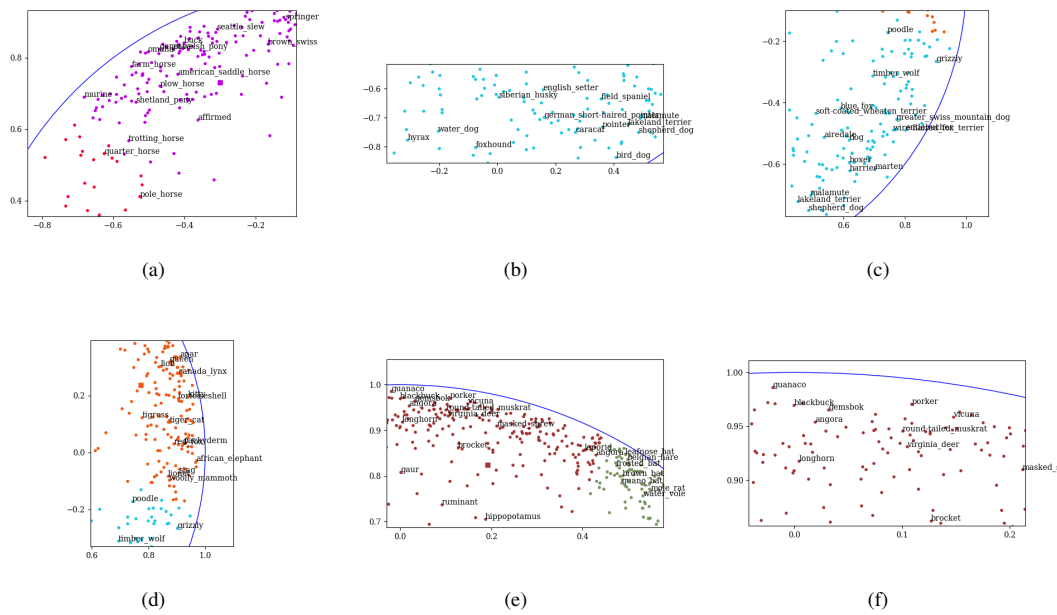


FIGURE 5 – Aperçu des étiquettes du sous-arbre des mammifères : sur les frontières de clusters distincts et à l’intérieur des clusters dans un voisinage du barycentre (ce dernier est représenté par un carré)

[20] R. A. Rossi and N. K. Ahmed. The network data repository with interactive graph analytics and visualization. In *Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence*, 2015.

[21] S. Said, L. Bombrun, Y. Berthoumieu, and J. H. Manton. Riemannian gaussian distributions on the space of symmetric positive definite matrices. *IEEE Trans. Information Theory*, 63(4) :2153–2170, 2017.

[22] S. Said, H. Hajri, L. Bombrun, and B. C. Vemuri. Gaussian distributions on Riemannian symmetric spaces : Statistical learning with structured covariance matrices. *IEEE Trans. Information Theory*, 64(2) :752–772, 2018.

[23] F. Sala, C. D. Sa, A. Gu, and C. Ré. Representation tradeoffs for hyperbolic embeddings. In *Proceedings of the 35th International Conference on Machine Learning, ICML 2018, Stockholmsmässan, Stockholm, Sweden, July 10-15, 2018*, pages 4457–4466, 2018.

[24] A. Tifrea, G. Bécigneul, and O.-E. Ganea. Poincaré glove : Hyperbolic word embeddings. *arXiv preprint arXiv :1810.06546*, 2018.

[25] C. Tu, H. Wang, X. Zeng, Z. Liu, and M. Sun. Community-enhanced network representation learning for network analysis. *CoRR*, abs/1611.06645, 2016.

[26] D. Wang, P. Cui, and W. Zhu. Structural deep network embedding. In *Proceedings of the 22Nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD ’16*, pages 1225–1234, New York, NY, USA, 2016. ACM.

[27] H. Zaatiti et al. Learning community embedding with Riemannian expectation maximisation algorithms. *Work in progress*, 2019.

[28] V. W. Zheng, S. Cavallari, H. Cai, K. C. Chang, and E. Cambria. From node embedding to community embedding. *CoRR*, abs/1610.09950, 2016.

# Trois conceptions du processus : les raisons d'un choix

G. Kassel

Laboratoire MIS, Université de Picardie Jules Verne  
33 rue Saint-Leu, 80039 Amiens Cedex 1

Gilles.kassel@u-picardie.fr

## Résumé

*L'ontologie des entités qualifiées d'« occurrentes » est un domaine actif de recherche. L'ontologie des processus, plus particulièrement, continue de connaître des développements, comme en témoigne l'ouvrage édité en 2018 par Rowland Stout 'Process, Action, and Experience'. Un constat notoire, toutefois, est qu'aucun consensus ne se dégage sur la nature des processus. Au contraire, trois figures du processus sont actuellement promues : (1) le processus comme continuant dynamique, proche de l'objet dans sa façon d'endurer dans le temps ; (2) le processus comme occurrent étendu temporellement et constituant la « matière » d'événements ; enfin, (3) le processus comme forme temporelle abstraite, réalisée par des occurrences concrètes. Dans cet article, nous approfondissons les bases conceptuelles et ontologiques de ces figures. En résultat, nous argumentons en faveur de (1) et en défaveur de (2). Par ailleurs, nous montrons que (3) est compatible avec (1), mais précisons qu'elle caractérise le mouvement plutôt que le processus.*

## Mots Clef

Ontologie formelle, processus physique, événement, changement, changement de Cambridge, changement réel, mouvement

## Abstract

*The ontology of entities described as "occurrences" is an active area of research. The ontology of processes, in particular, continues to experience developments, as evidenced by the book published in 2018 by Rowland Stout's 'Process, Action, and Experience'. However, a notorious finding is that there is no consensus on the nature of the processes. On the contrary, three figures of the process are currently promoted: (1) the process as dynamic continuant, close to the object in its way of enduring in time; (2) the process as a temporally extended occurrent, constituting the "matter" of events; finally, (3) the process as an abstract temporal form, realized by concrete occurrences. In this article, we delve deeper into the conceptual and ontological bases of these figures. As a result, we argue in favor of (1) and against (2). Moreover, we show that (3) is compatible with the (1), but specify that it characterizes the movement rather than the process.*

## Keywords

Formal ontology, physical process, event, change, Cambridge change, real change, movement

## 1 Introduction

Récemment, nous avons entamé un travail consistant à élaborer une nouvelle ontologie des entités dites « occurrentes », parmi lesquelles on distingue couramment les *processus*, *états* et *événements* [21,22,23]. Le cadre ontologique que nous proposons repose sur trois thèses principales : (i) le monde physique est peuplé de particuliers – *objets* et *processus* – qui, en endurent, assurent sa stabilité tout autant que sa dynamique ; (ii) ces particuliers portent temporairement des *propriétés*, ces exemplifications de propriétés correspondant à des *faits* existant dans le monde physique ; par ailleurs, (iii) 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. Avec ce cadre, nous ne cherchons pas à édifier une ontologie de l'*objet* en général ou du *processus* en général, mais nous contentons de caractériser les *objets physiques* et les *processus physiques*. L'inventaire que nous dressons des entités « occurrentes » diffère de l'inventaire courant en ontologie formelle comme en ontologie appliquée. Nous distinguons ainsi, d'une part, les *processus* (mais ceux-ci *endurent* plutôt qu'ils *occurrent*), d'autre part, les *faits dynamiques* (par exemple les faits de *perpétuation* de processus) et, enfin, les *événements* existant pour des sujets et qui *occurrent* lorsqu'ils sont réalisés par des processus [23].

Pour sa part, l'ontologie des processus continue d'être un domaine actif de travaux, en témoigne l'ouvrage collectif édité par Rowland Stout [38]. Par contre, le constat troublant est que les développements proposés ne s'accordent pas autour d'une même figure du processus. Au contraire, ils mettent en scène principalement trois conceptions a priori opposées : (1) le processus comme continuant dynamique, proche de l'objet dans sa façon d'endurer dans le temps [35,36,37,14,15] ; (2) le processus comme occurrent étendu temporellement, constituant la « matière » des événements [26,17,7,8] ; enfin, (3) le processus comme pattern abstrait temporel, réalisé par des occurrents concrets [15,16]. Dans nos travaux, nous promovons la figure (1) mais, ce faisant, nous retenons une classe de processus plus restreinte que celle considérée par Antony Galton et Stout.

Dans cet article, notre objectif est d'approfondir les bases conceptuelles et ontologiques de ces trois figures et de dégager les raisons du choix que nous avons effectué, renforçant par là-même notre cadre ontologique. Pour mener cette analyse, le fil conducteur que nous suivons est l'analyse du *changement* qualitatif (à savoir le fait pour une substance de porter des propriétés contradictoires à différents temps), et plus particulièrement du mouvement

(en tant que changement de localisation spatiale). Nous allons tout d'abord en Section 2 rappeler la contribution d'Aristote à l'analyse du changement, qui est l'une des premières dont nous disposons et qui contient en germes les théories les plus récentes. En visant à justifier la réalité du changement, un apport d'Aristote a été de dégager le *processus* en proposant une analyse *dispositionnelle*, toujours d'actualité. En Section 3, nous effectuons un saut de plus de 2000 ans dans le temps avec la conception du changement promue au début du 20<sup>e</sup> siècle par des philosophes de Cambridge (McTaggart, Russell). Nous en rappelons les critiques et présentons l'analyse dispositionnelle 'révisée' du processus proposée par Carol Cleland [5] pour définir le changement « réel ». En adoptant cette analyse, nous distinguons le '*processus du changement*' du '*changement*' lui-même. Ce préambule étant établi, en Section 4 nous évaluons les figures du processus (1),(2) et (3) précitées. En résultat, nous argumentons en faveur de (1) et en défaveur de (2). Par ailleurs, nous montrons que (3) est compatible avec (1), mais précisons qu'elle caractérise le *mouvement* (et, plus généralement, le *changement*) plutôt que le *processus du changement*.

## 2 Le traitement du changement chez Aristote

Aristote nous livre sa conception du changement dans son ouvrage *Physique*. Ses motivations sont alors doubles : 1) défendre la *réalité* du changement en réponse à Parménide (et plus largement à l'école d'Élée) niant l'existence du changement ; et 2) élucider la *nature* du changement, là où Platon s'était contenté d'en décrire les causes et différentes formes.

La réalité du changement, Aristote entreprend de la défendre avec sa théorie de l'*être* (ou de la *substance*) comme composé de *matière* et d'une *forme*. La matière est ce qui permet à l'être de subsister dans le temps. Quant à la forme, celle-ci détermine la nature (ou *essence*) de l'être. Elle consiste en des attributs (ou *qualités*) dont certains peuvent varier en leur contraire. Ceci explique qu'une substance puisse se trouver dans des états différents pour une même qualité : un être peut être tantôt froid ou chaud, tantôt blanc ou noir et se trouver tantôt en un lieu ou un autre. Aristote distingue trois espèces de changements que sont l'accroissement ou la diminution en taille d'un corps, le changement de qualité (l'altération) et le changement de lieu (locomotion)<sup>1</sup>.

Le changement est par nature, et de façon générale, l'acte du possible, c'est-à-dire la réalisation d'un possible existant en tant que potentialité dans l'être (dans sa forme). Les êtres naturels sont ainsi mus par eux-mêmes. Par exemple, les corps pesants tombent comme une pierre, tandis que les corps légers s'élèvent comme le feu. Dans ces deux cas, selon Aristote, ces corps comportent en privation leur « lieu naturel » vers lequel ils se dirigent. Ces changements consistent en la réalisation de ce lieu naturel. Par ailleurs, les êtres (naturels ou non) peuvent changer du fait d'une cause qui leur est externe, comme dans le cas de phénomènes d'agence. Des exemples favoris d'Aristote sont le fait, pour un morceau de bronze dans l'atelier d'un sculpteur, de devenir une statue, ou le

fait, pour des pierres et du mortier, de devenir une maison. Pour rendre compte de ces changements, naturels ou non, Aristote propose la définition suivante (dans son *livre III*, et selon la traduction d'Ursula Coope [6]) : *Change is the actuality of what is potentially in some particular different state, qua such*.

Aristote propose ainsi une analyse dispositionnelle du changement avec les notions d'*actualité* et de *potentialité* (en des termes contemporains, on parle d'*exercice* d'une *disposition*). Ces notions comportent notamment des parts d'ombre, en témoignent les diverses traductions, interprétations et amendement fournis par les commentateurs d'Aristote, y compris les plus contemporains. Pour éclairer ces notions, nous retenons principalement l'analyse de Coope [6], que nous modulons avec celle d'Andreas Anagnostopoulos [1].

Commençons par la *potentialité*. Selon ces auteurs, cette notion est à entendre au sens d'un potentiel d'une substance à être dirigé vers un état différent, que cet état corresponde à une qualité (potentialité à être froid) ou à une substance (potentialité à être une statue). Ceci suppose que la substance porte en elle cet état différent. Cette caractérisation entraîne que la potentialité disparaisse dès lors que la substance actualise le nouvel état.

L'*actualité*, pour sa part, est à entendre au sens où un potentiel d'une substance à être dans un état différent est *actuel*. L'actualité permet de distinguer deux situations dans lesquelles se trouve une substance, à savoir : (1) le fait de posséder – de façon dormante – un potentiel à être dans un état différent, et (2) le fait que s'exerce – de façon active – ce potentiel. L'actualité est ainsi une activité exercée par une substance.

Si l'on s'en tient à cette caractérisation du changement, la question qui se pose est de savoir si Aristote a effectivement réussi à défendre l'existence du changement et à en préciser la nature. Répondre à cette question suppose que l'on se donne des critères de succès. Il serait ici peu pertinent de passer la théorie d'Aristote au crible des critères contemporains de l'ontologie formelle. Au lieu de cela, tâchons d'évaluer l'apport d'Aristote. Nous apportons de fait une réponse nuancée : sur un plan ontologique (celui de l'existence), Aristote pose des conditions nécessaires et suffisantes pour l'existence du changement, mais ne propose pas de critère d'identité satisfaisant ; sur un plan métaphysique (celui de la nature), la notion de 'potentiel à être dans un état final donné' est, pour le moins, imprécise. Sur un plan ontologique, fondamentalement (selon Anagnostopoulos [1]), l'apport d'Aristote est de proposer avec la notion de potentialité une explication causale du devenir : se limiter (comme c'était le cas des prédécesseurs d'Aristote) à considérer pour une substance deux états '*être*' ou '*ne pas être*' et à constater des différences ne permet pas de rendre compte du devenir des substances, par exemple d'expliquer qu'un gland devienne chêne ou qu'un caillou projeté en l'air par un promeneur retombe par terre ; selon Aristote, le gland et le caillou possèdent respectivement une *potentialité à être un chêne* et une *potentialité à être sur le sol*. Il reste à toutefois à préciser la nature de cette potentialité, à définir un critère d'identité, et c'est là que le bât blesse.

Une des raisons notamment (selon Coope [6]) tient à l'analyse des phénomènes d'agence. Aristote, en effet,

<sup>1</sup> Ce faisant, il écarte la génération et la destruction d'un être au prétexte que, dans de tels phénomènes, aucune substance ne subsiste. Par contre, il

reconnaît qu'un changement puisse être *partiel* ou *absolu* suivant que la substance concernée soit affectée en partie ou en totalité.

identifie le *changement d'une substance à l'action de changer cette substance*. Une telle identification conduit à questionner la nature du potentiel et, tout particulièrement, la confusion régnant entre le physique et le mental.

Une seconde raison tient à la précision de la finalité du changement. Si, dans le cas d'une action, il paraît admissible qu'une finalité précise existe (être une statue ou une maison), celle-ci étant fixée par l'agent, en revanche, dans le cas d'un changement naturel d'une substance, il paraît questionnable qu'un potentiel comporte en soi un état précis, comme le souligne Coope en prenant l'exemple d'une feuille poussée par le vent (*ibid.*)<sup>2</sup> :

*It is hard, then, to see what ground there is for thinking that this [such] movement is the actuality of a potential to be in some particular place, rather than another (...)  
Consider, for instance, a dead leaf that is blown across the street by the wind. Is it really plausible to suppose that its movement is the actualization of some specific potential it has for being on the other side of the street?*

En conclusion, que pouvons-nous retenir du traitement d'Aristote ? Essentiellement, une analyse métaphysique innovante – l'analyse dispositionnelle – rendant compte du caractère dynamique, fluant et causal du changement. Nous lui devons d'avoir initié une caractérisation du processus se déroulant, à savoir notre figure (1). Par contre, son traitement ne rend pas compte d'une dimension importante du changement, à savoir sa caractérisation temporelle et spatiale, par exemple, dans le cas du mouvement, le fait pour une substance d'être dans des états différents à des temps différents. Il va falloir attendre les XVII<sup>e</sup> et XVIII<sup>e</sup> siècles (Descartes, Galilée et Newton) pour que cette dimension soit mise en avant avec la théorie cinématique du mouvement [11].

### 3 Simples « changements de Cambridge » vs « changements réels »

Dans cette section, nous enjambons quelques 2000 ans d'avancées en physique et métaphysique. La conception contemporaine du changement, héritière de la physique classique, fait dépendre le changement du temps. En des termes actuels, le changement se définit comme le fait qu'une substance puisse porter des propriétés contradictoires (*F* et *non F*) à des temps différents. Par exemple, un objet *O* est froid à un temps  $T_1$  et chaud à un temps  $T_2$ . Il s'agit là d'un changement impliquant une qualité de l'objet, considérée comme propriété intrinsèque. Dans le cas d'un mouvement, un objet *O* possède une certaine localisation à un temps  $T_1$  et une autre à un temps  $T_2$ . La localisation correspond à une relation externe que l'objet entretient avec son environnement.

Au début du XX<sup>e</sup> siècle, Bertrand Russell propose la conception suivante du mouvement continu, en tant qu'un changement de localisation spatiale : *Motion consists*

*merely in the occupation of different places at different times* [36]. Ainsi, un mouvement continu d'un objet *O* n'est rien de plus qu'une série de faits correspondant à l'occupation par *O* de différentes positions  $Pos_i$  à des instants successifs  $I_j$  :  $\langle O, Pos_1, I_1 \rangle$ ,  $\langle O, Pos_2, I_2 \rangle$ , etc. Le terme « changement de Cambridge » a été proposé par Peter Geach [18, p. 71] pour dénoter cette conception du changement continu promue par des philosophes de Cambridge dont John McTaggart et Bertrand Russell<sup>3</sup>.

Cette conception a été critiquée par plusieurs philosophes au début du 20<sup>e</sup> siècle, au motif qu'elle ne rend pas compte du caractère dynamique du déplacement. Comme Henri Bergson [3] a pu l'exprimer, le mouvement est traité *comme si il était fait d'immobilités*. La question qui se pose est de savoir si, au-delà du sentiment intuitif que le « compte n'y est pas », il existe des arguments pour faire une place sur le plan ontologique à une entité correspondant à un « processus physique de transition », permettant de parler d'un « changement réel » (pour reprendre le terme de Geach).

On peut tout d'abord noter que la métaphysique contemporaine donne crédit à l'existence de *Faits*, ceux-là mêmes auxquels il est fait référence dans les formulations de la théorie *at-at*. La thèse de l'existence des *Faits* a été défendue notamment par Kit Fine [9] et David Armstrong [2]. Un *Fait* (ou « circonstance », selon la terminologie de Fine, ou encore « état d'affaires », pour reprendre le terme d'Armstrong), est une entité *complexe* constituée d'une substance (ex : 'Paul'), d'une propriété (ex : 'Être à côté de Marie') et d'un temps (ex : 'Maintenant') :  $\langle Paul, Être à côté de Marie, Maintenant \rangle$ <sup>4</sup>. L'existence simultanée à un instant donné d'une substance et d'une propriété ne signifie pas pour autant que la substance exemplifie la propriété à cet instant. Le *Fait* correspond à un lien interne unissant, à un instant donné, substance et propriété/relation en une entité à part entière. L'argument principal de l'existence des *Faits* est qu'ils constituent un *véri-facteur*, autrement dit ce qui rend vrai dans le monde des propositions comme 'Paul est à côté de Marie'. Les *Faits* auxquels il est fait référence dans la théorie *at-at* ont pour constituant une propriété de localisation spatiale d'une substance.

Ontologiquement parlant, on tient donc la brique de base d'une série de faits successifs. Par contre, comme le précise la philosophe Carol Cleland [5], dont nous allons reprendre les propositions, quelle que soit la théorie retenue (y compris celle récente de Graham Priest [28] défendant la thèse qu'un objet peut occuper plusieurs places à un même moment), aucune ne solutionne les limites de la théorie *at-at* à rendre compte de la dynamique du mouvement (*ibid.*, p. 264) :

*Both the at-at theorist and Priest share a common assumption, viz., that moving can be completely analyzed in terms of the mere occupancy of places. As a consequence neither is able to provide us with an account of the transition in place ostensibly involved in the flight of an arrow. What is needed is an account of*

<sup>2</sup> Le même questionnement vaut pour les substances se mouvant par elles-mêmes vers leur lieu naturel : dans quelle mesure possèdent-elles ce 'potentiel à être dans leur lieu naturel' ? Comme on le sait, cette thèse au coeur de la physique d'Aristote a été balayée par les lois de la physique classique, notamment la loi de la gravitation exprimant que des corps s'attirent mutuellement.

<sup>3</sup> Cette conception porte également le nom de théorie « at-at ». Pour rendre compte du caractère continu du changement, la plupart des versions privilégient comme temps l'*instant* indivisible.

<sup>4</sup> Pour des raisons de place, dans cet article, nous supposons sans justifications l'existence réelle, et même physique, du temps et de propriétés comme celles rendant compte d'une localisation spatiale.

*how it is possible for an arrow to get into and out of different places, as opposed to a description of which places just happen to be occupied at different times*

Pour rendre compte de la dynamique du mouvement, la démarche de Cleland, dans son [5], consiste à affiner l'analyse dispositionnelle d'Aristote en la soumettant aux théories physiques contemporaines. Ainsi, pour rendre compte de la notion d'effort, au cœur de la physique Newtonienne, Cleland pose l'existence d'une entité correspondant à un « état actif de mouvement », qu'elle nomme « tendance opérante » (*ibid.*, p. 266) :

*In order to distinguish the sort of tendency which seems to be involved in « endeavouring » from the passive tendencies (or latent capacities) ordinarily associated with dispositional properties, I will frequently refer to “endeavourings” as “operative tendencies”.*

L'existence d'une tendance opérante est justement ce qui distingue le fait, pour un objet, de passer dynamiquement à travers des états plutôt que d'être statiquement dans des états différents. Toutefois, selon Cleland, l'existence d'une tendance opérante, qu'elle nomme également « processus causal de transformation de propriété », ne signifie pas pour autant qu'elle soit suivie d'effets. Ainsi, par exemple, l'existence d'une « tendance opérante à changer de place » reste insuffisante pour qu'un tel changement existe, si cette tendance se voit contrariée dans ses effets par l'existence concomitante d'autres tendances (*ibid.*, p. 273) :

Indeed the failure of some of these tendencies to terminate in the changes towards which they are directed can readily be explained in terms of their lawful interactions with other tendencies; in the case of the globe, the outward centrifugal force is said to be exactly balanced by the inward centripetal force.

L'exemple du globe évoqué ici correspond à une expérience que chacun peut mener en faisant tourner autour de soi un globe (ou tout autre objet) maintenu par une ficelle. Selon Cleland, le fait que l'on puisse sentir la tension dans la ficelle, autrement dit le fait que la tendance de l'objet à être éjecté soit un observable mesurable, constitue un argument décisif pour son existence physique (*ibid.*, p. 273) :

Given their crucial role in physical explanation and theory, I propose that we admit operative tendencies to be elsewhere into our ontology as primitive properties of physical objects. We can think of them as physicists think of instantaneous vector quantities, viz., as uneliminable proclivities of varying degrees of strength.

Finalement, suivant Cleland, nous devenons capables de distinguer ontologiquement les « simples changements de Cambridge » des « changements réels ». Un *changement de Cambridge* est une série temporelle de faits correspondant, pour un objet, à des possessions successives de propriétés distinctes. Un *changement réel* est un changement de Cambridge occasionné, au sein de l'objet, par un processus de transformation de propriétés.

Le fait de préciser « au sein de l'objet » est important en posant comme condition qu'un processus est actif de la part de l'objet et qu'il est la cause du changement de Cambridge. Cela permet d'exclure des situations où, par exemple dans le cas d'un mouvement, un objet est transporté (il se meut donc) sans pour autant que l'objet soit causalement et énergiquement responsable du mouvement. A contrario, des processus peuvent exister sans produire de changement, lorsque leurs effets se voient contrariés par ceux d'autres processus. Un tel engagement ontologique conduit finalement à distinguer le *processus du changement* du *changement* lui-même.

À titre d'illustration, considérons la situation d'un objet exerçant une pression sur un autre objet, par exemple ma main poussant une porte. Tant que la porte résiste à ma poussée, aucun mouvement n'existe. Pourtant, deux processus opposant leurs effets existent. À supposer que ma poussée s'intensifie en force, au point d'ébranler la porte. Nous considérons que cet ébranlement correspond à la naissance d'un nouveau processus – le processus de rotation de la porte sur ses gonds – venant s'ajouter aux deux processus déjà présents (ma poussée de la porte et sa réaction), lesquels continuent d'exister. Ce processus de mouvement conduit à l'ouverture de la porte.

Pour conclure, dressons le bilan de l'avancée que nous fait faire Cleland, par rapport à Aristote, dans l'analyse du changement (qualitatif). Nous avons conclu en Section 2 que Aristote avait contribué à rendre compte du changement *continu* plutôt que du changement *complet*, en évoquant les deux figures ontologiques contemporaines que sont le *processus* et l'*événement*. Avec Cleland, nous précisons en quelque sorte le changement continu pour distinguer, d'un côté, le *processus du changement*, et de l'autre, le *changement* lui-même – ce dernier correspondant au changement Aristotélicien (restreint aux non actions). Dans la suite de l'article, nous allons confronter ce résultat aux dernières conceptions du processus présentées dans la littérature.

#### 4 Trois conceptions du processus : les raisons de notre choix

À l'instar de Cleland, plusieurs auteurs ont défendu ces dernières années une figure du processus comme *continuant dynamique* [35,36,37,12,13,17]. Récemment, nous avons suivi cette voie [21,22,23]. Cette figure continue néanmoins d'être contestée par d'autres auteurs, s'arque boutant au contraire sur une vue du processus comme *occurrent* ne pouvant changer dans le temps [32] et, plus spécifiquement, comme *constituant* la matière d'événements [7,8]<sup>5</sup>. En Section 4.1, nous entendons contribuer aux débats en cours en proposant une caractérisation précise du processus physique sous forme de quatre propriétés essentielles.

Parallèlement, dans ses dernières publications, Galton a proposé une caractérisation des processus les situant en dehors de la strate physique pour en faire des entités abstraites mentales [14,16]. En Section 4.2, nous analysons

<sup>5</sup> En réponse à Stout, la philosophe Helen Steward a également pris part récemment au débat en proposant une position médiane [33,34]. Steward maintient que les processus sont des *occurrents* mais concède que ceux-ci peuvent changer dans le temps, à la manière de *continuant*s. Il s'avère que

la classe des processus considérée par Steward est beaucoup plus large que celle que nous allons retenir, ce qui explique sa position. Quoi qu'il en soit, pour des raisons de place dans cet article, nous ne pourrions pas exposer ses arguments.



sa figure du processus abstrait et proposons de la retenir pour définir le changement lui-même.

#### 4.1 Le processus comme continuant dynamique

La thèse que nous entendons défendre quant à la nature du processus physique comporte un volet intensionnel et un volet extensionnel. En effet, les propriétés que nous attribuons aux processus dépendent étroitement de la classe des individus considérée.

En intension, nous caractérisons le *processus physique* comme quelque chose :

- (p\_i) existant pleinement à des instants ;
- (p\_ii) portant des propriétés à des instants ;
- (p\_iii) pouvant changer dans le temps ;
- (p\_iv) *énaeté* par un objet physique.

En extension, des exemples sont : un processus de déplacement d'un objet physique ou de rotation de l'objet sur lui-même ; la croissance en taille d'un corps physique ; le mûrissement d'un fruit ; l'oxydation d'un objet métallique ferreux ; la fonte d'un glacier. À ces exemples s'ajoutent des processus ayant une cause intentionnelle : marcher, courir. Intuitivement, marcher et courir sont des espèces de processus de déplacement d'une personne.

En défense de cette thèse, nous allons considérer tour à tour chacune des propriétés faisant partie de notre caractérisation des processus physiques.

##### 4.1.1 Les processus existent pleinement à des instants

Avant de prendre position, il convient de préciser en quel sens nous entendons l'expression « exister pleinement ». Cette même expression qualifie habituellement les objets physiques.

Fine [10] a apporté des précisions dans sa défense du 3-dimensionalisme des objets physiques<sup>6</sup>. Selon Fine, lorsque nous affirmons à propos des objets physiques qu'ils « existent pleinement à des temps » (ou qu'ils « existent dans leur entièreté à des temps »), cette expression fait référence à deux notions d'*existence* – respectivement dans le temps et dans l'espace – qu'il convient de distinguer. Dans l'ordre, l'objet existe dans le temps, cette existence n'admettant pas de degré (n'étant pas une question de « plus ou moins ») : on peut penser ici à une existence de l'objet dans sa pleine *identité* (au sens d'*essence*). Par ailleurs, du fait qu'il existe, l'objet matériel est *étendu* dans l'espace et, cette fois, l'expression « pleinement » (ou « dans son entièreté ») traduit le fait qu'il *occupe complètement* une région spatiale (tout en occupant partiellement chaque partie de cette région).

Dans l'énoncé de notre propriété (p\_i), il convient d'entendre l'expression « existant pleinement » au sens d'une existence dans sa pleine *identité*. De ce fait, nous

rapprochons les processus des objets en les assimilant à des continuants 3-D. Ceci pose la question de l'extension temporelle des processus et, corrélativement, de leur existence à des *intervalles* de temps – alors que nous affirmons qu'ils existent à des *instants*.

Rappelons à ce sujet que, dans le domaine de l'ontologie appliquée, l'extension temporelle des processus et la théorie du perdurantisme pour expliquer la persistance des processus sont des principes gravés dans le marbre. Dans l'ontologie BFO [32], le processus est identifié à une entité 4D occurrente. Pour l'ontologie DOLCE, ces principes, déjà solidement ancrés dès la version initiale [25], ont été repris récemment par Nicola Guarino [19], qui a même proposé une analyse plus détaillée des parties temporelles des processus et des événements. Pour Galton, enfin, il ne fait aucun doute qu'un processus est étendu temporellement et que, corrélativement, le processus existe à un intervalle de temps<sup>7</sup>.

La défense du 3-dimensionalisme des processus que nous opposons se fonde avant tout sur la classe des processus considérée. Rappelons que, pour nos processus, nous avons retenu la notion de *processus de transformation de propriété* de Cleland [5]. Notre défense se fonde également sur le fait de considérer des événements *abstrait*s.

Pour la majorité des auteurs précisant actuellement la catégorie de processus physique, celle-ci recouvre des processus tels 'Ecrire une lettre', 'Remplir un formulaire' ou 'Donner une conférence' qui, à l'évidence, ne correspondent pas à de simples « processus de transformation de propriété ». Une analyse courante de ces phénomènes d'agence téléiques (comportant en soi une fin) revient à identifier, d'une part, un événement comme occurrence complète (ex : le fait accompli d'écriture d'une lettre) et, d'autre part, un processus le constituant. Une telle analyse, comme nous l'avons rappelé en Section 2 en nous référant à Coope [6], est clairement héritée de l'analyse Aristotélicienne des processus, avant d'avoir été renforcée par la thèse de la constitution d'événements par des processus [26]. Cet héritage, consistant à considérer des processus téléiques ayant vocation à s'achever en un événement complet, est notamment assumé par Charles [4]. L'argument retenu pour justifier l'existence de tels processus est de nature linguistique. Comme on peut le voir chez Stout [45], les processus sont définis comme étant ces entités occurrentes pour lesquelles on peut dire qu'elles étaient/sont/seront en train d'*occurrer* (*avoir lieu, se passer*) : [*Processes*] are things that were, are or will be happening – like someone reading or my writing this paper for instance.

Dans [21,23] nous avons proposé un cadre ontologique étoffé (comportant, à côté des processus physiques, des *événements abstraits* et des *faits physiques* temporaires) permettant de dénier l'existence de tels 'processus d'événements'. Dans cet article, pour des raisons de place,

<sup>6</sup> Le 3-dimensionalisme s'oppose à la thèse du 4-dimensionalisme selon laquelle les objets sont étendus à la fois spatialement et temporellement, et possèdent de fait des parties temporelles. Une *même* personne (selon la vue 3D) existant à des temps différents  $T_1$  et  $T_2$  est considérée selon la vue 4D comme *deux* personnes *différentes* à ces temps  $T_1$  et  $T_2$ . Plus exactement, selon la vue 4D, la personne au temps  $T_1$  et la personne au temps  $T_2$  sont considérées comme étant des parties temporelles distinctes d'une même et unique personne. En conséquence, un objet 4D ne peut changer qualitativement, selon la conception du changement (de sens commun) que

nous avons adoptée dans l'article. Le lecteur intéressé trouvera une défense du 4-dimensionalisme dans [30].

<sup>7</sup> Selon Galton, les processus n'existent pas à des instants – comme l'exprime notre propriété (p\_i) – mais à des intervalles de temps, et ceci vient contredire la théorie 'présentiste' du temps selon laquelle seuls des instants-présents existent [15, p. 167] : *This idea [selon laquelle les processus sont des particuliers concrets] raises problems for the traditional instant-based model of time, since processes, being inherently temporally extended, can only exist over intervals, not at instants.*

nous nous contentons de résumer notre analyse ontologique. En premier lieu, signalons que, pour la notion d'événement, nous retenons un événement *abstrait* correspondant à un construit humain. Concernant l'analyse des phénomènes d'agence comme 'Écrire une lettre', nous identifions bien un événement, par contre notre conception non téléique des processus nous interdit de considérer une espèce de 'macro' processus. À la place, nous considérons que la *réalisation* de phénomènes d'agence tel 'Écrire une lettre' donne lieu à de nombreux processus physiques (ex : des gestes d'écritures, des processus impliquant un crayon). L'événement n'est pas *constitué* de ces processus mais *est réalisé par eux*, lorsqu'il occure. De fait, l'argument linguistique évoqué par Stout (cf. supra) devient un argument en faveur de l'existence d'événements. Nous affirmons que les entités *occurentes* pour lesquelles on peut dire qu'elles étaient/sont/seront en train d'*occure* (*avoir lieu, se passer*) sont des événements, et non des processus. Il s'agit d'événements en train d'être réalisés, comme pour : « Je suis en train d'écrire une lettre », « j'étais en train de traverser la rue », « je serai en train de partir en vacances », ou encore « une bagarre entre deux hommes à l'extérieur du nightclub est en cours »<sup>8</sup>.

En restreignant ainsi notre classe de processus, par rapport aux théories courantes, non seulement nous défendons le 3-dimensionalisme de nos processus mais encore nous défendons qu'ils puissent exister à des instants et non à des intervalles de temps.

#### 4.1.2 Les processus portent des propriétés à des instants

En retenant comme notion d'existence celle d'une *identité/essence* endurent dans le temps, nous admettons que les processus portent des propriétés à des temps – notre propriété (p<sub>ii</sub>), particularisée aux instants. Certaines propriétés sont essentielles en constituant l'identité du processus, tandis que d'autres sont contingentes et correspondent à des manières d'être temporaires des processus.

La phénoménologie des processus nous indique que ceux-ci peuvent être, à des instants : *rapides/lents, bruyants/silencieux, chaotiques/réguliers*. On peut noter qu'ils ont des manières d'être distinctes de celles des objets [20] : un processus n'a pas de couleur, de masse ou de volume, en revanche il se caractérise par sa vitesse, sa direction, sa sonorité, son amplitude. Ce constat justifie du reste que les objets et les processus soient considérés comme appartenant à deux classes distinctes de continuants.

En complément de porter des propriétés, on notera que les processus entretiennent temporairement des relations avec d'autres processus. Reprenons notre exemple d'une situation où un objet, en exerçant une poussée sur un autre objet, le fait se mouvoir. Selon notre analyse, dans une telle situation, un processus de mouvement du premier objet 'perpétue' un processus de mouvement du second objet, cette relation correspondant à une propagation de causalité.

<sup>8</sup> Dans cette dernière phrase, nous considérons que le terme 'bagarre entre deux hommes' se réfère à un événement. Ce même exemple est traité par

Dans la littérature récente, deux auteurs – Barry Smith [32] et Thomas Crowther [7,8] – dénie aux processus toute possibilité de porter des propriétés à des instants. Leur thèse (non p<sub>ii</sub>) est fondée sur le fait d'assimiler les processus à des *occurents* 4D. Au-delà, chacun des chercheurs met en avant des arguments différents. Pour illustrer les arguments convoqués, considérons la situation suivante : *Marie marche à la vitesse de 4km/h à l'instant T*. Rappelons que, selon notre traitement, un processus particulier de marche est actif et ce processus porte la propriété '*A pour vitesse 4km/h*' à l'instant *T*. De son côté, Smith préconise dans BFO de considérer la vitesse comme une propriété de l'objet se mouvant [32, p. 479] : *Note that we could view speed in BFO terms as a (non-rigid) quality of the moving object, a view conformant with our way of speaking when we talk, for example, of the speed of light, or the speed of the earth, or the speed of a billiard ball*. Le problème que nous voyons est le manque de sémantique de cette propriété. Lorsque nous parlons de la vitesse d'une boule de billard, il est implicite qu'il s'agit de sa vitesse de déplacement. Mais, à supposer qu'un objet participe simultanément à plusieurs processus, comment faire pour savoir à quel processus rattacher la propriété (sauf à considérer une propriété comme 'vitesse de marche', ce que Smith s'interdit) ? Ce problème a été soulevé par Galton et Mizoguchi [17, p. 79] :

To say that Mary is slow at a particular time is meaningless unless we specify in what respect she is slow; this could be any range of activities such as walking, speaking, thinking, etc. and since one such activity must be specified in ascribing slowness to Mary it is clear that it is the activity rather than Mary herself that is described as slow.

Pour Crowther [7], plutôt que de considérer qu'un processus de marche porte la propriété '*A pour vitesse 4km/h*' au temps *T*, il convient de considérer que l'objet se mouvant – Marie – porte la propriété '*Marcher à la vitesse de 4km/h*', toujours au temps *T*. On notera qu'un tel traitement n'est pas opposé au nôtre, à condition d'admettre les deux propriétés. La seconde propriété est pour nous une propriété relationnelle mettant en relation Marie et un processus particulier de marche, ce dernier ayant une propriété de vitesse. Mais, si cette dernière attribution est interdite, comment rendre compte de la sémantique de la propriété '*Marcher à la vitesse de 4km/h*' ? Crowther peut identifier la propriété '*Marcher*' à un type de processus, mais quel lien faire avec une vitesse, si le processus est assimilé à un *occurent* étendu ? Nous estimons que le problème est repoussé, plutôt que résolu.

#### 4.1.3 Les processus peuvent changer dans le temps

Venons-en à notre propriété (p<sub>iii</sub>), à savoir la possibilité pour les processus de changer en portant des propriétés contraires à des temps différents. Cette thèse a notamment été soutenue par Galton [12, p. 6] :

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 rythm, or may stop, only

Stout au moyen de la phrase : « *This is a process of two men fighting outside the nightclub* ». On notera que la référence au processus contraint à recourir à un type de phrase ne relevant pas du langage courant.

to start again later. The flow of the river becomes turbulent; the wind veers to the north-west.

Considérons, dans la lignée des exemples que nous venons de prendre, un épisode de marche d'une personne et le fait qu'au cours de cet épisode cette personne à un instant « hâte le pas », autrement dit accélère. Notre analyse ontologique est la suivante : à chaque instant de l'épisode de marche, un processus de marche est pleinement présent (p\_i) ; ce processus porte à chaque instant une vitesse (p\_ii) ; à l'instant où la personne hâte le pas, la vitesse du processus de marche change de magnitude. C'est donc bien le processus qui a changé.

Cette propriété (p\_iii) découle directement de (p\_i) et (p\_ii). Pour renforcer cette thèse, nous ajouterons le constat selon lequel nous pouvons agir sélectivement sur des processus de sorte à les modifier, voire à les bloquer. Par exemple, nous pouvons agir sur un fruit pour (1) ralentir ou au contraire accélérer son mûrissement, mais également, et le cas échéant, pour (2) stopper sa chute. D'une façon générale, nous avons l'habitude, dans nos actions courantes, de moduler nos efforts pour faire varier proportionnellement des processus, par exemple lorsque nous poussons avec plus ou moins d'intensité des objets. Dans une situation de conduite automobile, en appuyant plus ou moins sur la pédale d'accélérateur nous faisons varier la vitesse de notre voiture.

Il est inutile de reprendre ici dans le détail les critiques formulées par les tenants de la vision 'occurrent 4D' des processus. Leur argument principal revient à dire que, comme un processus est déjà un changement, il ne peut lui-même changer. De notre côté, a contrario, c'est bien le fait de distinguer le 'processus du changement' du 'changement' lui-même qui nous permet d'accréditer la thèse du changement du processus. Une question reste toutefois en suspend qui est de savoir si le changement de processus peut être rapproché du changement d'objet. La réponse que nous apportons, suivant toujours (p\_i) et (p\_ii), est qu'il s'agit dans les deux cas d'un changement de propriétés dans le temps.

#### 4.1.4 Les processus sont énoncés par des objets physiques

Pour finaliser la caractérisation de nos processus physiques, évoquons un dernier engagement ontologique, à savoir le fait qu'un processus ne soit pas un continuant flottant dans l'air, mais soit ancré dans un objet support (notre propriété (p\_iv)) : 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'énonction introduite par Galton et Mizoguchi [17]. Pour ces auteurs, dire qu'un objet « énonce » un processus revient à dire qu'un objet porte un processus « externe » ou « comportement » (*ibid.*, 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'énonction par Galton et Mizoguchi repose sur une conception de l'objet comme interface entre des processus internes et externes. Cette conception, à son tour, met en avant une double hiérarchie de relations de constitution (et donc de dépendances existentielles) entre processus et objets situés à différents niveaux de la strate physique : un processus de marche d'une personne n'est possible que si des processus physiologiques énoncés par les organes de la personne existent concomitamment ; ces derniers ne sont possibles que si des processus énoncés par des tissus, des cellules, des molécules, etc. existent concomitamment<sup>9</sup>.

Dans cette section, nous venons de défendre la figure du processus comme continuant dynamique concret en réfutant celle du processus comme occurrent constituant d'événements. Ce faisant, pour revenir à notre fil conducteur du changement, nous avons identifié le processus à un « moteur » du changement ou 'processus du changement'. La figure du processus abstrait défendue par Galton [14,16] va nous donner l'occasion d'accorder au changement lui-même un statut ontologique.

## 4.2 Le processus comme pattern abstrait d'occurrence

Depuis 2012, Galton promeut une figure ontologique différente du processus. Nous pouvons même dire « très différente »<sup>10</sup> puisque, d'une entité concrète, Galton en fait une entité abstraite [14, Abstract] : *We regard processes as abstract patterns of behaviour which may be realised in concrete form as actually occurring states and events*. Pour appréhender cette nouvelle figure, voyons en quel sens Galton entend les termes « abstrait », « patron de comportement », « réalisation », « état » et « événement ». Dans [14], Galton prend soin de préciser deux notions, celles d'*abstraction* et de *réalisation*. Notons tout d'abord que, pour Galton, états et événements sont respectivement des continnants et des occurrents concrets (Galton continue de souscrire à la notion Davidsonienne d'événement). Plus précisément, pour Galton, un *état* est caractérisé comme un comportement *continuable* (ex : marcher) tandis qu'un *événement* est un comportement *répétable* (ex : marcher jusqu'à la gare)<sup>11</sup>. Dès lors, un *processus* est un type d'état/événement, la relation type-instance correspondant à une *réalisation* (*ibid.*, §6 Realisations) : *Such instantiations are concrete realisations of these continuable or repeatable behaviours. As such they are fully determinate with respect to their spatial, temporal, and indeed all other characteristics*. Le processus est ainsi une entité abstraite :

The category of process is neither subordinate to nor

<sup>9</sup> Dans [31], Peter Simons défend une thèse de l'endurance des objets physiques mettant en scène cette même double hiérarchie de relations de constitution entre des objets et des 'occurrences', ces dernières s'avérant correspondre à nos processus.

<sup>10</sup> Pour Galton, l'intention n'est pas de compléter le cadre précédemment établi dans [12,13,17], mais bien de le réviser radicalement. Ainsi écrit-il [14, p. 35] : *In this paper, I advocate a point of view which is in some respects utterly at variance from those expressed by the authors*

*represented in Figure 1 [Allen, Moens & Steedman, Mourelatos, Pustejovsky, Sowa] – and indeed from my own previous publications on this subject.*

<sup>11</sup> La différence est qu'un comportement comme 'marcher' peut être poursuivi dans l'instant (cette poursuite étant toutefois conditionnée à une décision de mettre fin au comportement ou à un état physique permettant cette poursuite), tandis que le comportement 'marcher jusqu'à la gare' ne peut être poursuivi, une fois arrivé à la gare, mais peut être répété quotidiennement.

superordinate to the categories of state and event; nor is it on the same footing as them with some immediate common superordinate category. On the contrary, processes belong in a completely different realm, the realm of abstract entities, patterns if you will, quite separate from the realm of spatio-temporal entities which includes both states and events.

Il reste à comprendre ce que Galton entend par “patron de comportement”. La définition de cette notion fait l’objet de [16]. En introduction, à titre d’exemple, Galton considère le processus de marche (*ibid.*, p. 41) :

The process of walking, for instance, may be characterized by a particular pattern of movement, alternating forward swings of the legs resulting in an overall forward movement of the body. An actual realization of this pattern, viewed synoptically, is an event which consists of someone’s starting to walk (that is, to realize the walking pattern), walking for a while, and then stopping. Viewed experientially, from moment to moment, we see a succession of instantaneous states, each of which may be characterized as a walking state, that is, a state in which the disposition and state of motion of the body parts is characteristic of one phase of the walking pattern.

Cette caractérisation du processus de marche appelle deux remarques. Tout d’abord, l’expression « patron de mouvement » (spécialisée, à l’occasion, en « patron de marche ») utilise le terme « mouvement », sans que cette notion soit définie. Plus précisément, Galton ne ménage pas de place au mouvement dans son inventaire ontologique. Par ailleurs, une nouvelle catégorie d’état est introduite. Nous avons (dans [14]) l’état de *comportement* en tant que *comportement continuable*. Désormais, nous avons une succession d’états *instantanés*, chacun d’entre eux étant « caractérisé comme un état de marche », autrement dit comme un *état de comportement*. On voit ici un problème d’incohérence ontologique : comment l’état de *marche*, qui est un particulier concret, peut-il « caractériser » (dans une relation ne pouvant être que type-instance) des *états instantanés* ?

Reprenons, pour notre part, l’analyse du mouvement là où nous l’avons laissée en fin de Section 3. Nous allons montrer qu’en cherchant à rendre compte complètement du mouvement, nous arrivons à faire une place dans notre inventaire ontologique à une entité proche du *patron abstrait d’occurrence* de Galton.

Pour caractériser le changement/mouvement continu, nous avons mis en avant (1) le processus du mouvement (notre processus physique), lequel (2) se manifeste par une série de faits instantanés (ces faits instantanés successifs résultent du processus). Nous avons fait une place dans notre inventaire ontologique aux faits instantanés (rappelons-le, en mobilisant les notions proches de *circonstance* de Fine [9] et d’état d’affaires d’Armstrong [2]). Par contre, nous avons laissé de côté la série – en tant que telle – de faits instantanés, pour la raison que nous n’acceptons pas dans notre inventaire

ontologique d’entités existant sur un intervalle de temps. De fait, nous avons laissé de côté le mouvement. Comment en rendre compte ?

Une façon simple de traiter le mouvement (continu) est de faire appel à notre capacité de perception. Notons tout d’abord que, sous réserve que le mouvement ne soit pas trop rapide ou trop lent pour que nous l’identifions, nous percevons les faits instantanés successifs de localisation de l’objet se mouvant. Dans le cas d’un mouvement global de l’objet (lorsque toutes ses parties changent de localisation spatiale), la série successive de faits n’est, ni plus ni moins, qu’une trajectoire dans une région spatio-temporelle. Or, et c’est une seconde donnée de la perception que nous mobilisons, cette trajectoire est quelque chose que nous percevons globalement, que nous unifions, et à laquelle nous attribuons une forme. Il paraît dès lors naturel d’assimiler le mouvement à cette trajectoire et de parler de la forme du mouvement, ce qu’attestent des expressions comme « mouvement *rectiligne/circulaire/oscillatoire* »<sup>12</sup>. Cette proposition nous paraît être en bonne cohérence avec la figure du processus abstrait comme patron d’occurrence défendue par Galton, tout en répondant aux deux remarques que nous formulons supra. D’une part, nous caractérisons ontologiquement le mouvement et, plus largement, le changement continu. Sur ce point on notera que, suivant la nature du processus et de sa manifestation, les faits instantanés peuvent varier et être, par exemple, des faits de sonorité distincts, mobilisant une modalité de perception différente, en l’occurrence l’ouïe. D’autre part, nous distinguons bien les deux catégories d’état confondues, ou en tout cas mal caractérisées, selon nous, par Galton. La première, l’état de *comportement*, correspond à notre *processus du changement*. La seconde, l’état *instantané*, correspond à nos *faits instantanés* successifs.

En synthèse, nous proposons de positionner l’ensemble des entités ontologiques que nous retenons en les illustrant sur l’analyse d’un épisode de marche comme suit (le lecteur pourra ainsi comparer avec l’analyse de Galton rappelée supra) :

Lors d’un épisode (continu) de marche d’une personne, un processus de marche énoncé par cette personne est pleinement présent à tout instant que dure l’épisode. Si l’intention de la personne est de se rendre d’un point A à un point B, le processus est le moyen de réaliser cet événement – son déplacement de A à B. Sans intention particulière ni raison de penser à sa marche, aucun événement n’existe. A noter toutefois qu’un observateur peut prêter à la personne l’intention d’un déplacement, auquel cas le processus sera considéré comme le moyen mis en œuvre pour réaliser l’événement – pensé cette fois par l’observateur. Au cours de l’épisode de marche, le processus peut être irrégulier et changer, par exemple, de vitesse ou de direction. Un observateur, percevant le déplacement en cours de la personne, l’identifie à un mouvement. Il reconnaît ainsi la forme caractéristique d’un mouvement de marche, caractérisée par la forme globale du déplacement mais aussi, et surtout, par la

<sup>12</sup> Pour conforter cette identification du mouvement à une conceptualisation d’une trajectoire, nous pouvons citer, en sémantique cognitive, les travaux du linguiste Ronald Langacker [24] portant sur la

caractérisation de verbes de ‘mouvement physique’. Selon Langacker, des verbes comme *aller*, *partir*, *grimper*, *rouler*, etc., ont pour sens une conceptualisation d’un élément mobile occupant successivement des positions différentes dans le temps.

forme des mouvements des jambes et du torse de la personne. Ces mouvements étant régulièrement coordonnés et répétitifs se traduisent par un motif se répétant dans la forme du mouvement de marche.

## 5 Remarques concluantes et perspectives

Dans cet article, nous avons précisé la nature ontologique du processus physique et, chemin faisant, ayant pris comme fil conducteur l'analyse du changement et du mouvement (en tant qu'une espèce de changement), nous avons en quelque sorte poursuivi le projet d'Aristote de défense de la réalité du changement. Adoptant toutefois une conception contemporaine de l'ontologie en admettant deux types de réalité – physique et mentale (et sociale) – nous avons positionné le changement dans la réalité mentale comme une espèce d'événement. La Figure 1 synthétise graphiquement le cadre ontologique global auquel nous parvenons.

Ce cadre ontologique renouvelle très largement la distinction entre 'continuants' et 'occurrents', prégnante en ontologie formelle et largement adoptée en ontologie appliquée. Rappelons qu'il repose sur des positions encore minoritaires à la fois en ontologie des processus (où la figure du processus comme matière constituante d'événements continue de régner) et en ontologie des événements (où la thèse d'événements concrets reste prédominante). Il repose également, en philosophie du temps, sur une thèse présentiste (soutenant que seul le présent existe, sous forme d'instant) faisant l'objet de nombreux débats. Il repose enfin sur la thèse de l'existence de faits physiques donnant lieu également à de nombreux débats. Par là-même, nous soulignons que la défense de ce cadre nécessite encore des efforts importants.

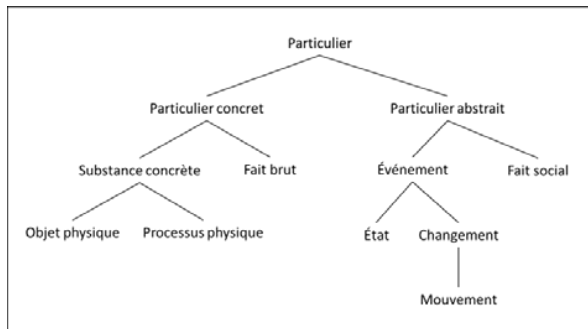


Figure 1 – Vue d'ensemble de notre cadre ontologique

## Références

[1] A. Anagnostopoulos, Change in Aristotle's Physics 3, *Oxford Studies in Ancient Philosophy*, 39, pp. 33-79, 2010.  
 [2] D.M. Armstrong, *A world of states of affairs*, Cambridge University Press, 1997.  
 [3] H. Bergson, *The creative Mind*, New York, Philosophical Library, 1946.  
 [4] D. Charles, Processes, Activities, and Actions, In R. Stout (ed.), *Process, Action, and Experience*, Oxford University Press, pp. 20-40, 2018.  
 [5] C.E. Cleland, The Difference Between Real Change and Mere Cambridge Change, *Philosophical Studies*, 60, pp. 257-280, 1990.  
 [6] U. Coope, Change and its relation to actuality of potentiality, In G. Anagnostopoulos (ed.), *A companion to Aristotle*, Blackwell Publishing, pp. 277-291, 2009.  
 [7] T. Crowther, The Matter of Events, *The Review of Metaphysics*, 65(1), pp. 3-39, 2011.  
 [8] T. Crowther, Processes as Continuants and Processes as Stuff, In R.

Stout (ed.), *Process, Action, and Experience*, Oxford University Press, pp. 58-81, 2018.  
 [9] K. Fine, First-Order Modal Theories III – Facts, *Synthese*, 53, pp. 43-122, 1982.  
 [10] K. Fine, In Defense of Three-Dimensionalism, *The Journal of Philosophy*, 3(12), pp. 699-714, 2006.  
 [11] A. Fouillée, Le Système du monde selon Descartes et selon la science contemporaine, *Revue des Deux Mondes*, Paris, 1982.  
 [12] A. Galton, On What Goes On: The ontology of processes and events, In R. Ferrario & W. Kuhn (eds.), proc. of the *Fourth International Conference on Formal Ontology in Information Systems (FOIS2006)*, pp. 4-11, 2006.  
 [13] A. Galton, Experience and History: Processes and their Relation to Events, *Journal of Logic and Computation*, 18(3), pp. 323-40, 2008.  
 [14] A. Galton, The ontology of states, processes, and events. In M. Okada & B. Smith (eds.), *Interdisciplinary Ontology: Proc. of the Fifth Interdisciplinary Ontology Meeting*, Open Research Centre for Logic and Formal Ontology, Keio University, Tokyo, Japan, pp. 35-45, 2012.  
 [15] A. Galton, The Dynamic Present, In P. Hasle, P. Blackburn & P. Ohrstrom (eds.), *Logic and Philosophy of Time: Themes from Prior*, Aalborg University Press, pp. 167-187, 2017.  
 [16] A. Galton, Processes as Patterns of Occurrence, In R. Stout (ed.), *Process, Action, and Experience*, Oxford University Press, pp. 41-57, 2018.  
 [17] A. Galton & R. Mizoguchi, The water falls but the waterfall does not fall: New perspectives on objects, processes and events, *Applied Ontology*, 4, pp. 71-107, 2009.  
 [18] P.T. Geach, *God and the Soul*, London: Routledge and Kegan Paul, 1969.  
 [19] N. Guarino, On the semantics of ongoing and future occurrence identifiers, In H.C. Mayr et al. (eds.), *Conceptual Modeling, Proc. of the 36th Int. Conf. ER 2017*, LNCS Springer, pp. 477-490, 2017.  
 [20] P.M.S.Hacker, Events and Object in Space and Time, *Mind*, 91, pp. 1-19, 1982.  
 [21] G. Kassel, Processus, événements et couplages temporels et causaux, *Revue d'Intelligence Artificielle*, 31(6), pp. 649-679, 2017.  
 [22] G. Kassel, Une alternative à la distinction 'continuant' vs 'occurrent', In S. Ranwez (ed.), *Actes des 29èmes Journées Francophones d'Ingénierie des Connaissances (IC 2018)*, Nancy, juillet, pp. 147-162, 2018.  
 [23] G. Kassel, Processes Endure, Whereas Events Occur, In S. Borgo et al. (eds.), *Ontology Makes Sense: Essays in honor of Nicola Guarino*, IOS Press, Frontiers in Artificial Intelligence and Applications, vol. 316, pp. 177-193, 2019.  
 [24] R.W. Langacker, Mouvement abstrait, *Langue française*, 76, pp. 59-76, 1987.  
 [25] 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.  
 [26] A.P.D. Mourelatos, Events, Processes, and States. *Linguistics and Philosophy*, 2(3), pp. 415-434, 1978.  
 [27] R. Poli, Levels of Reality and the Psychological Stratum, *Revue internationale de philosophie*, 2(236), pp. 163-180, 2006.  
 [28] G. Priest, Inconsistencies in Motion, *American Philosophical Quarterly*, (22), pp. 339-346, 1985.  
 [29] B. Russell, *Principles of Mathematics*. Cambridge, UK: Cambridge University Press, 1903.  
 [30] T. Sider, *Four-dimensionalism: An Ontology of Persistence and Time*, Oxford University Press, Oxford, 2001.  
 [31] P. Simons, Continuants and Occurrents, In Proc. of the *Aristotelian Society*, Supplementary Volumes, 74, pp. 59-75, 2000.  
 [32] B. Smith, Classifying Processes: An Essay in Applied Ontology, *Ratio*, 25(4), pp. 463-488, 2012.  
 [33] H. Steward, Processes, Continuants, and Individuals, *Mind*, 122(487), pp. 781-812, 2013.  
 [34] H. Steward, What is a continuant? In Proceedings of the *Aristotelian Society*, Volume: LXXXIX, pp. 109-123, 2015.  
 [35] R. Stout, Processes. *Philosophy*, 72(279), pp. 19-27, 1997.  
 [36] R. Stout, The life of a process, In G. Debrock (ed.), *Process Pragmatism: Essays on a Quiet Philosophical Revolution*, Rodopi, pp. 145-57, 2003.  
 [37] R. Stout, The category of occurrent continuants, *Mind*, 125(497), pp. 41-62, 2016.  
 [38] R. Stout (ed.), *Process, Action, and Experience*, Oxford University Press, 2018.



# Vers un désenchevêtrement de l’ambiguïté de la tâche et de l’incertitude du modèle pour la classification avec option de rejet à l’aide de réseaux neuronaux

T. Lorieul<sup>1</sup>A. Joly<sup>1</sup><sup>1</sup> Zenith, LIRMM, Université de Montpellier, Inria, France

titouan.lorieul@gmail.com

## Résumé

La classification avec option de rejet est un moyen d’aborder le problème de l’estimation de l’incertitude d’un classifieur. Les approches récentes s’attaquant à ce problème utilisent des critères basés sur une mesure, soit, de confiance, soit, de dispersion. Cependant, aucune d’entre elles ne combine explicitement les deux principales sources d’incertitude : l’ambiguïté de la tâche, intrinsèque à celle-ci, et l’incertitude du modèle, découlant de l’échantillonnage des données et de la stochasticité de l’apprentissage. Dans cet article, nous explorons comment ces deux quantités peuvent être fusionnées afin d’établir des critères de rejet plus efficaces. En particulier, nous proposons une série de méthodes combinant des mesures de désaccord et des estimations de l’ambiguïté en utilisant un ensemble de modèles. Des expériences sur des jeux de données synthétiques construits pour modéliser différents types d’incertitudes indiquent que ces nouveaux critères ont des performances similaires aux méthodes de référence. Néanmoins, des analyses plus approfondies montrent des indices empiriques qui mettent en avant l’existence d’information supplémentaire dans la distribution des résultats de l’ensemble. Dans les faits, le réjecteur idéal peut être une fonction plus complexe que les critères précédents, et peut même parfois être contre-intuitif.

## Mots Clef

Classification avec option de rejet, estimation d’incertitude, réseaux neuronaux, ensemble de modèles.

## Abstract

Classification with reject option is a way to address the problem of estimating the uncertainty of a classifier. Recent approaches to this problem use criteria based on either a confidence or a dispersion measure. However, they do not explicitly combine the two main sources of uncertainty : the ambiguity of the task, inherent to it, and the uncertainty of the model, resulting from data sampling and stochasticity of learning process. In this article, we explore how these two quantities can be merged to build more effective rejection criteria. In particular, we propose methods for combining disagreement measures and ambiguity estimates using

an ensemble of models. Experiments on synthetic data sets constructed to model different types of uncertainties indicate that these new criteria have similar performance to the baselines. Nevertheless, more in-depth analyses show empirical evidence that highlights the existence of additional information in the distribution of the overall results. In practice, the ideal rejector may be a more complex function than the previous criteria, and may even be counter-intuitive at times.

## Keywords

Classification with reject option, uncertainty estimation, neural networks, ensembles.

## 1 Introduction

Il est important de disposer de mesures précises de l’incertitude des prévisions d’un modèle dans de nombreux scénarios pratiques où l’on ne peut pas se permettre de commettre des erreurs. Cela est en particulier vrai dans des applications médicales, de conduite autonome, etc. Cependant, quantifier précisément cette information d’incertitude est un problème difficile, surtout lorsque le processus d’apprentissage n’est pas entièrement compris comme cela est le cas pour les réseaux neuronaux. Une façon d’assouplir cet objectif ambitieux, tout en progressant dans cette direction, est de permettre aux classifieurs de refuser de donner une réponse pour une entrée donnée. Ceci est connu sous le nom de *classification avec une option de rejet* [5, 14]. Cette décision de rejeter peut tirer parti de l’incertitude prédictive sans avoir à la modéliser complètement et ainsi permettre de mieux comprendre ce qui est nécessaire pour construire des modèles fournissant des informations d’incertitude plus complètes.

La classification avec option de rejet consiste à fournir deux fonctions : un prédicteur  $h : \mathcal{X} \rightarrow \mathcal{Y}$  et un réjecteur  $r : \mathcal{X} \rightarrow \{0, 1\}$ , où  $\mathcal{X}$  et  $\mathcal{Y}$  sont, respectivement, les espaces d’entrée et de sortie de la tâche. La fonction apprise est alors

$$(h, r)(x) = \begin{cases} h(x) & \text{si } r(x) = 0, \\ \textcircled{\mathbb{R}} & \text{sinon (} r(x) = 1), \end{cases}$$

où  $\textcircled{\mathbb{R}}$  désigne le refus de répondre.

Parmi les approches classiques du rejet, on peut distinguer deux grandes familles de méthodes. La première considère que  $h$  ne donne pas seulement une prédiction en  $\mathcal{Y}$  mais fournit également une forme d'information de confiance, par exemple dans  $\mathbb{R}$ , qui peut être utilisée pour effectuer un rejet en seuillant sur sa valeur [1]. C'est le cas, par exemple, de la régression logistique et des réseaux de neurones qui donnent une distribution de probabilité catégorielle en sortie, ou des SVMs qui fournissent une distance à la frontière de décision. Nous les qualifierons de méthodes *basées sur la confiance*. L'autre catégorie tente de mesurer les fluctuations du classifieur et favorise le rejet dans les zones de l'espace d'entrée où sa variabilité est la plus élevée. L'hypothèse avancée par ces méthodes est que le prédicteur est plus susceptible d'être incertain dans sa décision dans ces zones-là. Ces approches ont été particulièrement explorées dans le contexte de l'apprentissage actif [21] et dans les réseaux neuronaux bayésiens [9]. Nous les qualifierons de méthodes *basées sur la dispersion*.

Nous affirmons que, en fait, ces deux approches sont incomplètes dans le sens où elles ne saisissent que partiellement l'information sur l'incertitude. En effet, cette dernière peut être divisée en deux catégories : *l'ambiguïté de la tâche* et *l'incertitude du modèle*. La première est intrinsèque à la tâche que nous voulons accomplir. Par exemple, une image de la fleur d'une plante ne contient qu'une information partielle qui peut ne pas être suffisante pour distinguer deux espèces similaires. Cette incertitude provient du bruit de la mesure (prendre une photo dans notre exemple) ainsi que du bruit dans le processus d'annotation. Elle est, en fait, directement liée à la fonction de régression

$$\eta_k(x) = \Pr [Y = k \mid X = x]$$

de l'apprentissage avec un superviseur imparfait.

Cependant, ce n'est pas la seule incertitude qui apparaît quand on essaie d'apprendre un prédicteur. En effet, lorsqu'on nous donne un ensemble de données, nous devons faire face à l'incertitude du processus d'échantillonnage : nous n'avons accès qu'à une vue partielle de la distribution latente des données. Un échantillonnage différent des données nous induira à choisir un autre prédicteur. De plus, l'algorithme d'apprentissage peut être intrinsèquement stochastique. Ainsi, lors de l'apprentissage de réseaux neuronaux, en raison de la non-convexité de l'objectif d'apprentissage, la ré-exécution de la descente de gradient stochastique en utilisant une initialisation différente et un brassage différent des données d'apprentissage nous donnera une solution différente qui est souvent comparable en termes de précision tout en apportant de la diversité. Cette incertitude n'est pas intrinsèque à la tâche et survient lors de l'apprentissage du modèle, nous l'appellerons donc *incertitude du modèle*.

Ces deux types d'incertitude, l'ambiguïté des tâches et l'incertitude du modèle, saisissent des informations différentes et complémentaires. Nous proposons de les utiliser explicitement pour construire de nouveaux critères de rejet en

essayant de les démêler avant de les fusionner de nouveau de manière adaptée.

Dans cet article, nous nous concentrerons principalement sur la classification binaire à l'aide de réseaux neuronaux. Nous présentons deux contributions principales. Premièrement, dans la Section 3, nous proposons de nouveaux critères de classification avec option de rejet en utilisant des mesures de l'ambiguïté des tâches et de l'incertitude du modèle. Deuxièmement, comme ces critères semblent avoir des performances similaires à celles des méthodes de référence dans nos expériences, dans la Section 4, nous montrons qu'il existe effectivement des informations supplémentaires à exploiter en démêlant les deux types d'incertitude présentés précédemment mais qu'elles peuvent être difficile à saisir sans apprendre un critère ad hoc sur un ensemble de validation.

## 2 Formulation du problème et état de l'art

La classification avec option de rejet a d'abord été introduite et étudiée dans [4, 5] en utilisant des modèles probabilistes. Des travaux plus récents ont étudié ce problème dans le cadre de la théorie de l'apprentissage statistique en définissant une fonction de risque adaptée. Il est habituellement exprimé pour un coût de rejet donné de  $\lambda$  [14, 6] par

$$R_\lambda(h, r) = \mathbb{E}_{X, Y} [\delta_{h(X) \neq Y} (1 - r(X)) + \lambda r(X)] \quad (1)$$

et est minimisé pour le classifieur de Bayes optimal ( $h^*, r^*$ ) suivant :

$$h^*(x) = \delta_{\eta(x) \geq \frac{1}{2}} \quad \text{et} \quad r^*(x) = \delta_{|\eta(x) - \frac{1}{2}| < \frac{1}{2} - \lambda} \quad (2)$$

Il s'agit donc d'un compromis entre le taux d'erreur et le taux de rejet pour ce coût  $\lambda$  et favorise le rejet dans les zones où l'ambiguïté de la tâche est plus grande.

Beaucoup de travaux théoriques partent de cette formulation et se concentrent principalement sur le cas binaire. Par exemple, [14] a étudié le taux de convergence des estimateurs plug-in et des minimiseurs du risque empirique. [1] a proposé une *hinge loss* pour les rejeteurs basés sur la confiance qui peut être mise en œuvre dans la pratique et en propose l'étude théorique. [25] étudie d'autres fonctions objectives convexes tandis que [6] étend cela à des rejeteurs d'une classe de fonctions différente de celle du prédicteur. Une autre approche consiste à apprendre séquentiellement des classifieurs qui peuvent rejeter avec un coût, une tâche appelée apprentissage séquentiel, comme étudié dans [22].

Parallèlement, [8] a utilisé une stratégie basée sur les désaccords pour apprendre un prédicteur parfait dans le cas réalisable, c'est-à-dire en l'absence de bruit et lorsque le classifieur parfait est dans la classe des hypothèses. [23, 24] ont ensuite étudié de telles approches pour le cas agnostique afin d'apprendre ce qu'ils appellent des classifieurs sélectifs point par point, c'est-à-dire que les prédicteurs ont le même taux d'erreur que celui optimal tout en essayant de



maintenir le taux de rejet le plus bas possible. De plus, ils proposent de mesurer la performance de ces modèles, non seulement pour un coût de rejet  $\lambda$  donné, qui pourrait ne pas être quantifiable en pratique, mais sur l'ensemble de la courbe de risque/couverture, ce qui revient à comparer la performance pour tous les choix de  $\lambda$  en même temps [8]. Cependant, ces travaux sont, soit purement théoriques, soit reposent sur une optimisation convexe. Cela peut être un problème pour les réseaux de neurones parce que leur processus d'apprentissage n'est pas encore entièrement compris. Ainsi, de nombreuses études empiriques et heuristiques ont été réalisées autour des critères de rejet et de l'estimation de l'incertitude pour ces modèles. Pour les réseaux neuronaux, les approches basées sur la confiance ont été étudiées depuis longtemps et continuent de l'être jusqu'à récemment dans [17, 16] par exemple. En ce qui concerne les méthodes fondées sur la dispersion, [9] interprète le *dropout* comme une approximation de l'inférence bayésienne et propose d'utiliser la variance du résultat comme mesure d'incertitude. Cependant, [10] montre que ces différentes mesures, basées sur la confiance ou sur la dispersion, peuvent être améliorées en choisissant, pendant la phase d'apprentissage, des modèles mieux adaptés à chaque zone de l'espace d'entrée. Plus largement, dans le contexte de l'apprentissage actif, des mesures basées sur la confiance et sur la dispersion ont été étudiées, [21] en propose une vue d'ensemble.

Les mesures de dispersions que nous introduisons dans ce papier sont basées sur le désaccord entre les modèles d'un ensemble. Cette notion de désaccord est importante dans l'apprentissage et a été exploitée à multiples reprises. En effet, elle a d'abord été utilisée pour obtenir un modèle plus performant que chacun des modèles individuels à travers le *bagging* [3] et le *boosting* [20]. Certains travaux tels que [11] fournissent une étude théorique de ces approches. La notion de désaccord est également importante dans d'autres paradigmes d'apprentissage tels que l'apprentissage actif [13] ou bien, comme développé précédemment, la classification avec rejet [24].

Enfin, les différentes sources d'incertitudes sont généralement séparées en incertitudes *aléatoires* et *épistémiques* [7, 15]. La première est définie comme la partie irréductible de l'incertitude, par opposition à la seconde qui peut être réduite en collectant plus de données ou en affinant le modèle. Cependant, comme cela est expliqué en détail dans [7], la distinction précédente a un sens dans un modèle pour lequel il est explicite quelle incertitude peut être réduite. Nous préférons la terminologie de *l'ambiguïté de la tâche* et *l'incertitude du modèle* parce que nous ne considérons pas ici quelle incertitude peut être réduite et comment mais plutôt si elle est intrinsèque à la tâche ou si elle provient du processus d'apprentissage.

### 3 Deux types d'incertitude

Dans cette section, nous étudions comment, à la fois, l'ambiguïté des tâches et l'incertitude du modèle peuvent être

utilisées pour obtenir de nouveaux critères de classification avec option de rejet.

#### 3.1 Incertitude dans le choix du prédicteur

En raison de la stochasticité de l'apprentissage et des processus d'échantillonnage des données, il y a une incertitude dans le choix de l'hypothèse  $h$ . Nous pouvons construire un critère basé sur la variabilité de la prédiction due à cette stochasticité.

Comme nous utilisons des classes d'hypothèses paramétrées, l'incertitude dans le choix de  $h = h_\theta$  provient en fait d'une incertitude dans le choix des paramètres  $\theta$  étant donné les données d'apprentissage  $\mathcal{D}_{\text{train}}$ , c'est-à-dire  $\Pr[\theta = \theta' | \mathcal{D}_{\text{train}}]$ . Ainsi, une première approche consiste à modéliser complètement cette distribution de probabilité sur les paramètres et une façon de le faire est d'utiliser des approches bayésiennes [18]. Dans ce cas, une distribution *a priori* sur tous les paramètres,  $\Pr[\theta = \theta']$ , est définie et sa la distribution *a posteriori*  $\Pr[\theta = \theta' | \mathcal{D}_{\text{train}}]$  est calculée en appliquant la règle de Bayes.

Cependant, selon le modèle utilisé, il peut être difficile d'établir une distribution *a priori* adaptée, de plus, le calcul de la distribution *a posteriori* peut s'avérer difficile. Il est possible d'utiliser des méthodes d'approximation mais, dans certaines situations, cela peut ne pas être suffisant. Cela est en particulier le cas pour les réseaux de neurones où l'application des méthodes bayésiennes est un domaine de recherche actif [18, 19, 2, 9].

Néanmoins, comme dans le présent article, nous nous intéressons uniquement à la classification avec option de rejet, il n'est pas nécessaire de modéliser cette distribution complexe. Nous pouvons utiliser une approche plus directe pour estimer nos critères en supposant que nous pouvons échantillonner les modèles de notre distribution  $\Pr[\theta = \theta' | \mathcal{D}_{\text{train}}]$ . On peut y parvenir en simulant les différentes sources d'incertitude du modèle, c'est-à-dire :

- la stochasticité de l'algorithme d'apprentissage en exécutant plusieurs fois l'apprentissage sur les mêmes données
- l'échantillonnage des données d'entraînement en utilisant des techniques de *bootstrap* et de sous-échantillonnage pour simuler la stochasticité du processus de génération des données.

Une fois qu'il est supposé que nous pouvons échantillonner à partir de  $\Pr[\theta = \theta' | \mathcal{D}_{\text{train}}]$ , nous pouvons alors construire des critères basés sur la dispersion. En particulier, nous étudions dans la section suivante un critère basé sur le désaccord entre les prédicteurs de cette distribution.

#### 3.2 Un critère pratique de désaccord

Les critères de désaccord ont été théoriquement étudiés dans le contexte de la classification avec option de rejet pour trouver un prédicteur ayant le même taux d'erreur que le classifieur parfait tout en ayant un taux de rejet faible [8, 23, 24], ainsi que dans le cadre de l'apprentissage actif [13]. Les stratégies proposés restent cependant théoriques et

ne sont pas implémentables en pratique. Dans cette sous-section, nous proposons un moyen de les rendre pratiques et de les généraliser.

L'idée générale est de quantifier le désaccord entre les modèles  $(h_\theta)_\theta$  d'un ensemble de modèles en regardant uniquement leurs prédictions  $\hat{y} = h_\theta(x) \in \mathcal{Y}$ . Pour cela, nous définissons la *mesure de désaccord*  $\zeta(x)$  comme

$$\zeta(x) = \Pr[\hat{Y} = 1 | X = x] = \mathbb{E}_\theta [\delta_{h_\theta(x)=1}],$$

étant donné une mesure de probabilité sur l'ensemble des paramètres  $\theta$ . Notez que cette quantité est différente de la fonction de régression  $\eta(x)$  qui mesure l'ambiguïté de la tâche elle-même,  $\Pr[Y = 1 | X = x]$ , indépendamment des modèles. De plus, cette quantité n'est pas à proprement parler une probabilité de désaccord, mais elle capture cette information. En effet, le désaccord est maximal quand  $\zeta(x)$  vaut 0.5 et minimal quand il vaut 0 ou 1.

Si nous modélisons complètement l'incertitude dans les paramètres  $\Pr[\theta = \theta' | \mathcal{D}_{\text{train}}]$ , il est alors possible de calculer le désaccord entre ces hypothèses en utilisant

$$\hat{\zeta}(x) = \int_{\theta'} \delta_{h_{\theta'}(x)=1} \Pr[\theta = \theta' | \mathcal{D}_{\text{train}}] d\theta'. \quad (3)$$

Cependant, si nous ne nous intéressons qu'à  $\zeta(x)$ , nous n'avons, en général, pas besoin de modéliser complètement la distribution sur les paramètres, ce qui peut être une tâche complexe comme le montre la sous-section précédente.

Au lieu de cela, parce que  $\hat{y} \in \mathcal{Y}$  est discret, si on dispose de  $C$  échantillons,  $\theta_1, \dots, \theta_C$ , nous pouvons estimer directement la distribution  $\Pr[\hat{Y} = 1 | X = x]$  avec une approche fréquentiste :

$$\hat{\zeta}_{\text{freq}}(x) = \frac{1}{C} \sum_{i=1}^C \delta_{h_{\theta_i}(x)=1}. \quad (4)$$

Néanmoins, bien que cet estimateur soit non biaisé, il pourrait nécessiter beaucoup d'échantillons, soit un grand  $C$ , pour réduire sa variance.

En général, la plupart des modèles  $h_\theta$  peuvent être décomposés en utilisant une fonction paramétrée  $f_\theta : \mathcal{X} \rightarrow \mathbb{R}$  sur lequel est appliquée une fonction de décision  $\delta_Z : \mathbb{R} \rightarrow \{0, 1\}$ . Typiquement,  $\delta_Z(z) = \delta_{z \geq 0}$ . Lors de l'utilisation d'une fonction de coût logistique,  $z$  est appelé logit et nous pouvons même décomposer  $h$  un peu plus en appliquant, par-dessus  $f_\theta$ , la fonction sigmoïde  $\sigma : \mathbb{R} \rightarrow [0, 1]$  pour transformer  $z$  en probabilité avant de prendre la décision qui devient  $\delta_P(p) = \delta_{p \geq \frac{1}{2}}$ . Ces décompositions peuvent être résumées comme suit :

$$x \xrightarrow{h_\theta} \hat{y} \Leftrightarrow x \xrightarrow{f_\theta} z \xrightarrow{\delta_Z} \hat{y} \Leftrightarrow x \xrightarrow{f_\theta} z \xrightarrow{\sigma} p \xrightarrow{\delta_P} \hat{y}$$

Sur la base de la décomposition précédente, nous pouvons en fait utiliser une approche intermédiaire entre les deux extrêmes que sont les Équations (3) et (4). En effet, parce que  $\hat{y}$  est une fonction (non paramétrique) de  $z$  et  $p$ , il suffit de modéliser l'incertitude dans l'espace des logits ou de la probabilité. A partir de ces distributions, nous pouvons

ensuite dériver la mesure de désaccord en utilisant

$$\hat{\zeta}(x) = \int_z \delta_{z \geq 0} \Pr[Z = z | X = x] dz = 1 - F_x^Z(0) \quad (5)$$

et

$$\hat{\zeta}(x) = \int_p \delta_{p \geq \frac{1}{2}} \Pr[P = p | X = x] dp = 1 - F_x^P\left(\frac{1}{2}\right) \quad (6)$$

où  $F_x^Z$  et  $F_x^P$  sont respectivement la fonction de répartition de la probabilité conditionnelle dans l'espace des logits et des probabilités.

Cette méthode permet d'introduire des hypothèses et des connaissances a priori qui sont plus faciles à tester dans la pratique que la modélisation complète de la distribution des paramètres. En même temps, si ces hypothèses sont vérifiées, l'estimateur résultant convergerait plus rapidement que  $\hat{\zeta}_{\text{freq}}(x)$ , ce qui signifie que nous pourrions choisir un  $C$  plus petit, rendant cette approche plus pratique. Finalement, cette formalisation nous permet de mieux comprendre comment fusionner les différentes statistiques et moments, comme le montre la section suivante.

### 3.3 Choix de la distribution

Nous étudions maintenant quels sont les choix de distributions appropriées dans les espaces des logits et des probabilités.

En supposant qu'on nous donne plusieurs logits,  $z_1, \dots, z_C$ , pour une entrée  $x$ , un choix courant est d'utiliser une loi normale pour modéliser leur distribution. Dans ce cas, nous pouvons ajuster les paramètres de la distribution en utilisant les estimateurs de maximum de vraisemblance habituels  $\hat{\mu}_z = \frac{1}{C} \sum_{i=1}^C z_i$  et  $\hat{\sigma}_z^2 = \frac{1}{C} \sum_{i=1}^C (z_i - \hat{\mu}_z)^2$ . L'estimateur de l'Équation (5) devient alors

$$\hat{\zeta}_{\text{norm}}(x) = 1 - \phi\left(-\frac{\hat{\mu}_z}{\hat{\sigma}_z}\right),$$

où  $\phi$  est la fonction de répartition de la distribution normale standard. Fait intéressant, ce critère est une fonction bijective de  $\frac{\hat{\mu}_z}{\hat{\sigma}_z}$ .

Alternativement, si nous considérons l'espace des distributions binaires  $p$ , un choix naturel est la loi bêta. Dans ce cas, il n'existe pas de forme close pour les estimateurs du maximum de vraisemblance de ses paramètres  $\alpha$  et  $\beta$ . Il faut soit s'appuyer sur un algorithme itératif, soit utiliser la méthode des moments. Dans ce dernier cas, les estimateurs de la méthode des moments sont égaux à  $\hat{\alpha} = \hat{\mu}_p \left( \frac{\hat{\mu}_p(1-\hat{\mu}_p)}{\hat{v}_p} - 1 \right)$  et  $\hat{\beta} = (1 - \hat{\mu}_p) \left( \frac{\hat{\mu}_p(1-\hat{\mu}_p)}{\hat{v}_p} - 1 \right)$  avec  $\hat{\mu}_p = \frac{1}{C} \sum_{i=1}^C p_i$  et  $\hat{v}_p = \frac{1}{C} \sum_{i=1}^C (p_i - \hat{\mu}_p)^2$ . L'estimateur de l'Équation (6) est alors égal à

$$\hat{\zeta}_{\text{beta}}(x) = 1 - I_{\frac{1}{2}}(\hat{\alpha}, \hat{\beta}),$$

où  $I_x$  est la fonction bêta incomplète régularisée.

Maintenant que nous disposons d'une mesure pratique de désaccord, nous pouvons établir un critère de classification avec une option de rejet telle que

$$C_{\text{disagree}}(x) = \max(\hat{\zeta}(x), 1 - \hat{\zeta}(x)). \quad (7)$$

Cependant, il ne s'agit là que d'une mesure de la dispersion des prédictions de l'ensemble, l'incorporation d'une mesure de l'ambiguïté de la tâche pourrait conduire à de meilleurs critères.

### 3.4 Critère de fusion

Si l'on examine la fonction de risque de la classification avec option de rejet de l'Équation (1), cette quantité est minimisée par le réjecteur optimal de Bayes de l'Équation (2). En estimant la fonction de régression  $\eta(x)$ , nous pouvons construire une règle *plug-in* qui nous permettrait d'effectuer un rejet basé sur notre estimateur  $\hat{\eta}(x)$ . Le taux de convergence théorique de cet estimateur *plug-in* a été étudié dans [14] où il a été démontré qu'il dépend de la qualité de l'estimateur  $\hat{\eta}(x)$  et de la structure et du niveau de l'ambiguïté  $\eta(x)$ . La construction d'un tel estimateur peut se faire en optimisant une fonction de coût *strictly proper* [12]. Il s'avère que la fonction de coût logistique est en fait *strictly proper*, une telle approche a par ailleurs été appliquée aux réseaux neuronaux dans [16].

Cependant, cette méthode ne tient pas compte de la variabilité de l'estimateur qui change en fonction des zones de l'espace d'entrée. Cette variabilité pourrait donner lieu à une prédiction différente  $\hat{y}$ . En utilisant le critère de désaccord de la sous-section précédente, nous pouvons construire un nouveau critère en marginalisant notre incertitude dans le choix de la prédiction :

$$c_{\text{fusion}}(x) = \Pr \left[ \hat{Y} = 0 \mid X = x \right] (1 - \hat{\eta}(x)) + \Pr \left[ \hat{Y} = 1 \mid X = x \right] \hat{\eta}(x).$$

S'il n'y a pas de désaccord, c'est-à-dire si  $\zeta(x) \in \{0, 1\}$ , ce critère est simplement l'ambiguïté estimée de la prévision :  $\max(\hat{\eta}(x), 1 - \hat{\eta}(x))$ . De plus, pour une valeur égale de  $\hat{\eta}(x)$ , ce critère rejetera d'abord les zones de l'espace d'entrée où le désaccord est le plus fort. Cette quantité a donc des propriétés intéressantes.

En injectant notre estimateur de  $\zeta(x)$  dans l'équation précédente, le critère devient

$$c_{\text{fusion}}(x) = (1 - \hat{\zeta}(x))(1 - \hat{\eta}(x)) + \hat{\zeta}(x)\hat{\eta}(x). \quad (8)$$

### 3.5 Expériences synthétiques

Afin de comparer les différents critères de rejet, nous utilisons des jeux de données synthétiques où nous pouvons contrôler la quantité d'ambiguïté  $\eta(x)$  et la densité des données  $p(x)$ . Nous utilisons trois jeux de données différents illustrés dans la Figure 1 :

- Fig.1a : absence d'ambiguïté, i.e.  $\eta(x) \in \{0, 1\}$ , mais il existe deux zones de densités uniformes différentes ;
- Fig.1b :  $x$  est distribué uniformément mais il y a des zones d'ambiguïté de quantité différente ;
- Fig.1c : mélange des deux scénarios précédents.

Ces jeux de données permettent de comprendre finement comment les prédictions du modèle et les critères de rejet se comportent sous différentes contraintes d'incertitude.

La mesure de performance que nous utilisons est la courbe RC [8] et, en particulier, l'aire sous cette courbe telle que définie dans [10] que nous désignons RC-AUC. La courbe de risque/couverture (RC) est définie comme suit

$$R(h, r) = \mathbb{E}_{X, Y} \left[ \delta_{h(X) \neq Y} \frac{1 - r(X)}{\phi(h, r)} \right],$$

$$\phi(h, r) = 1 - \mathbb{E}_X [r(X)].$$

Le risque  $R$  quantifie le taux d'erreur des échantillons qui ne sont pas rejetés. La couverture  $\phi$  mesure le taux d'acceptation. Plus cette quantité est faible, meilleur est le compromis entre précision et taux de rejet.

Toutes les fonctions de rejet que nous considérons effectuent un seuillage se basant sur un critère  $c$  :

$$r(x) = \delta_{c(x) \geq \tau}. \quad (9)$$

Nous utilisons comme références les critères suivants :

- pour les approches basées sur la confiance : moyenne de  $|\mu_p - \frac{1}{2}|$  dans l'espace de probabilité, notée "prob. mean", et moyenne de  $|\mu_z|$  dans l'espace logit, notée "logit mean"<sup>1</sup>
- pour les approches basées sur la dispersion : la variance de  $|\mu_p - \frac{1}{2}|$  notée "var. prob." et de  $|\mu_z|$  notée "var. logit" et la moyenne de la divergence de Kullback-Leibler (KL) des prévisions des modèles individuels de l'ensemble comparé à la prévision moyenne telle que définie dans [21]

Notez que pour les critères basés sur la dispersion, c'est leur opposé qui est seuillé. Nous intégrons en outre le classifieur optimal de Bayes de l'Équation (2) pour fournir la plus faible valeur de la RC-AUC réalisable.

Nous comparons d'abord les différents critères basés sur les mesures de désaccord des Équations (7) et (8) dans la colonne de gauche de la Figure 2. Avec suffisamment de données, les critères de désaccord convergent vers le critère optimal en l'absence d'ambiguïté alors qu'en présence d'ambiguïté, ce n'est pas le cas. Cela est attendu car ces critères ne prennent pas en compte cette partie de l'incertitude, mais seulement le désaccord entre les modèles de l'ensemble. L'utilisation du critère de fusion pour incorporer l'ambiguïté à ces critères les améliore et les fait converger vers la performance du critère optimal. Parce que nous utilisons un ensemble de taille 1000, les approches fréquentistes peuvent être considérées comme les meilleurs estimateurs des critères de désaccord et de fusion. Les critères basés sur la loi bêta sont toujours assez proches et, par conséquent, cette dernière peut être considérée comme un bon choix pour modéliser la distribution des probabilités produites par l'ensemble. A contrario, les critères basés sur la loi normale ont une performance médiocre en l'absence d'ambiguïté tout en étant au même niveau ou légèrement meilleurs autrement. Cette loi ne semble donc pas adaptée pour modéliser la distribution dans l'espace des logits.

1. ici, parce que nous considérons des tâches de classification binaire, les autres critères basés sur ces moyennes, par exemple, l'entropie etc, sont équivalents aux critères que nous utilisons

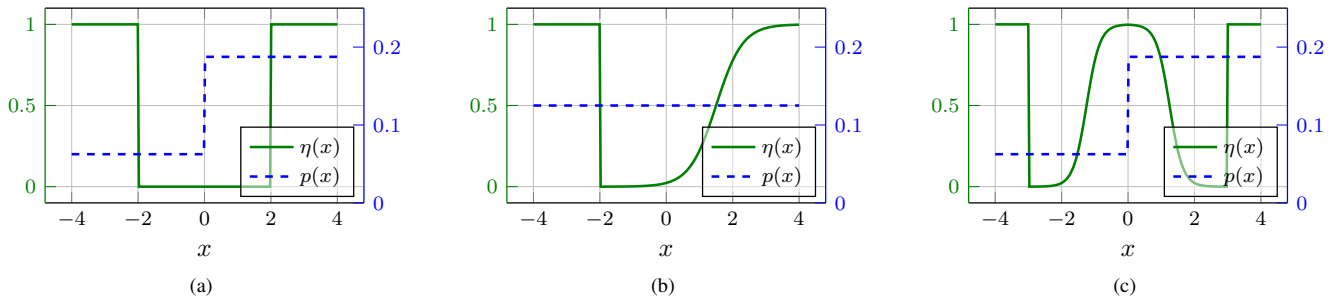


FIGURE 1 – Jeux de données synthétiques avec un degré variable d’ambiguïté et de densité des données. La distribution de densité de probabilité de  $x$  est indiquée en pointillés bleus tandis que la fonction de régression  $\eta(x)$  est indiquée en vert.

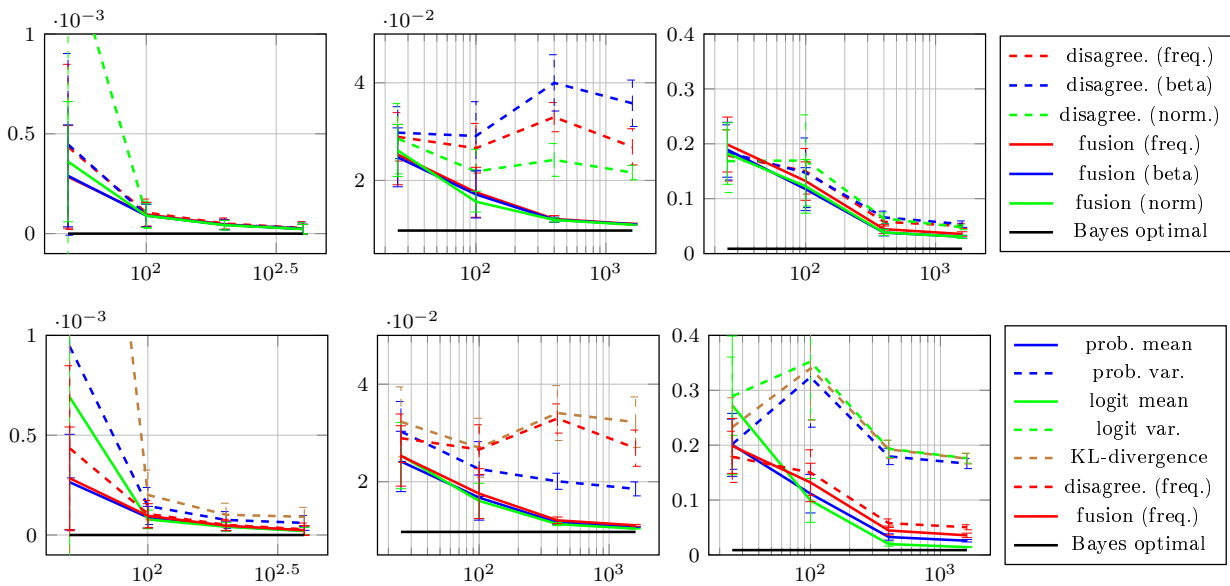


FIGURE 2 – Aire sous la courbe des courbes risk/couverture (RC-AUC) pour comparer les différents critères en fonction de la taille du jeu d’apprentissage. Les colonnes correspondent aux trois jeux de données synthétiques de la Figure 1. La première ligne compare les critères de désaccord et les critères de fusion entre eux. La seconde ligne les compare aux critères de base.

Lorsque ces critères sont comparés à ceux de référence, les critères de fusion donnent des résultats proches de la moyenne de la probabilité et se situent toujours à l’intérieur des fluctuations statistiques comme le montre la colonne de droite de la Figure 2. De plus, il est surprenant de constater que les autres critères basés sur la dispersion peuvent avoir de mauvais résultats et, dans certains cas, ils peuvent même ne pas apparaître sur le graphique.

Deux conclusions principales se dégagent de ces expériences. Premièrement, le bon comportement du critère fondé sur la moyenne de probabilité indique qu’il semble saisir une partie de l’incertitude du modèle, ce qui implique qu’il s’agit d’un estimateur biaisé en ce sens de  $\eta(x)$ . Par exemple, en l’absence d’ambiguïté, il est toujours performant alors que l’ambiguïté satisfait  $\eta(x) \in \{0, 1\}$  et n’apporte donc aucune information sur l’incertitude dans ce

cas. Deuxièmement, le critère de fusion proposé ne semble pas tenir compte de l’incertitude supplémentaire, du moins dans le cadre de ces expériences synthétiques. Toutefois, dans la section suivante, nous analysons plus en détail ces jeux de données contrôlés afin de comprendre s’il existe effectivement des informations supplémentaires qui peuvent être exploitées et comment y parvenir.

## 4 Étude approfondie des rejeteurs

### 4.1 Visualisation des frontières de décision

Les critères précédemment étudiés peuvent être considérés comme une simple décision prise par seuillage comme explicité dans l’Équation (9). La plupart d’entre eux définissent des frontières de décision, soit sur la moyenne et la variance dans l’espace de probabilité,  $(\mu_p, v_p)$ , soit

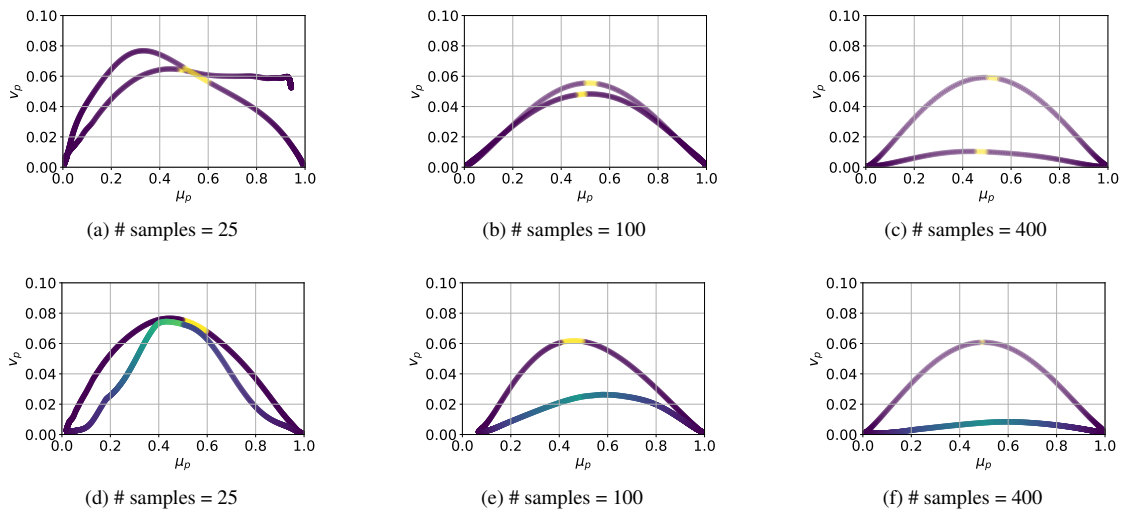


FIGURE 3 – La moyenne et la variance des prédictions dans l’espace de probabilité,  $(\mu_p, v_p)$ , pour chaque point de l’espace d’entrée, c’est-à-dire  $-4 < x < 4$ . La couleur de la courbe indique le taux d’erreur pour ces valeurs, le jaune correspondant à un taux d’erreur élevé et le mauve à un taux faible, le bleu étant une valeur intermédiaire. Les première et deuxième lignes correspondent respectivement au premier et au deuxième jeu de données synthétique.

sur la moyenne et l’écart-type dans l’espace logit,  $(\mu_z, \sigma_z)$ . Afin de visualiser la complexité de la frontière de décision idéale, la Figure 3 montre les valeurs de  $(\mu_p, v_p)$  pour  $x$  variant sur l’espace d’entrée, de  $-4$  à  $4$ , pour les deux premiers jeux de données synthétiques. À chacun de ces points est associé la probabilité d’erreur en couleur.

Ces figures montrent que cette frontière de décision idéale peut en fait être assez complexe et varie en fonction de la taille du jeu d’apprentissage. De plus, la variance de la zone avec moins d’incertitude, c’est-à-dire pour  $x$  autour de  $-2$ , augmente avec le nombre d’échantillons d’apprentissage et cette zone finit par être celle avec la variance la plus élevée. Au premier abord, cela semble plutôt contre-intuitif. En effet, si l’on s’attend à ce que la variance donne directement l’information de l’incertitude du modèle, cette quantité devrait toujours diminuer au fur et à mesure que la taille des données d’entraînement augmente. De plus, il devrait être plus faible dans la zone de non-ambiguïté parce que le modèle devrait la capturer plus rapidement.

Si l’on examine de près la distribution des fonctions de prédiction des modèles dans l’ensemble de la Figure 4, cela est en fait logique. Dans les zones où l’ambiguïté est faible, lorsqu’on dispose de suffisamment de données d’apprentissage, l’incertitude de la fonction de prévision est faible et nous savons presque exactement où se produit le passage d’une classe à une autre. Cependant, comme tous les modèles de l’ensemble ont compris qu’il n’y a pas d’ambiguïté, chacun d’entre eux prédit ou bien 0, ou bien 1, avec une “confiance” élevée, ce qui entraîne une variance élevée de cette valeur de confiance à ce point précis tout en ayant un faible risque d’erreur. A contrario, dans les zones ambiguës, la variance est plus faible car chaque modèle a connaissance de cette ambiguïté mais est incertain sur l’en-

droit exact où la transition de la classe 0 à la classe 1 a lieu. La fonction de décision finale semble plus incertaine mais la variance calculée en un point reste relativement faible comparé à la zone non-ambiguë.

Les Figure 5 et Figure 6 montrent les frontières de décision des critères de rejet utilisés dans la section précédente, respectivement, dans l’espace de probabilité et l’espace logit. Les critères de désaccord et de fusion que nous proposons tiennent compte à la fois de la moyenne et de la variance dans ces espaces et sont donc plus susceptibles de rejeter des zones de forte variance même pour une moyenne constante ou le contraire. Cependant, ces graphiques mettent en lumière les limites des différents critères de base et des critères que nous proposons. En effet, aucun d’entre eux n’a une forme adaptée pour effectuer un rejet efficace sur les jeux de données synthétiques. Cela est attendu, car plus la variance est élevée, plus il est probable que nous rejettions. Mais cette hypothèse n’est pas la bonne ici, ce qui souligne la complexité de l’élaboration d’un critère théorique. Cependant, comme on peut le remarquer dans la Figure 3, les zones de taux d’erreur différents peuvent toujours être séparées dans l’espace de  $(\mu_p, v_p)$ . Nous pouvons donc essayer d’apprendre de manière supervisée à mieux rejeter.

## 4.2 Apprentissage supervisé du réjecteur

Étant donnée que la frontière de décision du réjecteur est une fonction complexe, comme nous l’avons vu dans la sous-section précédente, nous proposons d’essayer de l’apprendre à partir des données en utilisant un jeu de données de calibration. Pour ce faire, nous devons revenir à la fonction de risque de la classification avec option de rejet de l’Équation (1). Dans notre cas,  $h$  est déjà appris et il suf-

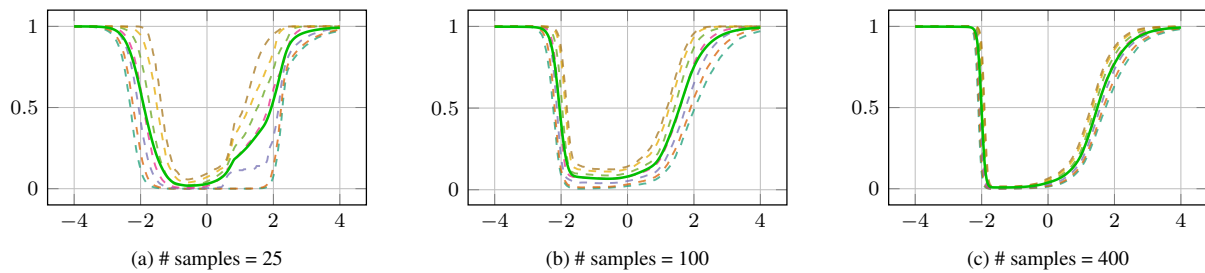


FIGURE 4 – Diagrammes en centiles des distributions binaires prédites par les modèles de l’ensemble pour plusieurs tailles du jeu d’apprentissage. Les centiles représentés en pointillés sont : 5%, 10%, 25%, 50%, 75%, 90% et 95%. La moyenne de ces prédictions est indiquée en vert.

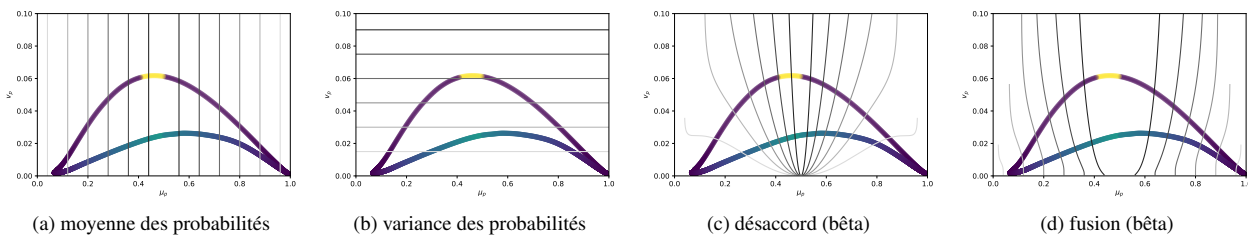


FIGURE 5 – Frontières de décision des différents critères dans l’espace  $(\mu_p, v_p)$  de l’espace de probabilité.

fit donc d’optimiser le risque précédent par rapport à  $r$ . Ce scénario a été étudié dans le contexte de l’apprentissage séquentiel [22]. Cela revient simplement à apprendre une tâche de classification binaire pondérée

$$R_\lambda^{\text{rej}}(h, r) = \mathbb{E}_{X, E} [w_E \delta_{r(X) \neq E}], \quad (10)$$

où  $E$  est la variable aléatoire correspondant à une erreur de prédiction,  $E = \delta_{h(X) \neq Y}$ , et les poids sont égaux à  $w_0 = \lambda$  et  $w_1 = 1 - \lambda$ .

Le réjecteur de Bayes optimal est dans ce cas égal à

$$r^*(x) = \delta_{\eta_E(x) > \lambda} \quad (11)$$

où  $\eta_E(x)$  est la fonction de régression de l’Équation (10), c’est-à-dire  $\eta_E(x) = \Pr [Y \neq h(X) | X = x]$ .

Nous pouvons alors apprendre n’importe quel modèle de classification habituel pour effectuer cette tâche. En faisant varier la valeur de  $\lambda$ , on peut alors reconstituer la courbe RC en réapprenant le réjecteur adapté à chaque fois.

Notre but ici n’est pas de construire le meilleur modèle de rejet mais plutôt de montrer qu’en changeant notre espace d’entrée pour le rejet de  $x \in \mathcal{X}$  à  $(\mu_p, v_p) \in \mathbb{R}^2$ , nous conservons encore suffisamment d’informations discriminantes pour effectuer le rejet. Nous ne prétendons pas qu’il s’agit du meilleur espace pour accomplir cette tâche mais qu’il est suffisant pour apprendre de meilleurs critères de rejet que les critères précédents.

Nous utilisons des classifieurs polynomiaux de degré 3 entraînés avec une fonction de coût logistique. Nous utilisons beaucoup de données d’étalonnage, soit 1000 échantillons, pour apprendre le réjecteur. Ce chiffre n’est pas réaliste dans un scénario réel où nous préférons utiliser ces don-

nées pour améliorer notre prédicteur. Cependant, notre but ici est de comprendre à quel point un dispositif de rejet peut devenir meilleur si nous lui permettons d’être plus complexe et d’être dépendant de la tâche. La Figure 7 montre la courbe RC-AUC du réjecteur par rapport aux critères précédents et aux réjecteurs de Bayes des Équations (11) et (2) sur le deuxième jeu de données synthétique. La fonction de rejet apprise est en effet bien meilleure que les critères précédents. Cela montre la richesse de l’information contenue dans le plan  $(\mu_p, v_p)$ .

## 5 Conclusion et travaux futurs

Les critères classiques, tels que la valeur maximale de la probabilité prédite par un réseau neuronal, fonctionnent bien dans la pratique, cependant ils n’utilisent pas toute l’information d’incertitude disponible. L’analyse des prédictions d’un ensemble met en lumière certains comportements contre-intuitifs qui soulignent notre mauvaise compréhension de l’information d’incertitude capturée par ces modèles. Trouver un bon critère de rejet qui tire parti de toute cette information et qui généralise à différentes tâches est un problème difficile.

Une suite naturelle de ces travaux est d’effectuer des analyses sur des données réelles afin de vérifier que les comportements présentés ici se produisent également dans des jeux de données usuels. De plus, ce travail ouvre la voie à l’élaboration d’un meilleur critère de rejet. Une piste de recherche en ce sens consiste à trouver un espace de représentation plus adapté pour l’apprentissage supervisé des réjecteurs afin de permettre à ces derniers d’être calibrés en

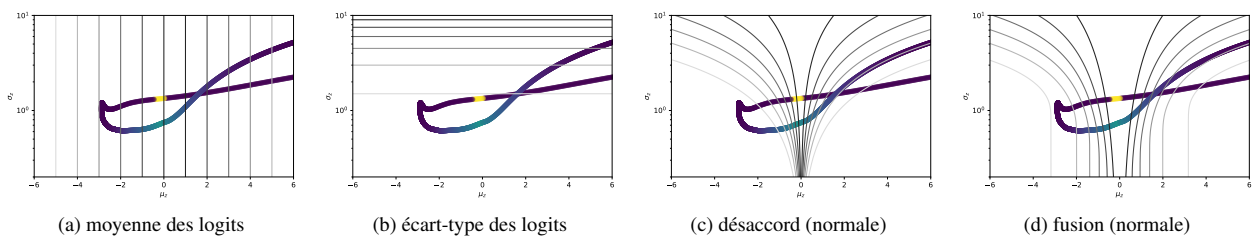


FIGURE 6 – Frontières de décision des différents critères dans l'espace  $(\mu_z, \sigma_z)$  de l'espace des logits.

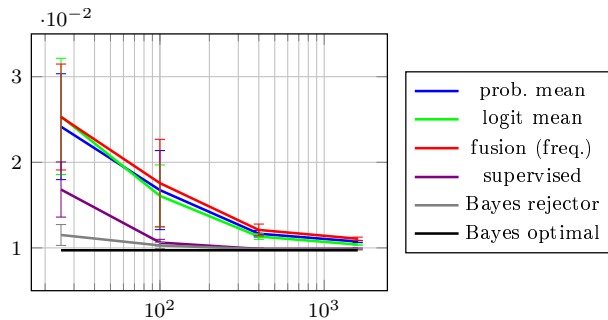


FIGURE 7 – Aire sous la courbe des courbes RC (RC-AUC) comparant le réjecteur supervisé aux différents autres critères avec une taille du jeu d'apprentissage variable sur le deuxième jeu de données synthétique.

utilisant beaucoup moins de données.

## Remerciements

Ce travail a été partiellement financé par le projet ANR WeedElec.

## Références

- [1] P. L. Bartlett and M. H. Wegkamp. Classification with a reject option using a hinge loss. *J. Mach. Learn. Res.*, 2008.
- [2] C. Blundell, J. Cornebise, K. Kavukcuoglu, and D. Wierstra. Weight uncertainty in neural network. In *ICML*, 2015.
- [3] L. Breiman. Bagging predictors. *Mach. Learn.*, 1996.
- [4] C. K. Chow. An optimum character recognition system using decision functions. *IRE T. Elec. Comp.*, 1957.
- [5] C. K. Chow. On optimum recognition error and reject tradeoff. *IEEE Trans. Inf. Theory*, 1970.
- [6] C. Cortes, G. DeSalvo, and M. Mohri. Learning with rejection. In *ALT*, 2016.
- [7] A. Der Kiureghian and O. Ditlevsen. Aleatory or epistemic? Does it matter? *Struct. Saf.*, 2009.
- [8] R. El-Yaniv and Y. Wiener. On the foundations of noise-free selective classification. *J. Mach. Learn. Res.*, 2010.
- [9] Y. Gal and Z. Ghahramani. Dropout as a Bayesian approximation : representing model uncertainty in Deep Learning. In *ICML*, 2016.
- [10] Y. Geifman, G. Uziel, and R. El-Yaniv. Bias-reduced uncertainty estimation for deep neural classifiers. In *ICLR*, 2019.
- [11] P. Germain, A. Lacasse, F. Laviolette, M. Marchand, and J.-F. Roy. Risk bounds for the majority vote : from a PAC-Bayesian analysis to a learning algorithm. *J. Mach. Learn. Res.*, 2015.
- [12] T. Gneiting and A. E. Raftery. Strictly proper scoring rules, prediction, and estimation. *J. Am. Stat. Assoc.*, 2007.
- [13] S. Hanneke et al. Theory of disagreement-based active learning. *Found. and Trends® in Mach. Learn.*, 2014.
- [14] R. Herbei and M. H. Wegkamp. Classification with reject option. *Can. J. Stat.*, 2006.
- [15] A. Kendall and Y. Gal. What uncertainties do we need in Bayesian Deep Learning for Computer Vision? In *NeurIPS*, 2017.
- [16] B. Lakshminarayanan, A. Pritzel, and C. Blundell. Simple and scalable predictive uncertainty estimation using deep ensembles. In *NeurIPS*, 2017.
- [17] A. Mandelbaum and D. Weinshall. Distance-based confidence score for neural network classifiers. *arXiv preprint arXiv :1709.09844*, 2017.
- [18] R. M. Neal. *Bayesian learning for neural networks*, volume 118. Springer Science & Business Media, 2012.
- [19] D. J. Rezende, S. Mohamed, and D. Wierstra. Stochastic backpropagation and approximate inference in deep generative models. In *ICML*, 2014.
- [20] R. E. Schapire and Y. Singer. Improved boosting algorithms using confidence-rated predictions. *Mach. Learn.*, 1999.
- [21] B. Settles. Active learning. *Synthesis Lectures on Artificial Intelligence and Machine Learning*, 2012.
- [22] K. Trapeznikov and V. Saligrama. Supervised sequential classification under budget constraints. In *AISTATS*, 2013.
- [23] Y. Wiener and R. El-Yaniv. Agnostic selective classification. In *NeurIPS*, 2011.
- [24] Y. Wiener and R. El-Yaniv. Agnostic pointwise-competitive selective classification. *J. Artif. Intell. Res.*, 2015.
- [25] M. Yuan and M. Wegkamp. Classification methods with reject option based on convex risk minimization. *J. Mach. Learn. Res.*, 2010.





# La gestion d'actifs augmentée par l'intelligence artificielle

Juliette MATTIOLI<sup>1</sup>

Sarah LAMOUDI<sup>2</sup>

Pierre-Olivier ROBIC<sup>3</sup>

<sup>1</sup> Thales,

<sup>2</sup> Consultant Expert pour le Pôle Finance-Innovation,

<sup>3</sup> Thales Global Services

juliette.mattioli@thalesgroup.com  
sarah.lamoudi@gmail.com  
pierre-olivier.robic@thalesgroup.com

## Résumé

*L'intelligence artificielle (IA) bouleverse tous les secteurs mais surtout le secteur secondaire. En effet, les moyens de productions industriels sont déjà équipés de capteurs qui servent à collecter des informations capitales. Les données vont alors aider à optimiser la production. Les banques, elles-aussi collectent de nombreuses données, comme celles relatives aux cours de bourse, à la santé des entreprises... L'analyse prédictive et prescriptive va permettre de prévoir et de quantifier la valeur des actifs et cela afin de prendre les meilleures décisions possibles. Quelle soit statistique, connexionniste ou symbolique, cet article présente comment l'IA contribue à la gestion d'actifs (GDA) tant dans le monde de la finance que celui de l'industrie.*

## Mots Clef

Gestion d'actifs, apprentissage automatique, décision dans l'incertain, optimisation multicritère, IA connexionniste, IA symbolique.

## Abstract

*Today, Artificial Intelligence (AI) disrupts all sectors and especially industry. Indeed, industrial assets are already equipped with sensors that are used to collect relevant data. Such data will help to optimize production. Finance collected such data for a long time, such as stock market prices, corporate data for enterprise quotation, etc. Predictive and prescriptive analysis will predict and quantify the value of assets in order to take the best possible decision. Whether AI is statistic, connectionist or symbolic, this article underlines how AI contributes to the management of assets both in the domains : finance and industry.*

## Keywords

Asset management, machine learning, decision under uncertainty, multicriteria optimisation, connectionist AI, symbolic AI.

## 1 La gestion d'actifs financiers et industriels

### 1.1 Contexte applicatif

Les avancées en intelligence artificielle (IA) de ces dernières années ont permis la mise en oeuvre de résultats prometteurs dans de nombreux domaines. Il devient possible de prédire, d'anticiper ou d'optimiser des comportements ou des processus métiers, dotant un système de capacité d'apprentissage, de raisonnement et de décision. Ainsi, l'IA peut contribuer à la gestion d'actifs allant de la gestion de portefeuille de biens aux décisions d'investissement tout en prenant en compte la conformité et les risques. Cependant, le terme **gestion d'actifs** (GDA - *Asset management* en anglais) est utilisé de manière très différente suivant le domaine d'usage concerné.

En effet, *en finance*, la GDA consiste à faire prospérer le patrimoine de ses clients via l'investissement, en respectant les obligations réglementaires et contractuelles et en appliquant des stratégies d'investissements, afin de dégager le meilleur rendement possible en fonction du risque choisi [1].

*L'industrie manufacturière* a repris ce terme pour décrire la maîtrise du cycle de vie de ses actifs. Une installation industrielle produit de la valeur grâce à ses performances techniques de disponibilité, de rendement et la qualité de sa production. Il est donc important de les préserver. Les actifs considérés sont alors les biens physiques. Cette approche est connexe aux approches de type *Product Life Cycle Management* (PLM littéralement gestion du cycle de vie des produits) ou *Service Life Cycle Management* (SLM), sur la composante services [2]. La GDA couvre ici les activités réalisées directement sur les actifs au cours de leur cycle de vie, de la conception ou acquisition à leur déclassement ou destruction en passant par leur exploitation, maintenance et renouvellement. Les activités de gestion nécessaires pour diriger, planifier, coordonner, favoriser, faciliter, soutenir et améliorer en continu l'efficacité et l'efficacité des activités effectuées directement sur ou

par les actifs sont aussi à prendre en compte. De plus, un actif physique industriel possède une valeur d'usage (valeur à neuf du bien, corrigée de la vétusté et/ou des maintenances ou réparations réalisés sur ce bien), et une valeur vénale, constituant ainsi un patrimoine de l'entreprise. Enfin, avec le mouvement "usine 4.0", l'usine intelligente est une usine productrice de données, ces dernières étant majoritairement délivrées par les technologies de l'Internet of Things (IoT). Il devient alors possible d'avoir des informations sur l'utilisation des actifs, sur leur degré d'obsolescence, de dégradation, ou sur leurs éventuelles défaillances pour optimiser la gestion des actifs de production en minimisant le coût du cycle de vie constitué des coûts directs et indirects (manque à gagner...) et obtenant le meilleur compromis entre les besoins à court et à long terme. L'IFRAMI (Institut Français d'Asset Management Industriel et Infrastructures) définit la GDA comme un processus impliquant l'équilibre des coûts, la prise en compte des risques et des opportunités ainsi que les avantages liés à une meilleure performance des actifs.

Un système de GDA est donc un système d'aide à la décision permettant d'aligner toutes les activités induites pour générer de la valeur, garantir une meilleure qualité et cohérence dans la prise de décision, définir des priorités dans les mesures à prendre, et permettre une meilleure gestion des risques, des incertitudes et des changements.

Le champ des possibles de l'IA est alors immense et ne cesse de s'étendre. Qu'elle soit symbolique, connexionniste ou statistique, et/ou combinée à la science des données (voir fig. 1), l'IA apporte des solutions à toutes les étapes de la GDA, de la gestion de portefeuille, aux décisions d'investissement en passant par la conformité réglementaire et la gestion des risques.

## 1.2 Structure de l'article

L'objectif de ce papier est de montrer comment les différentes techniques de l'intelligence artificielle (IA) sont déjà opérationnelle pour la GDA tant dans le domaine de la finance que dans l'industrie.

Notons que ce panorama pour le domaine de la finance est une synthèse des travaux menés dans le cadre du livre blanc "*Intelligence Artificielle, Blockchain et Technologies Quantiques au service de la finance de demain*" [3]. Pour le domaine de l'industrie, cet article reflète les actions autour de la transformation numérique de Thales.

Face à la complexité croissante des systèmes (financiers ou industriels) et aux évolutions souvent imprévisibles des environnements dans lesquels ils sont plongés, le décideur a de plus en plus besoin d'outils pour une GDA efficace et efficace. Mais pour prendre de bonnes décisions, il est nécessaire de bien comprendre la situation courante, puis de faire de bonnes prévisions<sup>1</sup>. Dans ce contexte, les al-

1. Par exemple, dans le contexte de la finance, une prévision est un pronostic effectué sur le comportement futur de variables d'intérêt. Par exemple, dans le cas de la finance, ces variables modélisent les rendements du marché boursier et la décision est l'action prise sur les marchés, comme l'action d'investir dans certains titres plutôt que d'autres.

gorithmes d'apprentissage apportent d'excellents résultats, comme le souligne la section 2.

Le caractère optimal de la GDA a amené les chercheurs à proposer une modélisation classique (ou monocritère) des problèmes de décision sous contraintes dont la solution représente le meilleur choix. Mais il paraît illusoire de parler d'optimalité car les critères de décisions devant être pris en compte sont multiples et parfois conflictuels. Ainsi, la GDA étant par essence de nature multicritère [4], l'aide à la décision multicritère fournit le cadre naturel à la résolution de tels problèmes comme le souligne la section 3.

Enfin, compte tenu de la complexité, les approches hybrides (mixant IA des données avec l'IA symbolique) prennent tout leur sens (voir section 4). Par exemple, pour obtenir une estimation de la valeur du portefeuille aussi réaliste que possible, le couplage de la prédiction estimée par apprentissage avec de la connaissance métier est pertinent. D'autres techniques d'IA peuvent apporter une vraie valeur ajoutée. En effet, la lecture et l'interprétation de toutes les réglementations requièrent beaucoup de temps de personnels qualifiés. Il est donc envisageable d'effectuer une lecture automatique par le biais de techniques de traitement automatique de la langue (TAL) autorisant une sélection ciblée des articles pertinents. De plus, couplé aux arbres de décision, le TAL est de plus en plus utilisés en GDA en particulier au travers des chatbots avec l'exemple des robots advisors (voir paragraphe 4.2) très répandus en finance.

## 2 Apprentissage automatique au service de la GDA

En finance, un nom incontestablement associé à la gestion d'actifs, est celui de l'économiste Harry Markowitz qui reçut le prix Nobel d'économie pour ses travaux en 1990. Son idée repose sur la construction d'un portefeuille d'actions qui permet d'obtenir un rendement donné avec un risque moindre comparé à celui de chaque action prise séparément, et cela en répartissant de manière optimisée son argent entre les différents actifs financiers. Ce problème d'ingénierie financière nécessite de déterminer une stratégie d'investissement parmi un ensemble d'actifs afin d'atteindre certains objectifs, tels que la maximisation de la richesse cumulée ou le rendement ajusté du risque à long terme. H. Markowitz propose [5] en 1952 son modèle Moyenne-Variance reposant sur l'assertion que *tout investisseur poursuit deux objectifs contradictoires qui sont la maximisation du rendement espéré et la minimisation du risque mesuré par la variance du rendement*. Sur cette base, de nombreuses variantes ont été développées [6] parmi lesquelles on trouve la sélection de portefeuille, la théorie de l'utilité espérée, les modèles d'équilibre d'actifs financiers ou la théorie du prix d'arbitrage. Chaque action est décrite par son rendement moyen et sa volatilité. Les estimations du rendement moyen et de la volatilité s'appuient en général, sur des approches d'IA statistique mais les arbres de décisions et les réseaux de neurones donnent

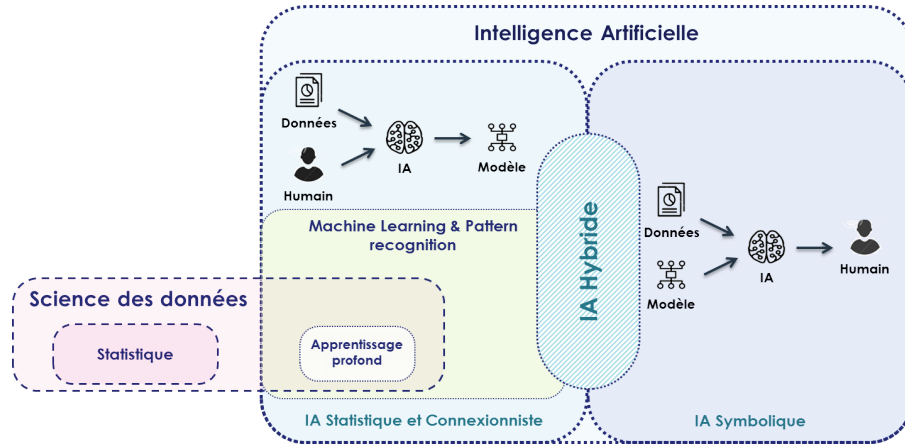


FIGURE 1 – Intelligence artificielle et science des données

aujourd’hui des résultats plus performants [7].

Dès 1997, l’utilisation de réseaux de neurones à base de perceptron multi-couches est proposée dans [8] pour la détection automatique d’entreprises en difficulté, et Y. Bengio [9] applique des techniques d’apprentissage par renforcement sur un problème réaliste d’allocation de 35 actifs. Il obtient d’excellente performance par rapport à un modèle entraîné à minimiser l’erreur de prévision (erreur quadratique). Dans les années 2000, des approches à base de *Pattern-Matching* ont été utilisés pour identifier dans l’historique du comportement du marché, une partie similaire à la situation actuelle pour optimiser le portefeuille [10]. En 2017, dans [11], les auteurs reprennent l’idée de Y. Bengio sur le *deep learning*, et appliquent avec succès des techniques d’apprentissage profond par renforcement (*reinforcement deep learning*) à la gestion de portefeuille, et cela sans aucune connaissance des marchés financiers.

Dans le cadre de la GDA industrielle, comme celle déployée dans Thales, le problème d’évaluation des risques est lié à la défaillance des actifs physiques. Cela consiste alors à détecter et à diagnostiquer des dégradations ou des baisses de performance des moyens de production. L’objectif est de pouvoir prédire (pronostiquer) des défaillances dans le futur, en inférant l’état de santé actuel des actifs dans le temps. Cette fonction *pronostic* se concentre sur l’évolution des défaillances progressives et exploite généralement l’état de santé actuel de l’actif (résultats du diagnostic prédictif) pour évaluer son état de santé futur, au regard du profil d’emploi estimé. Bien qu’il n’existe pas de définition unique de la notion de pronostic dans le domaine de la GDA industrielle, les différentes définitions existantes convergent vers une fonction capable de déterminer la durée de vie restante avant la défaillance, nommée RUL (*Remaining Useful Life*), ou la probabilité d’un système à fonctionner pendant un certain temps. Ce concept est spécifique à la maintenance prédictive [12]. Par nature, le pronostic vise à anticiper l’état d’un système dans le futur et amène des incertitudes sur la valeur du RUL. En effet, le pronostic de l’état futur de l’actif considéré doit non

seulement prendre en compte son état actuel, mais également des données relatives à ce système dans le futur (prévisions d’utilisation, informations sur l’environnement, futures opérations de maintenance exécutées...), qui influent sur son état de dégradation [13]. De nombreux travaux ont été réalisés ces dernières années afin d’améliorer la performance de l’estimation du RUL. [14] présente un état de l’art assez complet des méthodes de calcul du RUL parmi lesquelles figurent les techniques de régression, les Hidden Markov Models et Hidden Semi-Markov Models. [15] utilise une approche de *Multiobjective Deep Belief Network Ensemble* basée sur de l’apprentissage automatique non supervisé. Toutefois, les méthodes les plus prometteuses sont celles qui combinent plusieurs approches, comme le souligne [16]. En effet, le pronostic doit prévoir l’état futur et fournir une durée de vie résiduelle à partir d’une connaissance à priori sur ce système (lois de dégradations, par exemple), d’une connaissance d’informations passées (historique des modes de fonctionnement passés), présentes (état courant) et futures, ce qui nécessite une hybridation de techniques d’IA symboliques avec des méthodes d’apprentissages (voir section 4).

### 3 GDA : problème de décision multicritère sous contraintes en avenir incertain

#### 3.1 Principes généraux de la décision multicritère en avenir incertain

**Aide à la décision.** L’aide à la décision est sans doute l’un des domaines où l’émergence de l’intelligence artificielle (IA) apporte un bouleversement majeur, tant sur la nature de l’aide fournie que sur l’échelle temporelle où cette information est susceptible d’être disponible. Elle peut même conduire à un changement de paradigme. Cela concerne tous les métiers de la finance mais aussi ceux de l’industrie. Mais qu’est-ce qu’une décision ? C’est le fait qu’un acteur (ou un ensemble d’acteurs) effectue un choix, si possible après réflexion, entre plusieurs solutions (*alter-*

*natives*) pour affronter une situation difficile, résoudre un problème délicat ou répondre à une question complexe.

**Décision en avenir incertain.** La décision en avenir incertain prend tout son sens lors d'opérations de gestion d'événements rares ou exceptionnels, et cela en raison des bouleversements de l'environnement dans lequel les organisations évoluent. En effet, on ne peut prévoir ni leur forme (ex. krach boursier, panne subite d'un équipement), ni le moment précis de leur déclenchement. De plus, la capacité de l'organisation à prendre des décisions justes et rapides est un facteur clé de succès vers l'issue la plus favorable. L'incertitude est alors le cadre de cette prise de décision. Il est impossible d'y échapper. Elle renvoie au fait que les conséquences d'une décision ne sont pas connues à l'avance, ce qui expose le décideur à un risque. Cependant, le risque encouru peut être quantifiable quand la probabilité d'occurrence des différentes conséquences possibles est objectivement connue à l'avance. On estime alors la perte encourue en multipliant le montant de la perte éventuelle par la probabilité que se réalise l'éventualité défavorable, en utilisant par exemple le RUL dans le cadre industriel ou la Value at Risk (VaR) mesure probabiliste de la perte financière possible sur un horizon donné. Toutefois, selon Knight [17], l'incertitude devient intrinsèque si les risques ne sont pas objectivement mesurables à l'avance. Ainsi, en finance, la rentabilité des investissements par le crédit bancaire et le rendement futur des titres obligataires ou des actions sont affectés par cette incertitude. Les méthodes sous-jacentes à cette estimation ont été présentées dans la section 2 et sont en général soit basées sur de l'IA statistique ou connexioniste soit sur de l'IA hybride.

**Décision multicritère.** De plus, un problème de décision est souvent multicritère. Il se caractérise par la prise en compte explicite de plusieurs objectifs à optimiser simultanément dans l'analyse des préférences, la comparaison des solutions et la détermination d'une ou des solution(s) optimale(s). Les problèmes induits varient selon la question posée. On peut distinguer les problèmes de choix où l'on cherche à déterminer les meilleures solutions, les problèmes de classement où l'on veut ordonner, au moins partiellement, les solutions selon un ou plusieurs critères et les problèmes de classification où l'on cherche à affecter les solutions à des catégories prédéfinies selon leur valeur intrinsèque. La problématique du choix est celle que l'on rencontre le plus fréquemment. Elle vise à trouver une solution qui optimise au mieux les différents critères, ou un sous-ensemble de solutions, aussi réduit que possible, contenant les meilleures solutions. Formulée ainsi, l'optimisation multicritère est alors un problème mal posé. En effet, de part la nature potentiellement conflictuelle des critères, comme c'est le cas pour la GDA, il n'existe généralement pas de solution optimisant tous les critères simultanément. La plupart des méthodes existantes en décision multicritère se basent sur la somme pondérée, et ce pour des raisons évidentes de simplicité, mais cet opérateur présente des biais qu'il n'est pas possible d'éliminer. Par exemple,

comme le montre [18], la somme pondérée ne permet pas de modéliser des phénomènes de veto ou des phénomènes de compensation entre critères. Le problème de l'agrégation des préférences consiste alors à synthétiser des informations traduisant des aspects ou des points de vues différents, parfois conflictuels, au sujet d'un même ensemble d'objets ou d'actions (performances, utilités, préférences). Il se pose de manière cruciale dans nombre de procédures d'évaluation, de comparaison ou de classification utilisées en aide à la décision multicritère [19]. Que ce soit un problème de choix ou de rangement, la question centrale est toujours un problème de comparaison. Ainsi, dans un problème de choix, l'identification du meilleur candidat nécessite d'être capable de le comparer à tous les autres ; dans un problème de rangement, on espère pouvoir comparer toute paire d'actions de manière à obtenir un classement complet. Un problème d'agrégation multicritère consiste alors à exploiter l'information préférentielle (performances, indices d'importance), pour construire un modèle de comparaison globale des alternatives sous la forme d'une relation de préférence  $\succeq$  sur l'ensemble des dites alternatives. Pour cela, on utilise nécessairement une règle d'agrégation permettant de construire la relation  $\succeq$  à partir des vecteurs de performances.

### 3.2 Décision multicritère appliquée à la GDA

En matière de gestion financière de portefeuille, l'ensemble des actions présentes sur un marché financier est discret, fini et défini en extension. Chaque action est alors représentée dans l'espace des critères. En général, six critères sont utilisés :

- Le *Return on Equity* (ROE), critère de rentabilité des capitaux propres à maximiser ;
- Le *current ratio*, critère de liquidité au sens strict qui doit être maximisé ;
- Le ratio cash flow/dette, critère de solvabilité à maximiser ;
- Le rendement mensuel moyen ;
- Le *Price Earning Ratio* (PER) mensuel qui doit être minimisé ;
- Et le bénéfice par action (EPS, *Earnings Per Share*) annuel, qui doit être maximisé.

Dans le monde de l'industrie, lors de l'élaboration d'un plan de renouvellement des actifs de production, les décideurs sont confrontés à plusieurs questions, parmi lesquelles :

- Quand est-il approprié de remplacer les actifs existants au lieu de continuer à les entretenir ?
- Que faire si le potentiel de production de l'actif est insuffisant ?
- Quels critères devraient être considérés lors du remplacement de l'actif ?

Ces problèmes font partie de la GDA et devraient être traités de façon appropriée lorsque l'on considère la question de renouvellement. Le taux de remplacement d'un

actif peut en effet, varier en raison de plusieurs facteurs dont la maintenance excessive, l'obsolescence, la détérioration physique, la vie de l'actif soumis à l'usure et les défaillances subites.

Dans ces deux cadres (finance et industrie), l'application de techniques d'aide à la décision multicritères permet de définir une fonction d'agrégation des différents KPI (*Key Performance Indicator*) intégrant les différentes dimensions de la GDA pour une optimisation multicritère. Les techniques de planification et de résolution de problèmes combinatoires comme la programmation par contraintes permettent alors de proposer au décideur, une ou des solutions valides.

## 4 IA hybride en GDA

### 4.1 Estimation de la valeur d'un portefeuille d'actifs

Si les modèles analytiques traditionnels basés sur les processus de fabrication du produit ou du service, restent encore largement utilisés en industrie, les modèles statistiques s'imposent progressivement. Ainsi, pour estimer la valeur d'un portefeuille d'actifs (industriels ou financiers) au cours du temps, le recours de techniques faisant appel aux approches stochastiques est devenue une pratique courante [20]. Ces modèles stochastiques doivent cependant être calibrés (par apprentissage par exemple) à partir des données ou mesures observables les plus récentes possibles [21]. Cependant, la complexité des approches sous-jacentes, surtout si elles sont dynamiques et stochastiques, rend difficile une implantation numérique, à moins de simplifier les modèles. Ainsi, en renonçant à un modèle très précis, [22] propose une approche reposant sur l'existence de moyennes ou de tendances et sur des séries temporelles pour l'estimation de la valeur d'un portefeuille d'actifs financiers. Signalons que [23] utilise dès 2002 des approches neuronales pour aborder ce problème, ou plus récemment avec l'apprentissage profond [11]. En 2004, CMU (*Carnegie Mellon University*) propose même *Warren*, un système de gestion intelligente de portefeuille reposant sur des approches multi-agents combinées à de l'analyse textuelle de rapports financiers [24].

Mais avec l'IA hybride couplant apprentissage et connaissances, de nouvelles méthodologies apparaissent, notamment en matière de pricing [20]. On considère alors que le prix d'un actif est corrigé par l'ajout de données complémentaires, comme les agrégats des vues, avis et réflexions délivrées par de nombreux individus (clients, prospects, partenaires, etc.) permettant de comprendre sa dynamique et d'anticiper son évolution future.

### 4.2 Robo-advisors support à la GDA

Depuis plusieurs années, une nouvelle forme de conseil financier a émergé bousculant le monde traditionnel : les *robo-advisors* [25]. Littéralement "robot-conseiller", le robo-advisor est un conseiller en gestion de patrimoine, impliquant un strict minimum d'intervention humaine, les ar-

bitrages étant effectués sur la base d'algorithmes croisant de grandes quantités de données. Plus précisément, il s'agit d'examiner comment segmenter les clients en matière de gestion de patrimoine, en fonction de leurs besoins et de leurs préférences, et développer une offre de conseil partiellement automatisée et personnalisée [26]. Pour ce faire, l'IA peut être utilisée dans le fonctionnement de ses services pour contrecarrer les deux points de blocages majeurs :

- Le coût : dans la finance traditionnelle, moins le client a de ressources, plus le conseil lui revient cher. Grâce à l'apprentissage automatique, les temps de calcul sont réduits et la recommandation du CGP (Conseil en Gestion de Patrimoine) est automatisée, baissant ainsi les coûts ;
- Le choix de produits et la personnalisation : la plupart des acteurs de l'épargne cherchent aujourd'hui à faire rentrer leurs clients dans des cases prédéfinies et n'offrent pas de personnalisation du service : c'est en effet la solution la plus couramment utilisée pour industrialiser la distribution. Grâce à l'aide à la décision multicritère, des systèmes de recommandations sur-mesure adaptées à leurs objectifs, à leurs ressources financières, à leur situation patrimoniale et à leur volonté de prise de risques seront possibles.

De nombreuses banques ont été séduites par les assistants virtuels, comme la banque singapourienne DBS qui a décidé de mettre Watson au service de ses conseillers en gestion du patrimoine. Dans ce cas, une combinaison de techniques d'apprentissage, de raisonnement à base de connaissances (modélisées par le biais d'ontologie) a permis de concevoir des robot-advisors ou des chatbots pertinents. Ainsi un système, reposant sur une hybridation d'algorithmes d'apprentissage comme le deep learning et IA symbolique, permet ainsi de bâtir une recommandation personnalisée. Par ailleurs, la collecte ciblée de données améliorera la pertinence de la recommandation.

## 5 Conclusions

Le secteur de la GDA aussi bien financière qu'industrielle est actuellement en pleine transformation, touché par l'augmentation de l'incertitude, de nouvelles contraintes réglementaires et l'objectif d'augmenter l'efficacité du portefeuille. Ainsi dans un environnement de plus en plus concurrentiel, l'IA devient alors incontournable. La mise en place de méthodes d'analyse de données historiques à base de machine learning permet ainsi d'évaluer les performances des actifs, l'estimation de leur valeur et de détecter des schémas qui se répètent et réallouer le portefeuille en conséquence. Dans ce processus de GDA, la dimension de l'affect humain est aussi extrêmement importante. L'apport de la décision multicritère (voir section 3) couplé à des techniques d'apprentissage est très attendu pour permettre de s'adapter à un environnement changeant où il n'existe pas de règles précises. Deux approches se profilent :

- L'approche *exo-squelette* consiste à utiliser des outils d'aide à la décision, permet d'augmenter la productivité du gérant d'actif par le recours à des outils techniques utilisant l'IA.
- La seconde approche, à plus long terme, inclura l'industrialisation du processus d'investissement et donc de la prise de décision. C'est un changement de paradigme significatif, analogue à la révolution industrielle qui a vu la disparition de la fabrication artisanale au profit d'unités de production où les tâches rationalisées et optimisées sont effectuées par des machines contrôlées par l'homme.

Pour y parvenir, il faut

- Travailler sur la qualité des données : les données considérées comme utiles à la prise de décision sont celles sur les actifs, son contexte d'emploi, et le sentiment des décideurs à l'égard de la valeur de l'actif. L'utilisation de données externes factuelles alternatives est intéressante et peut donner un avantage concurrentiel, mais elles sont sujettes à interprétation positive ou négative selon le contexte (par exemple : des stocks importants peuvent vouloir dire qu'il y a une surproduction et donc des invendus - signal négatif ou que l'entreprise a engrangé beaucoup de commandes - signal positif). Le lien entre l'information et la décision à prendre n'est pas linéaire et l'enjeu est donc de sélectionner attentivement l'information pertinente. Avant toute mesure de mise en oeuvre, il est important de définir les objectifs et usages des données à traiter en fonction des enjeux selon le principe du "fitness for use" : dans certains cas, une tolérance à l'erreur est acceptable (comme des systèmes de recommandation d'achat en B2C) alors que dans d'autres (comme les systèmes critiques ayant un impact juridique, médical, financier...), l'ensemble du système d'information devra être traité avec la plus grande rigueur. Il s'agit donc d'une problématique liée au domaine métier et à l'usage :
  - en amont, il faut que les producteurs de données deviennent des sources qualifiées, voire certifiées. On parle alors de cotation des sources ;
  - en aval, il s'agit de fournir aux utilisateurs tous les éléments utiles afin qu'ils puissent en faire un usage pertinent et durable.
- Etre sélectif sur les techniques d'IA : comme nous l'avons vu, il n'existe pas de méthode unique pour adresser le problème de la GDA et il faut donc trouver la meilleure hybridation. Cela implique de décomposer le processus en tâches simples et d'appliquer la technique la plus efficace. Pour cela, une logique holistique et systémique est extrêmement importante ;
- Et incorporer l'humain : l'estimation de la valeur des actifs résulte de l'agrégation de plusieurs points de vue et donc de l'affect des différentes parties

prenantes intégrant parfois des facteurs à la fois rationnels et émotionnels. L'ambition est de comprendre et d'anticiper les motivations de décisions basés sur des perceptions ou des sentiments.

Enfin, pour pouvoir passer de preuves de concept à une réelle mise en production, il est nécessaire de travailler sur l'ingénierie de l'IA pour garantir des propriétés tels que la robustesse... En particulier, la validation, la vérification et l'évaluation doivent être traitées du point de vue système et pas seulement algorithmique (cf. fig. 2).

## Remerciements

Ces réflexions ont été nourries par de nombreuses discussions avec Bertrand Braunschweig (directeur de la mission Inria de coordination du programme national de recherche en intelligence artificielle) que les auteurs remercient chaleureusement. En effet, pendant les réunions de travail du Livre Blanc sur *Intelligence Artificielle, Blockchain et Technologies Quantiques au service de la finance de demain* [3], l'analyse de retours d'expériences issus d'autres industries que la finance a permis d'identifier de nombreuses synergies, en particulier celui de la gestion d'actifs.

## Références

- [1] Robert A Haugen and Robert A Haugen. *Modern investment theory*, volume 5. Prentice Hall Upper Saddle River, NJ, 2001.
- [2] Hasan Burak Cavka, Sheryl Staub-French, and Erik A Poirier. Developing owner information requirements for bim-enabled project delivery and asset management. *Automation in construction*, 83 :169–183, 2017.
- [3] Pôle Finance Innovation, editor. *Intelligence Artificielle, Blockchain et Technologies Quantiques au service de la finance de demain*, Livre blanc, 2019.
- [4] Christian Hurson and Constantin Zopounidis. *Gestion de portefeuille et analyse multicritère*. FeniXX, 1996.
- [5] Harry Markowitz. Portfolio selection. *The journal of finance*, 7(1) :77–91, 1952.
- [6] Mark Rubinstein. Markowitz's portfolio selection : A fifty-year retrospective. *The Journal of finance*, 57(3) :1041–1045, 2002.
- [7] Bo Wahlberg, Stephen Boyd, Mariette Annergren, and Yang Wang. An admm algorithm for a class of total variation regularized estimation problems. *IFAC Proceedings Volumes*, 45(16) :83–88, 2012.
- [8] Philippe Paquet et al. L'utilisation des réseaux de neurones artificiels en finance. *Document de recherche*, 1, 1997.
- [9] Yoshua Bengio. Training a neural network with a financial criterion rather than a prediction criterion. In *Decision Technologies for Financial Engineering : Proceedings of the Fourth International*

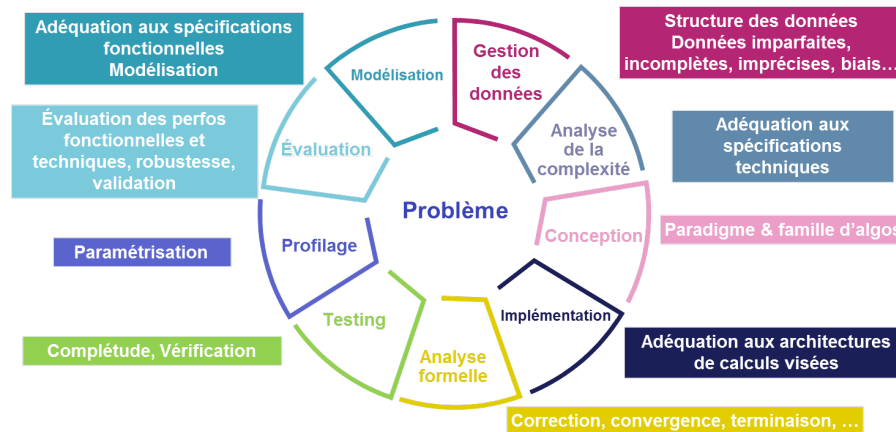


FIGURE 2 – Vers l'ingénierie algorithmique de l'IA

- Conference on Neural Networks in the Capital Markets (NNCM'96)*, World Scientific Publishing, pages 36–48, 1997.
- [10] László Györfi, Gábor Lugosi, and Frederic Udina. Nonparametric kernel-based sequential investment strategies. *Mathematical Finance : An International Journal of Mathematics, Statistics and Financial Economics*, 16(2) :337–357, 2006.
- [11] Zhengyao Jiang and Jinjun Liang. Cryptocurrency portfolio management with deep reinforcement learning. In *2017 Intelligent Systems Conference (IntelliSys)*, pages 905–913. IEEE, 2017.
- [12] 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.
- [13] 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.
- [14] Xiao-Sheng Si, Wenbin Wang, Chang-Hua Hu, and Dong-Hua Zhou. Remaining useful life estimation—a review on the statistical data driven approaches. *European journal of operational research*, 213(1) :1–14, 2011.
- [15] Chong Zhang, Pin Lim, AK Qin, and Kay Chen Tan. Multiobjective deep belief networks ensemble for remaining useful life estimation in prognostics. *IEEE transactions on neural networks and learning systems*, 28(10) :2306–2318, 2017.
- [16] Juliette Mattioli, Pierre-Olivier Robic, and Thomas Reydellet. L'intelligence artificielle au service de la maintenance prévisionnelle. In *4ème conférence sur les Applications Pratiques de l'Intelligence Artificielle APIA2018*, 2018.
- [17] Frank H. Knight. Risk, uncertainty and profit, 1921. *Library of Economics and Liberty*, 1971.
- [18] Michel Grabisch and Christophe Labreuche. A decade of application of the choquet and sugeno integrals in multi-criteria decision aid. *4OR*, 6(1) :1–44, 2008.
- [19] Jacky Montmain and Christophe Labreuche. Amélioration multicritère d'options dans les systèmes complexes. *LFA'2009, Rencontres Francophones sur la Logique Floue et ses Applications*, 2009.
- [20] Arash Bahrammirzaee. A comparative survey of artificial intelligence applications in finance : artificial neural networks, expert system and hybrid intelligent systems. *Neural Computing and Applications*, 19(8) :1165–1195, 2010.
- [21] Michel Fliess, Cedric Join, and Frédéric Hatt. A-t-on vraiment besoin d'un modèle probabiliste en ingénierie financière ? In *Conférence Méditerranéenne sur l'Ingénierie Sûre des Systèmes Complexes, MISC 2011*, 2011.
- [22] Michel Fliess. A mathematical proof of the existence of trends in financial time series. *Systems Theory : Modelling, Analysis and Control*, pages 43–62, 2009.
- [23] Yang Liu, Xiaohui Yu, and Jiqing Han. Sharpe ratio-oriented active trading : A learning approach. In *Mexican International Conference on Artificial Intelligence*, pages 331–339. Springer, 2002.
- [24] Young-Woo Seo, Joseph Giampapa, and Katia Sycara. Financial news analysis for intelligent portfolio management. Technical report, Carnegie-Mellon Univ Pittsburgh PA Robotics Inst, 2004.
- [25] Marika Salo and Helena Haapio. Robo-advisors and investors : Enhancing human-robot interaction through information design, 2017.
- [26] Claire Castanet and Camille Planes. Finance et intelligence artificielle : une révolution en marche. *Enjeux numériques*, page 15, 2018.





# Apprentissage de plongements de mots dynamiques avec régularisation de la dérive

Syrielle Montariol<sup>1,2</sup>

Alexandre Allauzen<sup>1</sup>

<sup>1</sup> LIMSI, CNRS, Univ. Paris-Sud, Univ. Paris-Saclay

<sup>2</sup> Société Générale

syrielle.montariol@limsi.fr

## Résumé

*L'usage, le sens et la connotation des mots peuvent changer au cours du temps. Les plongements lexicaux diachroniques permettent de modéliser ces changements de manière non supervisée. Dans cet article nous étudions l'impact de plusieurs fonctions de coût sur l'apprentissage de plongements dynamiques, en comparant les comportements de variantes du modèle Dynamic Bernoulli Embeddings. Les plongements dynamiques sont estimés sur deux corpus couvrant les mêmes deux décennies, le New York Times Annotated Corpus en anglais et une sélection d'articles du journal Le Monde en français, ce qui nous permet de mettre en place un processus d'analyse bilingue de l'évolution de l'usage des mots.*

## Mots Clef

Diachronie, Plongements lexicaux, analyse bilingue

## Abstract

***Learning dynamic word embeddings with drift regularization***

*Word usage, meaning and connotation change throughout time. Diachronic word embeddings are used to grasp these changes in an unsupervised way. In this paper, we use variants of the Dynamic Bernoulli Embeddings model to learn dynamic word embeddings, in order to identify notable properties of the model. The comparison is made on the New York Times Annotated Corpus in English and a set of articles from the French newspaper Le Monde covering the same period. This allows us to define a pipeline to analyse the evolution of words use across two languages.*

## Keywords

Diachrony, word embeddings, cross-lingual analysis

## 1 Introduction

Les langues sont des systèmes dynamiques : l'usage des mots évolue au cours du temps, reflétant les nombreux aspects des évolutions de la société, qu'ils soient culturels, technologiques ou dûs à d'autres facteurs [1].

La diachronie désigne l'étude de ces variations temporelles d'usage et de sens au sein d'une langue. Ici, nous étudions un corpus journalistique d'une plage temporelle de deux décennies : l'usage des mots évolue suite à des événements ayant un retentissement médiatique. Par exemple, l'usage du mot "Katrina" a connu un important changement au cours de ces deux décennies. Si par le passé il fut exclusivement utilisé comme un prénom féminin, dès 1999 son sens se rapproche de celui d'*ouragan*, avec l'arrivée du premier orage tropical éponyme. Puis à partir de 2005 où le cyclone Katrina eut lieu, ce qui fut un prénom féminin partage le même champs lexical que les mots *désastre*, *dévastation* et *inondation*.

Détecter et comprendre ces changements avec le concours de méthodes d'apprentissage automatique est utile à la recherche linguistique, mais aussi à de nombreuses tâches de traitement automatique des langues. Ajouter une notion temporelle aux représentations de mots permet d'étudier des corpus qui s'étendent sur des plages temporelles longues avec une plus grande acuité. Le problème se pose particulièrement aujourd'hui, alors qu'un nombre croissant de documents historiques sont numérisés et rendus accessibles ; leur analyse conjointe à celle de corpus contemporains, pour des tâches allant de la classification de documents à la recherche d'information, nécessite de prendre en compte la diachronie.

De nombreuses méthodes de représentations vectorielles de mots ont été mises au point depuis deux décennies [13]. Elles permettent de représenter les mots par des vecteurs continus de faible dimension : les plongements lexicaux ou *word embeddings*. Néanmoins, ces plongements lexicaux reposent sur l'hypothèse que le sens d'un mot est inchangé sur l'ensemble du corpus. Cette hypothèse d'une représentation statique peut s'avérer limitée. Ainsi en supposant qu'un changement dans le contexte usuel d'un mot reflète un changement dans la signification de ce mot, il est possible d'entraîner des plongements de mots diachroniques : qui évoluent au cours du temps en suivant les changements d'usage des mots.

Un tel modèle a été proposé récemment : *Dynamic Bernoulli Embedding* (DBE) [15]. Il apprend des représenta-

tions de mots qui évoluent au cours du temps selon les strates temporelles d'un corpus, en caractérisant la dérive de ces représentation d'une strate à l'autre au moyen d'un processus aléatoire gaussien. D'autres approches dans la littérature impliquent d'apprendre des plongements lexicaux pour chaque strate temporelle sans les relier chronologiquement [8, 7]; ou d'apprendre les plongements diachroniques de façon incrémentale sans contrôler la dérive de ces plongements [9].

Dans cet article nous prenons pour base le modèle DBE, qui présente un bon compromis entre simplicité et modularité, pour questionner l'importance de ces différents choix de modélisation. Dans ce but, nous analysons le comportement des plongements de mots appris à partir de ce modèle (décrit à la section 3) sur deux tailles de strates temporelles – mensuelle et annuelle – et l'appliquons (dans la section 4) à des corpus dans deux langues différentes : français et anglais. Les données en anglais proviennent du *New York Times Annotated Corpus*<sup>1</sup> [17], qui s'étend de 1987 à 2006; le corpus en français est constitué d'articles du journal *Le Monde* collectés de façon à couvrir la même période. Dans un deuxième temps, nous étudions l'évolution d'un mot à travers ces deux corpus conjointement.

## 2 État de l'Art

Les premières méthodes automatiques d'étude de la diachronie se basent sur la détection de changements dans les co-occurrences des mots, puis sur des approches basées sur la similarité distributionnelle [6] en construisant des mesures d'information mutuelle à partir de matrices de co-occurrences.

L'usage de méthodes d'apprentissage automatique basées sur les plongements lexicaux est récent et a connu une forte hausse d'intérêt depuis deux ans [11, 18].

Dans un des premiers articles employant ce type de méthode [9], les auteurs estiment des plongements lexicaux pour la première strate temporelle  $t_0$  puis mettent à jour ces plongements pour les strates temporelles suivantes, considérant les plongements au temps  $t - 1$  comme initialisation pour la strate  $t$ . D'autres travaux ont ensuite vu le jour, reposant sur l'apprentissage de façon indépendante des plongements lexicaux pour chaque strate temporelle. Néanmoins, les plongements ainsi obtenus ne sont pas directement comparables car appartiennent à des espaces vectoriels différents. Deux approches sont alors envisageables : d'une part, déterminer la meilleure transformation linéaire afin d'aligner les espaces de représentation à travers les périodes [7, 4, 10]; d'autre part, calculer la similarité cosinus entre chaque paire de mot à l'intérieur d'une strate temporelle, les similarités étant alors comparables d'une strate sur l'autre sans nécessiter d'alignement [9].

Les méthodes dites dynamiques constituent un second type d'approche. Le corpus d'étude est toujours divisé en strates temporelles, mais cette fois les plongements lexicaux diachroniques sont appris de façon conjointe sur l'ensemble

des strates. Ils sont ainsi placés dans un même espace de représentation dès l'apprentissage. Pour cela, des modèles bayésiens d'apprentissage de plongements lexicaux peuvent être utilisés [2] : les vecteurs sont liées à travers les périodes à l'aide d'un processus de diffusion temporel qui contrôle leur évolution. Poursuivant le même objectif, différentes méthodes ont été proposées [20, 15, 8] afin de mettre en évidence de façon jointe l'évolution continue du sens des mots. Ces méthodes permettent de s'affranchir de la limite de volume de données par strate temporelle lors de l'apprentissage.

La majorité de ces modèles sont évalués sur des corpus en anglais. À notre connaissance, bien que plusieurs auteurs ont expérimenté sur d'autres langues que l'anglais [7, 5], aucun travaux n'a tenté de comparer l'évolution de mots à travers plusieurs langues à l'aide de méthodes de plongements diachroniques.

## 3 Modèles de plongements de mots dynamiques

Nous partons du modèle *Dynamic Bernoulli Embeddings* (DBE) [15]. Il se base sur les plongements de mots de la famille exponentielle [16], qui sont une généralisation probabiliste du modèle *Continuous Bag-of-Words* (CBOW) [13]. Nous en fournissons une brève description avant de présenter les différentes variantes mises en place.

### 3.1 Le modèle DBE

L'objectif de ce modèle est de prédire un mot à partir de son contexte. Afin de désigner un mot  $v$  parmi un vocabulaire de taille  $V$ , on considère un vecteur de variables aléatoires binaires  $\mathbf{x}_v \in \{0, 1\}^V$ , où seule la composante associée au mot  $v$  vaut 1. Le mot  $v$  à la position  $i$  dans le corpus est donc représenté par le vecteur binaire  $\mathbf{x}_{iv}$  et son contexte  $\mathbf{c}_i$  est constitué des  $C$  mots avant et des  $C$  mots après ( $C$  étant la taille de la fenêtre). Ainsi,  $\mathbf{x}_{\mathbf{c}_i}$  regroupe l'ensemble des points constituant le contexte du mot  $i$ . Le modèle DBE prédit le vecteur binaire du mot  $\mathbf{x}_{iv}$  à partir de son vecteur de contexte  $\mathbf{x}_{\mathbf{c}_i}$  selon la loi de Bernoulli suivante :  $\mathbf{x}_{iv} | \mathbf{x}_{\mathbf{c}_i} \sim \text{Bern}(p_{iv})$ . Le paramètre de la loi de Bernoulli  $p_{iv}$  est calculé à partir des plongements lexicaux  $\rho_v$  du mot à prédire et  $\alpha_{v'}$  des mots du contexte :

$$p_{iv} = \sigma \left( \rho_v^T \left( \sum_{j \in \mathbf{c}_i} \sum_{v' \in V} \alpha_{v'} \mathbf{x}_{jv'} \right) \right). \quad (1)$$

Ainsi la somme sur  $v'$  sélectionne les plongements  $\alpha_{v'}$  des mots du contexte, qui sont ensuite additionnés (somme sur  $j$ ) afin de créer un vecteur représentant le contexte  $\mathbf{c}_i$ . Le paramètre de Bernoulli résulte de l'application de la fonction sigmoïde  $\sigma$  au produit scalaire de ce vecteur avec le plongement  $\rho_v$  du mot à prédire.

**Vers un modèle dynamique.** Pour rendre ce modèle dynamique, considérons un corpus composé de  $T$  strates temporelles indicées par  $t$ . Dans chaque strate, chaque mot  $v$  a deux types de représentations : celle en tant que mot de

1. <https://catalog.ldc.upenn.edu/LDC2008T19>

contexte  $\alpha_v$ , et celle en tant que mot central  $\rho_v$ . Le vecteur  $\alpha_v$  est considéré comme invariant : il est commun à toutes les strates temporelles. Seuls les plongements  $\rho_v$  évoluent au cours du temps selon la marche aléatoire gaussienne suivante :

$$\begin{aligned} \rho_v^{(0)} &\sim \mathcal{N}(0, \lambda_0^{-1} I), \\ \text{puis } \forall t \geq 1, \rho_v^{(t)} &\sim \mathcal{N}(\rho_v^{(t-1)}, \lambda^{-1} I). \end{aligned} \quad (2)$$

Le paramètre  $\lambda$ , nommé *dérive*, est le même pour l'ensemble des strates et contrôle l'évolution du vecteur  $\rho_v$  d'une strate temporelle sur l'autre.

**Apprentissage.** L'apprentissage de ce modèles'appuie sur une variante de la stratégie du *negative sampling* [13]. L'objectif est d'optimiser la fonction suivante :

$$\mathcal{L}(\rho, \alpha) = \mathcal{L}_{pos}(\rho, \alpha) + \mathcal{L}_{neg}(\rho, \alpha) + \mathcal{L}_{prior}(\rho, \alpha). \quad (3)$$

Le premier terme  $\mathcal{L}_{pos}$  représente la log-probabilité associée aux exemples positifs, tandis que le second ( $\mathcal{L}_{neg}$ ) correspond à celle associée à des exemples négatifs tirés aléatoirement. Le troisième terme agit comme un terme de régularisation sur  $\alpha$  et sur la dynamique des plongements  $\rho$ , et consiste à pénaliser le vecteur  $\rho_v^{(t)}$  lorsqu'il s'éloigne trop fortement du vecteur  $\rho_v^{(t-1)}$ , de la manière suivante :

$$\begin{aligned} \mathcal{L}_{prior}(\rho, \alpha) = & -\frac{\lambda_0}{2} \sum_v \|\alpha_v\|^2 - \frac{\lambda_0}{2} \sum_v \|\rho_v^{(0)}\|^2 - \\ & \frac{\lambda}{2} \sum_{v,t} \|\rho_v^{(t)} - \rho_v^{(t-1)}\|^2. \end{aligned} \quad (4)$$

### 3.2 Variantes de régularisation

La première variante du modèle se rapproche du principe d'apprentissage incrémental [9]. Elle consiste à supprimer la régularisation sur la dérive des plongements de mots. Dans ce cas, la fonction de coût ne prend en compte que les deux premiers termes de la log-prior ainsi que  $\mathcal{L}_{pos}$  et  $\mathcal{L}_{neg}$ . Par la suite, nous intituleons cette variante DBE-I (Incrémental).

La seconde variante consiste à abolir l'obligation de chronologie dans les vecteurs temporels successifs. En remplaçant la troisième composante de  $\mathcal{L}_{prior}$  par  $\sum_{v,t} \|\rho_v^{(t)} - \rho_v^{(0)}\|^2$ , on force le vecteur  $\rho_v^{(t)}$  à rester proche du plongement d'origine  $\rho_v^{(0)}$ . Ce principe est similaire à celui de [8], où les plongements de mots diachroniques sont appris de façon indépendante sur chaque strate temporelle. Cette variante est désignée par DBE-NC (Non Chronologique). Une autre version de cette dernière fonction est mise en place, afin de prendre en compte l'éloignement temporel. Dans ce but, la troisième composante de la log-prior  $\sum_{v,t} \|\rho_v^{(t)} - \rho_v^{(0)}\|^2$  est multipliée par un facteur temporel : le coefficient devient  $-\frac{\lambda}{2} * t$  et permet de contrôler l'éloignement à la prior sans ajouter de dépendance entre les strates temporelles successives. Cette dernière version est nommée DBE-SC (Semi Chronologique).

## 4 Expérimentation

Nous expérimentons à partir de notre propre implémentation du modèle DBE en PYTORCH. Dans un premier temps, nous analysons de façon quantitative le comportement des différentes variantes du modèle DBE définies précédemment. Puis nous observons de plus près la dérive des mots ; en particulier, nous mettons en place un processus pour comparer les évolutions des mots dans deux langues de façon conjointe.

### 4.1 Données et hyper-paramètres

Les données du *New York Times Annotated Corpus* sont composées de 1 855 000 articles s'étalant sur une période d'environ 20 ans, du 1<sup>er</sup> janvier 1987 au 19 juin 2007. Le journal *Le Monde* est un des quotidiens les plus lus en France ; nous en collectons des articles entre le 1<sup>er</sup> janvier 1987 et le 31 décembre 2006. Ces deux corpus sont divisés en  $T = 20$  strates temporelles annuelles<sup>2</sup> et  $T = 240$  strates temporelles mensuelles.

Pour construire le vocabulaire, nous sélectionnons pour les deux langues  $V = 40\,000$  mots selon leurs fréquences après avoir retiré les mots-outils. Nous sous-échantillonons les mots fréquents en retirant chaque mot  $i$  avec une probabilité  $p = 1 - \sqrt{\frac{10^{-5}}{\text{fréquence}(i)}}$ . Dans le corpus *LeMonde*, le nombre moyen de mots par strate temporelle est d'environ 3.5 millions pour les strates annuelles et 300k pour les strates mensuelles. Dans le corpus *NYT*, ce nombre est d'environ 9 millions pour les strates annuelles et 750k pour les strates mensuelles. Le corpus est ensuite divisé en échantillons d'apprentissage, de validation et de test. Ces derniers comprennent chacune 10 % des données tirées aléatoirement. Les embeddings sont entraînés avec 1000 mini-batches par strates temporelles pour l'analyse annuelle et 100 mini-batches pour l'analyse mensuelle.

Les hyper-paramètres sont sélectionnés à partir de l'étude de la log-probabilité sur les exemples positifs  $\mathcal{L}_{pos}$  calculée sur l'échantillon de validation de chaque corpus, à partir du modèle DBE classique. Afin de permettre la comparaison, les valeurs de  $\mathcal{L}_{pos}$  sont mises à l'échelle selon la règle suivante :

$$\text{Échelle} = \frac{\text{Nb de mots dans l'échantillon de validation}}{\text{Nb de mots dans chaque mini-batch}}.$$

Dans un premier temps, le modèle est entraîné sur l'ensemble du corpus sans composante temporelle (modèle statique). Ainsi, les plongements lexicaux  $\rho$  et  $\alpha$  peuvent servir par la suite d'initialisation pour les modèles temporels. Suite à l'évaluation sur l'échantillon de validation, la fenêtre de contexte choisie est  $C = 4^3$ . La dimension des plongements lexicaux est de 100 et le nombre d'exemples négatifs tirés pour chaque exemple positif est fixé à 10. Pour finir, la dérive  $\lambda = 1$  est celle qui offre les meilleurs résultats. La dérive initiale  $\lambda_0$  est fixée à  $\frac{\lambda}{1000}$  [15].

2. La dernière année du corpus *NYT* étant incomplète, elle n'est pas prise en compte dans l'analyse.

3. Deux mots précédant et deux mots suivant le mot central.

## 4.2 Évaluation quantitative

Dans un premier temps, nous analysons l’effet des différentes variantes de la fonction de coût sur les performances du modèle mesurées en terme de log-vraisemblance et sur la distribution des dérives des mots.

**Évolution de la log-vraisemblance.** Dans cette partie, nous calculons la log-probabilité du modèle DBE sur les exemples positifs  $\mathcal{L}_{pos}$  sur les données de test de chaque corpus, *NYT* et *LeMonde*, pour les deux tailles de strates temporelles. Nous appliquons le terme d’échelle décrit dans la partie 4.1.

Dans un premier temps, les plongements lexicaux sont appris sur l’ensemble du corpus de façon statique. Puis ils sont utilisés sur chaque strate annuelle du corpus pour calculer  $\mathcal{L}_{pos}$ . La courbe obtenue est plus basse que la courbe associée au modèle dynamique appris en initialisant les vecteurs à partir de ce modèle statique (Figure 1). On constate logiquement que l’apprentissage dynamique permet aux plongements d’être adaptés à chaque strate temporelle, et donc plus efficace pour prédire les données de test.

À l’inverse, pour les deux corpus, le modèle dynamique avec initialisation aléatoire a la performance la plus faible. Cette tendance est confirmée par la log-probabilité moyenne sur l’ensemble des strates temporelles (Table 1). Une explication se trouve peut être dans le faible volume de données sur chaque strate. Quand à la performance du modèle statique, elle dépend de l’homogénéité temporelle du corpus étudié; nos deux corpus couvrent une plage de temps relativement faible, justifiant la performance du modèle statique par rapport à celle du modèle dynamique avec initialisation aléatoire.

Comme le montre le tableau 1, les variantes du modèle définies par les différentes fonctions de coût ont des performances très proches; l’ajout du coefficient de dérive croissant (DBE-SC) au modèle non chronologique (DBE-NC) permet d’augmenter légèrement sa performance, mais dans l’ensemble, c’est le modèle sans régularisation sur la dérive (DBE-I) qui obtient le score le plus élevé quelle que soit la taille de la strate temporelle.

L’étude de la log-probabilité ne reflète qu’une vision globale des performances du modèle; afin de mieux comprendre son comportement, nous observons ensuite la distribution des dérives pour chaque variation du modèle.

**Caractérisation de la dérive des plongements.** Dans le but d’analyser plus en détail le comportement du modèle et l’effet des variations de la fonction de coût, nous représentons les histogrammes superposés des dérives successives observées sur le corpus *LeMonde* (Figure 2). Les histogrammes pour le corpus *NYT* présentent des tendances similaires, de même que le cas des strates temporelles mensuelles. La dérive de chaque mot est calculée à partir de la distance euclidienne entre le plongement du mot au début du corpus  $\rho_v^{(t_0)}$  et ses plongements successifs  $\rho_v^{(t)}$  à chaque nouvelle strate temporelle  $t$ . Sur les histogramme, les cou-

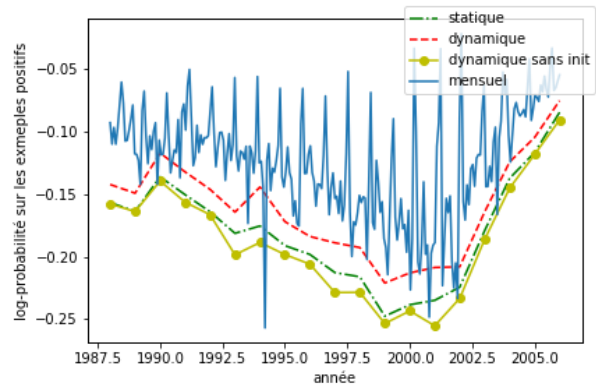


FIGURE 1 – Log-probabilités sur les exemples positifs sur l’échantillon de test du corpus *LeMonde* à partir des modèles statique, dynamiques, et dynamiques sans pré-entraînement pour des strates annuelles, et dynamique pour des strates mensuelles.

	NYT		Le Monde	
	annuel	mensuel	annuel	mensuel
Statique	-0.09875	-0.06771	-0.1794	-0.1259
DBE	-0.08476	<b>-0.06720</b>	-0.1606	-0.1231
DBE init aléatoire	-0.10774	-0.07305	-0.1873	-0.1498
DBE-I	<b>-0.08448</b>	-0.06752	<b>-0.1593</b>	<b>-0.1227</b>
DBE-NC	-0.08517	-0.06817	-0.1607	-0.1236
DBE-SC	-0.08455	-0.06752	-0.1598	-0.1228

TABLE 1 – Log-probabilités moyennes sur l’ensemble des strates temporelles, sur l’échantillon de test des deux corpus, pour les différentes variantes d’apprentissage et de fonction de coût du modèle DBE.

leurs plus claires correspondent aux distributions des dérives aux strates récentes : ainsi, la courbe la plus claire représente la distribution des dérives de mots calculées entre  $t_0 = 1987$  et  $t = 2006$  tandis que la plus sombre représente la distribution des dérives entre  $t_0 = 1987$  et  $t = 1988$ .

Une première propriété intéressante est le caractère dirigé des dérives. Comme le montre le premier histogramme de la figure 2, les valeurs des dérives augmentent à travers le temps pour le modèle DBE classique. Cela signifie que le modèle capture principalement des dérives possédant une tendance, plutôt que de brefs changements de plongements suivis de retours à la normale. Ainsi le terme de régularisation décrit par l’équation 4 réalise bien le compromis attendu, en considérant comme partie de l’objectif la détection des grandes tendances d’évolution du sens des mots et en omettant leurs brèves variations.

Ces brèves variations sont dues à des événements qui modifient temporairement le contexte dans lequel apparaît un mot sans avoir un impact à long terme sur son sens. Elles sont plutôt capturées par la version DBE-NC du modèle, dont l’histogramme ne présente pas d’évolution dirigée de la dérive en fonction de la distance à  $t_0$ , donc ne distingue pas ces "bruits" de la tendance générale d’évolution des

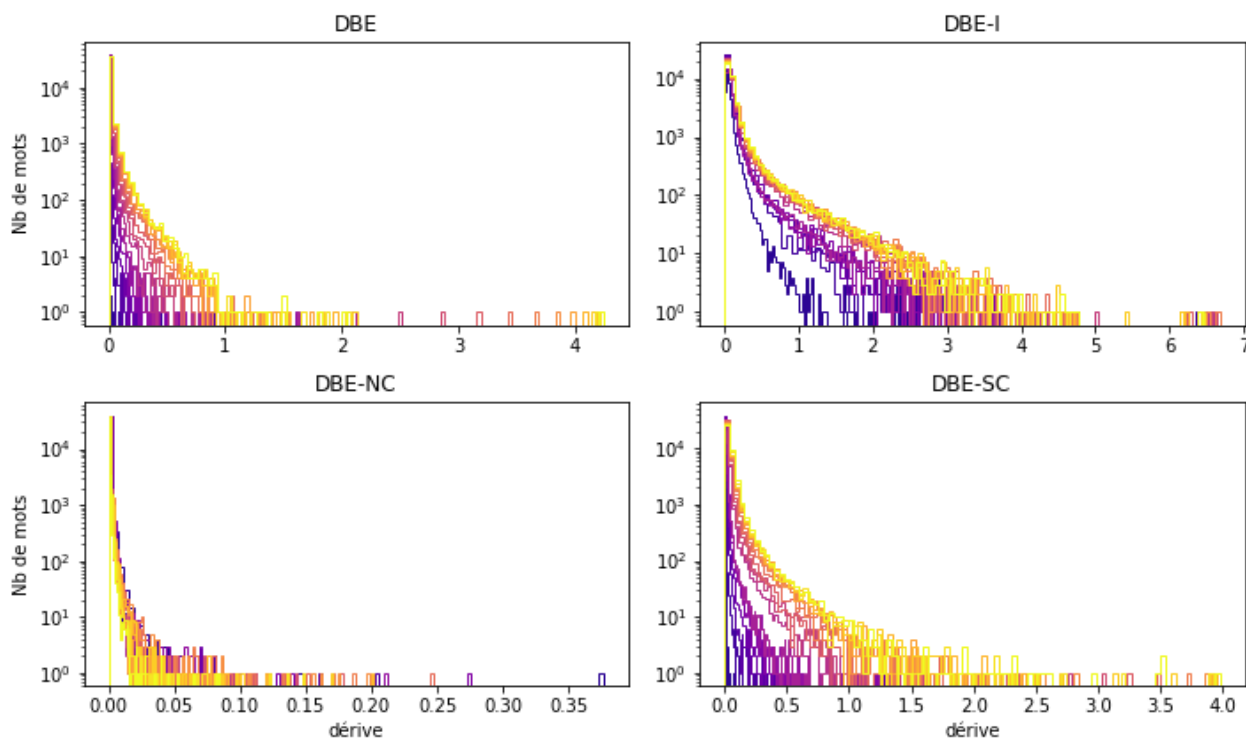


FIGURE 2 – Histogramme des dérives entre les plongements de mots à  $t_0 = 1987$  et à chaque strate temporelle annuelle successive du corpus *LeMonde*, pour le modèles DBE et ses trois variantes. Plus la couleur est claire, plus la différence est calculée par rapport à une strate temporelle récente. Les nombres de mots (en ordonnée) sont en échelle logarithmique.

mots. Pour finir, malgré l’absence de terme de régularisation sur la dérive, le modèle DBE-I capture naturellement une dérive relativement dirigée dans le temps bien que l’histogramme montre une plus grande sensibilité au bruit.

La seconde propriété mise en évidence est la capacité du modèle à distinguer les mots stables des mots dont l’usage évolue. En effet dans un intervalle de deux décennies, la majorité des mots est supposée peu évoluer. Le modèle DBE-NC, en introduisant une régularisation par rapport aux plongements de mots initiaux, permet de forcer le respect cette propriété : une grande part des mots sont presque invariants sur tout le corpus, et seule une sélection de dérives se démarque. Le modèle DBE classique permet aussi, dans une certaine mesure, de garder une faible dérive pour une grande partie des mots ; de même pour le modèle DBE-SC. Seul le modèle DBE-I ne distingue pas naturellement les mots qui dérivent peu.

### 4.3 Évaluation qualitative

La seconde étape de l’analyse est d’observer directement l’évolution des mots. À notre connaissance, il n’existe pas de corpus annotés permettant une évaluation directe des modèles diachroniques. Il est par contre possible d’observer l’évolution de certains mots choisis, permettant un premier diagnostic sous la forme d’une évaluation qualitative

et subjective [18]. Ainsi pour chaque variante du modèle, nous observons les mots qui dérivent le plus pour mieux en comprendre le comportement. Puis, afin d’étudier de façon conjointe l’évolution des mots sur les deux corpus, nous mettons en place un processus d’analyse diachronique inter-langues.

**Analyse des fortes dérives.** Nous nous concentrons ici sur le corpus *LeMonde*. La période étudiée ne couvrant que deux décennies, on observe principalement des évolutions de contexte liées aux événements ayant un impact médiatique ; les mots subissant de fortes dérives sont en majorité des entités nommées, et sont liés au contexte politique de la période, un thème récurrent dans ce journal.

Nous listons les 10 mots dont l’usage a le plus dérivé au cours des deux décennies selon chaque modèle dans le corpus, pour les deux tailles de strates temporelles. Du corpus *NYT*, nous ne reportons que les 10 mots ayant le plus varié d’après le modèle DBE classique sur des strates temporelles annuelles (Table 2).

Dans le cas du modèle DBE, pour les strates mensuelles et annuelles, les mots qui dérivent le plus sont associés à des concepts ayant subi un changement notable et continu au cours de la période (*euros*, *al-qaïda*, *internet*...). Cette observation est en accord avec la propriété observée précédemment au sujet du caractère dirigé des dérives. À l’in-

NYT	Le Monde					
Annuel	Annuel			Mensuel		
DBE	DBE	DBE-I	DBE-NC	DBE	DBE-I	DBE-NC
google	euros	clearstream	royal	euros	rfa	sez nec
skilling	ump	arcelor-mittal	sarkozy	sarkozy	euros	tramway
bloomberg	villepin	raimond	gdf	ump	ségolène	pinochet
email	rfa	shultz	euros	francs	sarkozy	hamas
katrina	sarkozy	ségolène	hezbollah	ségolène	ump	euros
cellphone	al-qaïda	outreau	liban	villepin	villepin	ahmadinejad
darfur	poutine	eads	thaksin	internet	monory	abbas
contras	gorbatchev	zapatero	ump	ue	climatique	révision
blog	katrina	villepin	islam	euro	contras	fibre
euros	internet	zidane	suez	bush	réévaluation	mahmoud

TABLE 2 – Listes des 10 mots ayant la plus grande dérive totale (distance entre la première et la dernière strate temporelle) pour les modèles DBE, DBE-I et DNE-NC sur le corpus *LeMonde* et DBE sur le corpus *NYT*. Les mots ayant un arrière-plan coloré sont communs à plus d’un modèle.

	Annuel		Mensuel	
	top 500	top 10	top 500	top 10
DBE	0.00642	0.00785	0.00356	6.280e-05
DBE-I	0.1152	0.0272	0.1149	0.0253
DBE-NC	0.5259	0.1054	0.0767	0.0530
DBE-SC	0.1096	0.0301	0.0715	0.0164

TABLE 3 – Valeurs moyennes pour les 10 mots et les 500 mots dérivant le plus, de leur dérive moyenne normalisée. Les dérives sont calculées pour les 4 variantes du modèle DBE sur le corpus *LeMonde*.

verse, les modèles DBE-I et DBE-NC annuels mettent principalement en valeur des dérives très fortes de mots liées à des événements uniques (*zidane*, *clearstream*, *royal*). C’est particulièrement le cas pour le modèle DBE-NC sur les strates mensuelles, où les dérives rapportées sont sujettes à un bruit important.

Afin de confirmer ces observations, le tableau 3 moyenne pour chaque modèle les rapports entre la dérive moyenne (sur toutes les strates temporelles) et la dérive totale (entre la première et la dernière strate) des 10 et 500 mots qui évoluent le plus. Notons que les valeurs des dérives moyennes normalisées ne sont pas directement comparables, selon que l’on considère les strates annuelles ou mensuelles (car le nombre de strates diffère). La valeur moyenne pour les 10 mots qui dérivent le plus est toujours plus faible que celle sur les 500 mots qui dérivent le plus. Les dérives importantes sont donc les plus dirigées quel que soit le modèle. Le modèle DBE-I présente des valeurs très proches de celles du modèle DBE-SC, montrant que l’absence de régularisation par rapport à la dérive permet, dans une certaine mesure, de conserver une certaine robustesse au bruit, en adéquation avec les observations de la figure 2.

Pour finir, remarquons que parmi les mots qui ont le plus dérivé au cours des deux décennies, certains sont communs aux deux corpus (*euros*, *google*, *katrina*). Nous proposons donc par la suite une méthode pour observer l’évolution conjointe de ces mots dans les deux langues.

**Analyse conjointe en anglais et français.** Dans cette partie, nous étudions l’évolution d’un mots en français dans le corpus *LeMonde* et de sa traduction anglaise dans le corpus *NYT*. Les modèles de plongements de mots étant appris de façon indépendante sur ces deux corpus, les vecteurs ne sont pas directement comparables. Nous effectuons alors un alignement des deux espaces de représentation en utilisant un dictionnaire bilingue comme outil de supervision [3]. Dans un premier temps, les plongements de mots des deux langues appris de façon statique sur l’ensemble du corpus sont normalisés; puis, l’espace de représentation des plongements en français est aligné sur l’espace vectoriel des plongements en anglais. Nous choisissons ce sens car les données du *NYT* sur lesquelles les plongements anglais sont appris ont une volumétrie plus élevée, permettant des plongements lexicaux plus robustes.

Nous utilisons l’outil MUSE<sup>4</sup> [12]. pour l’alignement. La supervision est effectuée au moyen d’un dictionnaire bilingue construit à partir des vocabulaires des deux corpus. Nous sélectionnons tous les mots ayant un équivalent dans l’autre langue à partir du dictionnaire fourni par MUSE, puis ajoutons manuellement une sélection de mots spécifiques aux données (principalement des entités nommées). À partir des deux vocabulaires de 40 000 mots, nous obtenons un vocabulaire bilingue de 27 351 mots. Pour finir, les plongements sémantiques alignés  $\rho_{align}$  and  $\alpha_{align}$  sont utilisés pour initialiser les modèles dynamiques entraînés sur chaque corpus.

Suite à l’apprentissage, pour chaque couple de mot dans le dictionnaire bilingue, nous calculons leurs dérives dans les deux corpus. Puis, nous calculons le cosinus comme une similarité entre les plongements des deux mots à la première et la dernière strate temporelle. Appelons cette valeur la similarité inter-langues. Nous calculons la dérive de cette similarité entre la première et la dernière strate temporelle, en mesurant la distance euclidienne entre ces deux valeurs. En observant la distribution de ces grandeurs, nous mettons en évidence 4 types de comportement de mot à travers les

4. <https://github.com/facebookresearch/MUSE>



deux langues :

1. Les mots qui dérivent dans la même direction dans les deux langues ;
2. Les mots qui dérivent dans les deux langues, mais dont le sens diverge (la similarité inter-langues décroît entre la première et la dernière strate temporelle) ;
3. Les mots qui dérivent dans une seule des deux langues, tandis que l'autre reste stable ;
4. Les mots qui sont stables dans les deux langues.

Nous différencions les classes de mots en utilisant la moyenne des dérive des mots de chaque langue ainsi que la moyenne de la dérive de la similarité inter-langues, et reportons leur répartition au sein du vocabulaire bilingue dans le tableau 4. La majorité des mots appartiennent à la catégorie (4) (mots stables dans les deux langues), ce qui confirme la propriété du modèle DBE énoncée à partir de la figure 2 ; tandis que les mots qui évoluent dans les deux langues (catégories (1) et (2)) sont les plus rares.

Classe	(1)	(2)	(3-fr)	(3-en)	(4)
%	5.4	5.5	16.1	15.2	57.8
Ex.	renouvelable	soviétique	francs	patrie	savon

TABLE 4 – Proportion des mots dans les différentes classes de comportement de dérive inter-langues, avec un exemple pour chaque classe.

Considérons à titre d'exemple le mot *Barbie*, qui appartient à la 3ème catégorie. L'espace aligné des plongements de mots est réduit à deux dimensions au moyen de la méthode t-SNE [19] pour représenter l'évolution de ce mot dans les deux langues (figure 3). Les mots les plus similaires au mot *Barbie* dans chaque langue, et à chaque strate temporelle, sont indiqués sur le graphique. En français (en rouge), le mot cible ne subit pas d'évolution notable. Il est majoritairement associé au criminel nazi *Klaus Barbie* et à son procès. Ces événements ont eu une couverture médiatique moins importante et plus ponctuelle dans les journaux américains ; l'équivalent anglais du mot cible évolue rapidement en direction du champs lexical de la mode, en association avec la célèbre marque de poupée homonyme. Son plongement sémantique se stabilise dans ce voisinage à partir des années 2000.

## 5 Discussion

Dans cet article, nous étudions en détail le comportement d'un modèle d'apprentissage de plongements lexicaux dynamiques. Le point de départ de notre étude est le modèle *Dynamic Bernoulli Embeddings*, dont nous définissons plusieurs variantes. Nous répliquons ainsi le comportement d'autres modèles d'apprentissage de plongements de mots diachroniques de la littérature. Deux propriétés nous paraissent importantes à distinguer pour bien caractériser ces modèles : la capacité à mettre en évidence des évolutions dirigées des plongements de mots, et la capacité

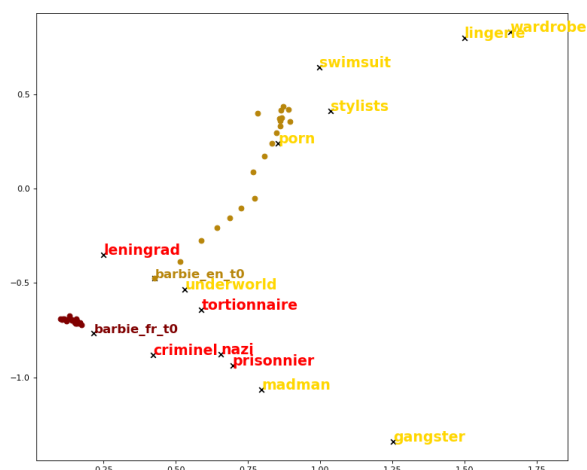


FIGURE 3 – T-SNE des espaces de plongements sémantiques alignés, représentant l'évolution temporelle du mot *Barbie* en français (rouge) et en anglais (jaune) ainsi que leurs plus proches voisins au cours du temps sur des strates annuelles.

à garder une partie du vocabulaire stable au cours du temps. Nous montrons ensuite qu'il est possible d'analyser l'évolution d'un mot dans deux langues de façon conjointe. Un processus d'analyse préliminaire est mis en place en initialisant les modèles dynamiques à partir de plongements de mots alignés, et en analysant la dérive de la similarité inter-langues.

Le domaine en plein essor qu'est l'apprentissage de plongements de mots dynamiques manque encore de la cohésion que possèdent les tâches plus anciennes du traitement automatique des langues. Les publications sur ce sujet portent sur des corpus très diversifiés et les évaluations se font le plus souvent de façon qualitative, en l'absence de base d'évaluation robuste. Notons qu'un cadre d'évaluation est difficile à définir dans le cas qui nous intéresse, tant les attentes applicatives vis-à-vis d'un modèle diachronique peuvent varier. De plus, un cadre mathématique commun et rigoureux n'a pas encore été défini [11] et devrait s'appuyer sur les modèles d'apprentissage conjoint sur toutes les strates temporelles tel que celui décrit ici.

En effet, l'apprentissage conjoint à travers toutes les strates permet de s'affranchir dans une certaine mesure de la nécessité d'avoir un grand volume de données dans chacune d'elle. Néanmoins, le caractère discontinu des strates temporelles induit le modèle à détecter seulement les dérives d'une strate à l'autre ; plus les strates sont larges, plus les variations internes sont cachées, ou du moins moyennées. Ainsi, la question de la juste granularité temporelle se pose et dépend de l'application visée. Il est cependant important que les modèles étudiés puissent travailler à différents niveaux de finesse temporelle. Par exemple, lors de la recherche de dérives sémantiques brusques et de court terme, un type de modèle en temps continu [14] pourrait être plus

adéquat, mais nécessite des informations temporelles très précises qui en pratique se retrouvent presque exclusivement dans les corpus issus de médias sociaux.

Une alternative est d'explorer l'usage de processus temporels de diffusion plus complexes, travaux initiés par exemple par [2] avec le processus d'Ornstein-Uhlenbeck. Enfin, l'emploi de strates temporelles non fixes dont les ruptures seraient apprises en même temps que les plongements lexicaux, assorti d'une régularisation sur la fonction de coût similaire à celle du modèle DBE-SC, serait une alternative à explorer. Néanmoins le cadre théorique approprié à cette sorte de quantification temporelle apprise par le modèle reste à définir.

## Références

- [1] J. Aitchison. Language change : Progress or decay ? In *Cambridge Approaches to Linguistics*. Cambridge University Press, 2001.
- [2] R. Bamler and S. Mandt. Dynamic word embeddings. In *Proceedings of the 34th International Conference on Machine Learning*, pages 380–389, 2017.
- [3] A. Conneau, G. Lample, M. Ranzato, L. Denoyer, and H. Jégou. Word translation without parallel data. *CoRR*, 2018.
- [4] H. Dubossarsky, D. Weinshall, and E. Grossman. Outta control : Laws of semantic change and inherent biases in word representation models. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 1136–1145. Association for Computational Linguistics, 2017.
- [5] S. Eger and A. Mehler. On the linearity of semantic change : Investigating meaning variation via dynamic graph models. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, pages 52–58, 2016.
- [6] K. Gulordava and M. Baroni. A distributional similarity approach to the detection of semantic change in the google books ngram corpus. In *Proceedings of the GEMS 2011 Workshop on GEometrical Models of Natural Language Semantics*, pages 67–71. Association for Computational Linguistics, 2011.
- [7] W. L. Hamilton, J. Leskovec, and D. Jurafsky. Diachronic word embeddings reveal statistical laws of semantic change. In *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics*, pages 1489–1501. Association for Computational Linguistics, 2016.
- [8] R. Han, M. Gill, A. Spirling, and K. Cho. Conditional word embedding and hypothesis testing via bayes-by-backprop. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 4890–4895. Association for Computational Linguistics, 2018.
- [9] Y. Kim, Y.-I. Chiu, K. Hanaki, D. Hegde, and S. Petrov. Temporal analysis of language through neural language models. In *Proceedings of the ACL 2014 Workshop on Language Technologies and Computational Social Science*, pages 61–65. Association for Computational Linguistics, 2014.
- [10] V. Kulkarni, R. Al-Rfou, B. Perozzi, and S. Skiena. Statistically significant detection of linguistic change. In *Proceedings of the 24th International Conference on World Wide Web*, pages 625–635. International World Wide Web Conferences Steering Committee, 2015.
- [11] A. Kutuzov, L. Øvrelid, T. Szymanski, and E. Velldal. Diachronic word embeddings and semantic shifts : a survey. In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 1384–1397. Association for Computational Linguistics, 2018.
- [12] G. Lample, A. Conneau, L. Denoyer, and M. Ranzato. Unsupervised machine translation using monolingual corpora only. 2017.
- [13] T. Mikolov, I. Sutskever, K. Chen, G. S. Corrado, and J. Dean. Distributed representations of words and phrases and their compositionality. In *Advances in Neural Information Processing Systems 26*, pages 3111–3119. 2013.
- [14] A. Rosenfeld and K. Erk. Deep neural models of semantic shift. In *NAACL 2018*, 2018.
- [15] M. Rudolph and D. Blei. Dynamic embeddings for language evolution. In *Proceedings of the 2018 World Wide Web Conference on World Wide Web*, pages 1003–1011, 2018.
- [16] M. Rudolph, F. Ruiz, S. Mandt, and D. Blei. Exponential family embeddings. In *Advances in Neural Information Processing Systems*, pages 478–486, 2016.
- [17] E. Sandhaus. The new york times annotated corpus. In *Philadelphia : Linguistic Data Consortium*. Vol. 6, No. 12., 2008.
- [18] N. Tahmasebi, L. Borin, and A. Jatowt. Survey of computational approaches to diachronic conceptual change. *CoRR*, 2018.
- [19] L. van der Maaten and G. Hinton. Visualizing data using t-SNE. *Journal of Machine Learning Research*, pages 2579–2605, 2008.
- [20] Z. Yao, Y. Sun, W. Ding, N. Rao, and H. Xiong. Dynamic word embeddings for evolving semantic discovery. In *Proceedings of the Eleventh ACM International Conference on Web Search and Data Mining*, pages 673–681. ACM, 2018.



# Application du Clustered Deep Q-Network aux politiques tarifaires

S. Pageaud<sup>1,2</sup>V. Deslandres<sup>1</sup>V. Lehoux<sup>2</sup>S. Hassas<sup>1</sup><sup>1</sup> Université de Lyon - Université Claude Bernard Lyon 1 - LIRIS CNRS UMR 5205, Lyon, France<sup>2</sup> NAVER LABS Europe - Meylan, France

simon.pageaud@liris.cnrs.fr

## Résumé

Ce travail présente une nouvelle approche multi-agent et multi-niveaux, nommée *Clustered Deep Q-Network (CDQN)*, avec pour objectif de répondre aux problèmes de passage à l'échelle et de non-stationnarité dans des contextes d'apprentissages décentralisés. Notre approche repose sur : 1) une gestion de chaque agent dans des clusters dynamiques avec une action jointe contrainte pour réduire la non-stationnarité ; 2) l'attribution d'un score de confiance joint pour évaluer la contribution individuelle de chaque agent. Les expérimentations et les résultats sur une politique urbaine montrent que notre modèle permet une coordination efficace d'agents indépendants en utilisant l'apprentissage par renforcement profond multi-agent et la réutilisation d'expériences pour augmenter à la fois le gain individuel et global.

## Mots-clés

Simulation multi-agent, apprentissage par renforcement multi-agent, apprentissage par renforcement profond, politiques urbaines.

## 1 Introduction

Les progrès en apprentissage par renforcement profond a permis aux agents d'atteindre un niveau de contrôle humain sur de nombreux domaines comme les jeux Atari [14]. Cependant, ces jeux requièrent uniquement un agent apprenant. Les améliorations au Deep Q-Network (DQN) dans un contexte coopératif permet à deux agents de jouer à des jeux Atari ensemble [22]. Leur approche repose sur l'Independent Q-learning où les agents apprennent leur propre Q-fonctions indépendamment en parallèle [21]. Les limites de leur travail est l'hypothèse que chaque agent est dans un environnement entièrement observable. Les travaux avec les Q-networks profonds récurrents [7] permettent à des agents de jouer à des jeux dans des environnements 3D partiellement observables [9] dans un cadre mono-agent. Plus récemment, certains travaux se concentrent sur la combinaison des observations partielles avec un cadre multi-agent. Le Q-network profond récurrent distribué [3] fournit au cadre multi-agents des outils pour initier des protocoles de communications et s'organiser entre eux pour résoudre des

énigmes.

Les cadres mono-agents passent difficilement à l'échelle avec l'augmentation de l'espace d'états. L'utilisation des approches multi-agents surmontent ce problème en utilisant des politiques décentralisées où les agents choisissent leurs actions uniquement à partir de leur historique local d'états/action. L'entraînement centralisé de politiques décentralisées est une approche classique de l'apprentissage par renforcement multi-agent [16] et l'apprentissage par renforcement profond [3]. Dans ce papier, nous considérons un apprentissage décentralisé sans informations supplémentaires sur l'état et sans communications entre agents. Une des approches les plus populaires en apprentissage par renforcement multi-agent, l'Independent Q-Learning [23], permet à l'agent d'apprendre ses Q-valeurs à partir de son ensemble d'états/actions en considérant les autres agents comme une partie de l'environnement. Le problème avec l'utilisation de l'apprentissage par renforcement profond avec les apprenants indépendants est que l'environnement reste non-stationnaire et impacte fortement la réutilisation d'expériences nécessaires au DQN.

Un autre défi lié à l'utilisation de politiques décentralisées est l'identification de la contribution des agents (*multi-agent credit assignment* [2]), où les actions jointes ne génèrent généralement que des avantages globaux, ce qui rend difficile pour chaque agent d'inférer sa propre contribution au gain global. Une solution consiste à concevoir des fonctions de récompense individuelles pour chaque agent. Cependant, dans les contextes coopératifs, cela incite les agents à adopter des comportements individuels plutôt que coopératifs.

Les travaux précédents mentionnés dans les contextes partiellement observables et multi-agents reposent sur le fait que les agents ont une connaissance des autres agents et qu'ils peuvent communiquer entre eux. Nous considérons le cas où les agents ne peuvent pas communiquer et ont aucune connaissance des actions effectuées par les autres agents. Dans ce contexte, les agents doivent pouvoir se coordonner afin de maximiser une mesure de performance ciblée.

Ce document présente l'architecture CDQN (Clustered Deep Q-Network) permettant aux agents contrôleurs de gé-

rer des clusters d'agents apprenants et de les coordonner efficacement pour améliorer leurs politiques. Ils apprennent des politiques décentralisées tandis que les agents contrôleurs ont pour rôle de réduire la non-stationnarité lors de la collecte de l'historique de l'action-observation locale. Nous montrons que la combinaison d'agents DQN indépendants avec des agents contrôleurs permet de résoudre le problème de la communication limitée dans des environnements multi-agents partiellement observables. Il permet également de réutiliser l'expérience et d'identifier la contribution individuelle de chaque agent plus facilement.

Nous introduisons quatre éléments d'importance primordiale pour l'efficacité du CDQN : i) un modèle de populations multi-niveaux et multi-agents dans lequel les agents contrôleurs gèrent un groupe d'agents apprenant ; ii) une attribution de récompense unique au lieu d'une récompense jointe. Les agents apprenants ont ainsi une récompense locale basée sur leurs propres observations ; iii) un score de confiance : la sélection et la gestion du comportement des agents apprenants via l'attribution de scores de confiance par les agents contrôleurs ; iv) des actions de contrôle permettant la réorganisation dynamique de groupes d'agents d'apprenants avec un mécanisme de fusion / séparation. La population d'agents de contrôle évolue à travers les ajouts et suppressions d'agents contrôleurs.

L'évaluation est réalisée au moyen de trois expériences basées sur des problèmes de politique de prix du monde réel utilisant les approches IDQL et CDQN. Un intervenant recherche une répartition satisfaisante des prix de stationnement afin de maximiser le gain dans une ville avec un découpage approprié. Nos résultats montrent que le CDQN augmente efficacement le gain cumulé global et parvient à coordonner les agents apprenants de manière décentralisée, grâce à leur organisation dynamique en clusters sans aucune communication entre eux.

## 2 Travaux similaires

Les méthodes d'apprentissage par renforcement multi-agents étaient auparavant axées sur les méthodes tabulaires [1] avant d'utiliser des techniques d'apprentissage par renforcement profond. Ces méthodes permettent d'avoir des espaces d'actions et/ou d'états plus vastes [22]. Pour la planification d'une politique urbaine, nous considérons un cadre coopératif.

Dans les environnements d'apprentissage par renforcement multi-agents profonds, nous distinguons deux approches principales. Une première s'attaque à l'apprentissage centralisé d'actions jointes, ce qui permet une meilleure coordination et évite la non-stationnarité. La limite de cette approche est la croissance exponentielle de l'espace d'actions jointes avec le nombre d'agents. Certains travaux antérieurs reposent sur la communication entre les agents [19] ou nécessitent une connaissance préalable étendue de la relation entre les agents [5].

Une autre approche considère l'apprentissage décentralisé d'actions jointes pour améliorer la gestion d'un espace

d'actions jointes croissant. Le  $Q$ -Learning (IQL) indépendant [23] propose que chaque agent apprenne sa propre fonction  $Q$  à partir de son propre historique d'observation/action en considérant les autres agents comme faisant partie de l'environnement. Chaque agent  $i$  observe l'état actuel  $s_t$ , choisit une action  $a_i(t)$  en fonction de sa politique, puis reçoit une récompense d'équipe  $r_t$  partagée entre tous les agents. Dans des environnements partiellement observables, l'IQL peut être implémenté en demandant à chaque agent d'apprendre son historique d'observation/action. L'IQL a récemment été utilisé dans les contextes d'apprentissage par renforcement profond [22] où il est combiné au DQN, produisant ainsi un DQN indépendant appliqué dans un jeu compétitif de Pong sur Atari. Tous les agents apprennent en parallèle une  $Q$ -fonction  $Q_i(s, a_i; \theta_i)$  où  $\theta_i$  est le paramètre du réseau de neurones. Cependant, le  $Q$ -learning indépendant rencontre des problèmes de convergence en raison de l'environnement non stationnaire induit par l'apprentissage des agents par rapport aux autres agents. Dans les approches tabulaires, les résultats empiriques montrent que la non-stationnarité est surmontée [12], mais reste présent dans l'apprentissage par renforcement profond. L'approche Deep Q-Network (DQN) [15] s'appuie sur la réutilisation d'expériences pour stocker l'expérience dans une mémoire tampon et l'utiliser pour entraîner le modèle [11]. Au fur et à mesure que les agents améliorent leurs politiques, les autres agents le font également, ce qui entraîne le problème du *moving target*, où l'expérience antérieure devient obsolète. Certains travaux proposaient un moyen d'éviter ce problème en limitant la taille du tampon d'expérience [10] ou même en les désactivant [3].

Il existe des approches hybrides utilisant un apprentissage centralisé et une exécution décentralisée pour entraîner des agents décentralisés [4, 6].

## 3 Clustered Deep Q-Network

Le Clustered Deep Q-Network (CDQN) est composé de populations multi-agents, multi-niveaux utilisant l'attribution d'un score de confiance et des récompenses locales pour coordonner les actions, dans des environnements partiellement observables et ne permettant pas de communication. Le but de la planification dans CDQN est de trouver une politique commune  $\pi = \langle \pi_1, \dots, \pi_n \rangle$  qui maximise le gain cumulé global. Le modèle formel CDQN est une extension du Dec-POMDP factorisé [17].

### 3.1 Agents multi-niveaux

La coordination multi-agents de notre approche repose sur deux populations d'agents décentralisés. Nous considérons l'ensemble fini d'agents augmentés  $\hat{\mathcal{D}} = \mathcal{D}^l \cup \mathcal{D}^c$  avec un ensemble fini d'agents  $\mathcal{D}^l = (d_1^l, \dots, d_m^l)$ , appelée *apprenants* dont leur nombre  $m$  est fixé par la simulation, et l'ensemble fini d'agents  $\mathcal{D}^c = (d_1^c, \dots, d_n^c)$ ,  $n \in [1, m]$ , appelés *contrôleurs*, chacun gérant un clusters d'apprenants. Le nombre de contrôleurs varie au cours de la simu-

lation avec l'ajout et la suppression des contrôleurs pour une meilleure répartition des clusters. À tout moment, un apprenant est géré par un seul contrôleur. Par conséquent, le nombre possible de configurations de contrôleurs correspond au nombre de partitions de l'ensemble des apprenants en sous-ensembles disjoints et non vides, c'est-à-dire le  $n^{\text{ème}}$  nombre de Bell.

Seuls les contrôleurs peuvent communiquer entre eux et les apprenants connaissent uniquement leur agent contrôleur courant.

L'ensemble fini d'actions augmentées est défini par  $\hat{\mathcal{A}} = \mathcal{A}^c \cup \mathcal{A}^l$  où  $\mathcal{A}^c$  est l'ensemble fini d'actions de contrôle disponibles pour les contrôleurs et  $\mathcal{A}^l$  l'ensemble fini d'actions environnementales disponibles pour les apprenants. Une action de contrôle  $a^c \in \mathcal{A}^c$  est une action qui façonne des clusters par opposition à une action environnementale  $a^l \in \mathcal{A}^l$  qui modifie l'état. À chaque pas de temps, un apprenant  $i$  suggère une action  $a_i^s \in \mathcal{A}_i^l$  à son contrôleur. Les actions suggérées par les apprenants d'un cluster  $j \in [1, n]$  produit  $\mathbf{a}_j^s$ , l'action jointe suggérée des apprenants du cluster  $j$ . Actuellement, le contrôleur  $j$  utilise un système de vote pour choisir l'action élue  $a_j^l$  parmi les actions  $\mathbf{a}_j^s$  avec le plus grand nombre de suggestions. L'action choisie résultante  $a_j^l$  est appliquée par chaque apprenant de ce cluster sur l'environnement formant l'action jointe appliquée  $\mathbf{a}_j^a = (a_j^l)_{i \in j}$ . C'est la première contribution du modèle CDQN : à un moment donné, chaque apprenant d'un même cluster agit dans la même direction en appliquant la même action, chacun ayant ainsi la même transition en mémoire. Cela réduit la non-stationnarité pendant l'apprentissage.

Un apprenant  $d_i^l, i \in [1, m]$  a un ensemble de variables d'états  $x_i$  et un réseau représentant sa fonction de valeur individuelle  $Q_i(o_i, a_i^l)$ . Nous représentons le réseau de l'apprenant par un DQN recevant l'observation individuelle actuelle  $o_i \in \mathcal{O}_i$  avec  $\mathcal{O} = x_i \mathcal{O}_i$  l'ensemble des observations jointes et son action appliquée  $a_i^l$  comme entrée à chaque pas de temps (figure 1). La fonction d'observation  $O$  spécifie les probabilités d'observations  $\Pr(o|a, s')$ .

Chaque apprenant possède la même fonction de récompense mais reçoit un gain individuel  $g_i^l(t)$  à chaque pas de temps en fonction de son état factorisé actuel  $s_m \in \mathcal{S}$  avec  $\mathcal{S} = \chi_1 \times \dots \times \chi_{|\mathcal{X}|}$  l'espace d'état factorisé. Les contrôleurs perçoivent le gain cumulé  $g_j^c(t)$  de leurs apprenants comme unique perception. Ceci est motivé par les *différence de récompenses* [25] où chaque agent apprend d'une récompense modifiée plutôt que d'une récompense globale et réduit l'impact du problème de *multi-agent credit assignment*. Nous utilisons des *récompenses tronquées* [24]  $r = \{-1, 1\}$  sur l'évolution relative du gain entre l'état précédent et l'état actuel. Cela présente deux avantages : (i) une représentation générique de l'attribution de récompense, quelle que soit la fonction de récompense utilisée et (ii) leur impact sur la vitesse d'apprentissage. Un apprenant  $d_i^l$  enregistre son gain local le plus élevé  $gb_i^l$  obtenu pendant l'exploration, son gain local au pas de temps précédent  $g_i^l(t-1)$  et de son gain local actuel  $g_i^l(t)$ . La récompense

est égale à 1 si  $g_i^l(t) > g_i^l(t-1)$  ou si  $g_i^l(t) = g_i^l(t-1)$  et  $g_i^l(t) = gb_i^l$  et égal à -1 si  $g_i^l(t) \leq g_i^l(t-1)$  et  $g_i^l(t) < gb_i^l$ . L'ensemble des récompenses factorisées est  $\mathcal{R} = \{R_1, \dots, R_n\}$ .

Les contrôleurs utilisent un score de confiance  $t_s \in [-1; 1]$  pour suivre le comportement de leurs apprenants avec l'action appliquée jointe  $\mathbf{a}_j^a$ . Le contrôleur  $j$  met à jour le score de confiance de l'apprenant  $i$  en utilisant la fonction de score de confiance incrémentielle  $\tau : \mathcal{A}^l \times \mathcal{A}^l \times \mathcal{S} \times \mathcal{S} \rightarrow \{-0.1, 0, 0.1\}$ , l'action proposée  $a_i^s \in \mathcal{A}^l$ , l'action appliquée  $\mathbf{a}_j^a \in \mathcal{A}^l$ , le gain cumulé à l'itération précédente  $g_j^c(t-1) \in \mathcal{S}$  et le gain cumulatif de l'itération actuelle  $g_j^c(t) \in \mathcal{S}$ . Alors que les récompenses sont fournies par l'environnement et utilisées par les apprenants, l'attribution du score de confiance incombe aux contrôleurs. La confiance est modifiée chaque fois qu'un apprenant propose une action.  $\tau = 0$  si l'action proposée est choisie au hasard;  $\tau = 0.1$  si  $a_i^s = a_j^l$  et  $g_j^c(t) \geq g_j^c(t-1)$ ;  $\tau = -0.1$  si  $a_i^s = a_j^l$  et  $g_j^c(t) < g_j^c(t-1)$  ou si  $a_i^s \neq a_j^l$  et  $g_j^c(t) \geq g_j^c(t-1)$ . Le décrement  $a_i^s \neq a_j^l$  et  $g_j^c(t) \geq g_j^c(t-1)$  est plus subtile à comprendre, il traduit notre hypothèse que toute action augmentant le gain cumulé du contrôleur est la meilleure pour l'état actuel. Cependant, cela est faux si deux actions différentes ont le même effet sur les performances. Pour résoudre ce problème, nous avons élaboré un moyen d'identifier de telles actions similaires. Comme le contrôleur ne dispose pas d'informations claires sur l'état actuel, nous considérons plutôt les transitions d'actions offrant le meilleur gain. Ces transitions fournissent des informations sur l'état précédent et l'état actuel via l'observation du gain du contrôleur. Une transition est décrite par  $tr = (g_j^c(t-1), a_j^l(t-1), g_j^c(t))$  et  $T$  est une fonction de transition spécifiant les probabilités de transition d'état  $\Pr(s'|s, a)$ . Deux actions sont considérées comme similaires si, pour deux transitions,  $tr_1$  et  $tr_2$ ,  $g_j^c(t-1, tr_1) \neq g_j^c(t-1, tr_2)$  et  $g_j^c(t, tr_1) = g_j^c(t, tr_2)$  et  $g_j^c(t, tr_1) = gb_j^c$  où  $gb_j^c$  est le meilleur gain obtenu par le contrôleur  $j$ . Avec cette notion d'actions similaires, nous ajoutons  $\tau = 0.1$  si  $a_i^s \equiv a_i^l$  et  $g_j^c(t) = gb_j^c$  pour le score de confiance.

### 3.2 Actions de contrôle

Les actions disponibles pour les contrôleurs sont définies par  $\mathcal{A}^c = \{\text{Fusion}, \text{Separation}, \text{Do\_Nothing}\}$ . Les contrôleurs utilisent ces actions pour gérer la répartition des apprenants dans les clusters afin d'améliorer leurs gains cumulés. Lorsqu'une *Fusion* ou une *Sparation* est appliquée, le score de confiance de chaque apprenant dans les clusters concernés est réinitialisé et la configuration avant l'application de l'action de contrôle est enregistrée. Avec ces actions vient un moyen d'identifier la pertinence de la nouvelle distribution sur l'environnement. Lorsqu'une action est appliquée, les contrôleurs disposent de plusieurs épisodes pour évaluer l'efficacité de l'action de contrôle. Pendant cette période d'évaluation, aucune autre action de contrôle (fusion, séparation) ne sera acceptée sur le clus-

ter en cours d'évaluation pour éviter de fausser l'évaluation de l'action de contrôle. Après évaluation, les contrôleurs concernés choisissent de conserver cette configuration ou de revenir à la configuration précédente. Dans tous les cas, toutes les actions de contrôle sont à nouveau disponibles. Nous voulons minimiser le nombre de clusters afin de réduire l'impact de la non-stationnarité dans un contexte d'apprentissage  $Q$ -Learning indépendant.

Toutefois pour être considérée comme efficace, une action de contrôle doit avoir un impact positif sur le gain cumulé global.

L'action *Separation* sépare les apprenants avec des scores de confiance positifs de ceux ayant des scores de confiance négatifs, créant ainsi deux clusters : l'actuel et un nouveau, augmentant d'un le nombre de contrôleurs. Le nombre d'apprenants reste le même. Pour autoriser une séparation, nous introduisons une zone dite de *clivage* dans laquelle aucun score de confiance ne doit entrer pendant un nombre spécifique d'itérations. Si un score de confiance entre dans la zone interdite, le compteur d'itérations permettant d'appliquer l'action de séparation est réinitialisé. Pour un contrôleur  $d_1^c$  qui se divise en contrôleurs  $d_2^c$  et  $d_3^c$ , l'action de séparation est efficace si  $g_{a_1^c}^c(bs) < g_{a_2^c}^c(as) + g_{a_3^c}^c(as)$  est vrai, où  $bs$  correspond aux cas avant la séparation et  $as$  après la séparation. La séparation est considérée comme non pertinente si  $g_{a_1^c}^c(bs) \geq g_{a_2^c}^c(as) + g_{a_3^c}^c(as)$ . L'égalité des gains est jugée non pertinente car généralement il est plus efficace de minimiser le nombre de contrôleurs.

L'action de *Fusion* fusionne deux contrôleurs en vérifiant des séquences équivalentes d'actions d'environnement. Le cluster du contrôleur résultant contient les apprenants des deux contrôleurs confondus : leur nombre est ainsi réduit.

À chaque étape, les contrôleurs vérifient si les autres contrôleurs ont la même séquence d'actions environnementales auquel cas, ils appliquent une action de fusion. Lorsqu'un contrôleur reçoit son gain cumulé le plus élevé, il commence à rechercher des *séquences équivalentes* chez les autres contrôleurs. Les séquences sont dites équivalentes si elles sont composées des mêmes actions, même non ordonnées. Ainsi pour les trois actions environnementales  $a, b$  et  $c$ , les séquences  $abc$  et séquence  $acb$  sont équivalentes.

Cependant, considérer uniquement les séquences similaires ne suffit généralement pas pour réussir les fusions. Si nous considérons que les états initiaux des contrôleurs sont différents, ils subiront un changement dans leurs états. Ensuite, nous ne pouvons utiliser la fusion que si les états des contrôleurs sont homogènes, ce qui signifie que l'état factorisé a les mêmes variables d'état. Lorsque la simulation réinitialise les états de l'environnement, les contrôleurs se trouvent dans un état initial homogène. Ainsi, nous considérons des séquences d'action qui débutent de tout nouvel état aléatoire généré de l'environnement et se terminent quand le contrôleur pense avoir atteint son état optimal. Une deuxième contrainte est liée au fait que des sé-

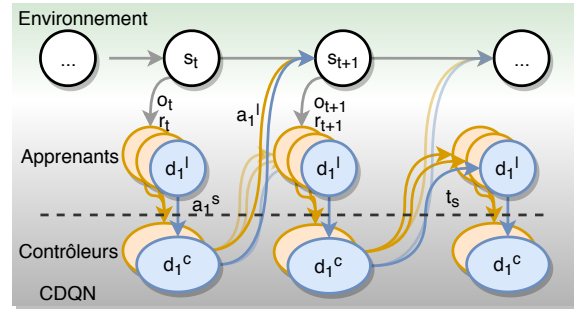


FIGURE 1 – Les apprenants sont séparés en deux clusters (bleu et orange). Un apprenant  $d_1^l$  observe  $o_t$  et suggère une action  $a_1^s$  à son contrôleur. Le contrôleur  $d_1^c$ , avec un système de vote, choisit une action  $a_1^l$  appliqué par les apprenants. À l'itération suivante,  $d_1^l$  observe  $o_{t+1}$  et reçoit la récompense  $r_{t+1}$ . Avec l'évolution de son gain,  $d_1^c$  modifie le score de confiance  $t_s$  des apprenants de son cluster.

quences équivalentes peuvent conduire à des états différents, en fonction du contexte. C'est pourquoi pour la fusion, nous nous limiterons aux séquences équivalentes qui conduisent aux mêmes états.

Dans notre contexte, certaines actions environnementales peuvent avoir des actions qui annulent une action précédente, appelée action opposée. Par exemple, si nous considérons nos trois actions environnementales précédentes et posons  $a$  le contraire de  $b$ . Alors les séquences  $abc$  et  $c$  débutant dans le même état initial conduiraient au même état mais seraient considérées comme différentes. Cela signifie que le contrôleur avec la séquence  $c$  a un cluster avec de meilleures actions suggérées que le contrôleur avec la séquence  $abc$ . Nous considérons donc que les deux groupes ne devraient pas fusionner.

Pour deux contrôleurs  $d_1^c$  et  $d_2^c$  fusionnés en  $d_3^c$ , l'action de fusion est pertinente si  $g_{a_1^c}^c(bf) + g_{a_2^c}^c(bf) \leq g_{a_3^c}^c(af)$  est vrai, où  $bf$  signifie avant la fusion et  $af$  pour après le Fusion. Cette règle a également pour conséquence de minimiser le nombre de contrôleurs car la fusion est considérée comme pertinente même si les gains avant et après Fusion sont égaux.

Pendant une période spécifiée, le contrôleur effectuant l'action observe si ses performances augmentent. À la fin de cette période, les contrôleurs utilisent les critères de pertinence introduits précédemment pour évaluer l'action et conserver la configuration actuelle ou revenir à la configuration précédant l'application de l'action de contrôle.

### 3.3 Algorithme du CDQN

Le CDQN lance une simulation avec la configuration initiale  $b_0$  décrivant l'état factorisé de chaque apprenant et sa répartition dans les différents clusters. À chaque pas de temps, un apprenant  $i$  reçoit une observation  $o_t^i$  et suggère l'action  $a_t^i$  produisant la Q-valeur la plus élevée avec une probabilité de  $1 - \epsilon$ . Chaque contrôleur reçoit des propositions de son cluster et choisit, avec son système de votes,

l'action  $a_t^l$  à appliquer sur l'environnement par les apprenants du cluster. À l'itération suivante, les apprenants reçoivent une observation  $o_{t+1}$  et une récompense  $r_t$ . La récompense est déterminée par l'évolution du gain local entre l'itération actuelle et l'itération précédente. Les contrôleurs reçoivent leurs gains cumulés et mettent à jour le score de confiance  $ts_{t+1}$  de chaque apprenant de leur cluster.

Si tous les apprenants ont un score de confiance de 1 pendant plusieurs épisodes, le cluster est alors considéré comme pertinent car chaque apprenant propose l'action qui améliore le gain cumulé. D'autre part à chaque pas de temps, l'apprenant stocke une expérience  $(o_t, a_t, r_t, o_{t+1})$  dans sa mémoire tampon.

Les contrôleurs ne peuvent appliquer qu'une et une seule action à chaque itération. Les agents apprenants sont pénalisés s'ils avaient suggéré une action différente de celle finalement choisie pour le cluster, et la mise à jour du score de confiance permet aux contrôleurs d'apprendre sur l'apport de chaque agent au gain du cluster.

Dans les environnements de vidéo-ludiques [13], les agents ont un état terminal lorsqu'ils apprennent. Il est souvent induit par la perte du jeu. Cela provoque un nouvel état initial et impose à l'agent d'accumuler de nouvelles connaissances via une exploration forcée. Dans notre contexte, l'environnement n'a pas d'état terminal. Si jamais c'est nécessaire, une idée serait d'ajouter un horizon fini avant lequel le contrôleur doit atteindre son état optimal, mais cela dépend de l'environnement. De plus, cela compromettrait l'autonomie de la simulation. Cependant, nous observons que sans états terminaux, lorsque les apprenants parviennent à une solution, ils l'exploitent et remplace la mémoire d'expériences par la meilleure solution, menant à l'oubli de la politique précédemment apprise. Par conséquent, nous introduisons un nouvel état factorisé aléatoire généré avec des variables d'états initiaux homogènes. Nous avons appelé un épisode le nombre d'itérations entre la réinitialisation de deux environnements. Au cours d'un épisode, chaque contrôleur tente de maximiser la somme des gains cumulés.

Le modèle CDQN est donc décrit par une version augmentée du Dec-POMDP factorisé  $\langle \mathcal{D}, \mathcal{S}, \mathcal{A}, T, \mathcal{O}, O, \mathcal{R}, t, \tau, \gamma \rangle$  où  $\gamma \in [0, 1)$  est un facteur de réduction.

## 4 Expérimentations

Le cadre principal de l'environnement est basé sur le modèle SmartGov [18] pour produire une simulation à plusieurs niveaux pour la régulation de la politique urbaine, dans laquelle des agents de personnalités différentes interagissent dans des environnements réalistes construits avec des données d'Open Street Map. Pour évaluer le CDQN, nous menons des expériences sur une instance de SmartGov et considérons les agents indépendants profonds (Independent Deep Q-Learning IDQL) comme base de comparaison.

SmartGov fournit au CDQN les états de l'environnement en tant qu'entrées pour les apprenants et en sortie, il reçoit

l'action jointe à appliquer sur l'environnement.

Nos expériences se situent dans un contexte de politique urbaine où les parties prenantes souhaitent améliorer des politiques urbaines. L'étude de cas concerne une politique de tarification des places de stationnement de la ville, pour des personnes se rendant chaque jour en ville, appelées navetteurs, en fonction de la personnalité et des préférences de stationnement des conducteurs. Chaque agent local contrôle le prix du stationnement d'un certain nombre de places d'un périmètre restreint, avec un front de rue composée de routes, d'emplacements payants et de bâtiments hétérogènes (maisons, bureaux, magasins, etc.) [18]. Les agents locaux appliquent des actions locales (augmenter, diminuer, etc., décrites ci-après), en fonction de la perception limitée de leur environnement. Ils doivent apprendre la politique qui optimise les récompenses. Dans ce contexte, les apprenants ont un espace d'états défini par des perceptions limitées et des variables d'état telles que  $\chi = \{\text{price}, \text{occupation}\}$ . La fonction de récompense choisie ici est le gain représenté par le produit du prix par place et du nombre de places occupées. Nous appelons *quartier* un espace où chaque navetteur a la même personnalité et les mêmes attentes. Nous notons  $C_1$  et  $C_2$  les ensembles de navetteurs avec des personnalités identiques. Sur la base de la configuration initiale du système, un cluster représente un ensemble d'apprenants ayant le même prix de stationnement initial. Le décideur politique a pour objectif de maximiser le gain cumulé global obtenu par le produit du prix et de l'occupation des emplacements, et il souhaite également que les clusters proposent un découpage de la politique de prix du stationnement en vigueur dans la ville. Le décideur juge de la pertinence du découpage de la politique proposée. Le CDQN gère efficacement ses clusters pour maximiser le gain cumulé sur ceux-ci.

Pour obtenir des résultats réalistes, nous avons utilisé un ensemble de trajectoires et de populations du monde réel pour essayer les actions de contrôle *Fusion* et *Separation* de manière indépendante dans un premier temps, puis ensembles. L'objectif principal de chaque expérience est de maximiser le gain global cumulé et de minimiser le nombre de clusters, les deux critères étant pris en compte par ordre lexicographique.

L'ensemble des actions disponibles pour l'apprenant  $i$  est  $\mathcal{A}_i^l = \{\nearrow, \searrow, =, +, -\}$  où  $\nearrow$  augmente le prix d'un emplacement,  $\searrow$  le baisse,  $+$  ajoute un emplacement,  $-$  en supprime un et  $=$  n'a aucun impact. Dans notre contexte, l'action  $-$  n'est jamais pertinente, mais elle est nécessaire pour évaluer la robustesse de la méthode actuelle et d'ajouter du bruit. Le nombre d'apprenants par clusters est déterminé par la topologie de la carte.

Nous avons évalué l'efficacité des actions *separation* et *fusion* sur trois scénarios différents (Tableau 1). Pour chaque scénario, nous avons comparé les performances du CDQN avec des IDQL dans lequel les agents IDQL sont des apprenants sans contrôleur.

Le scénario de fusion vise à identifier des séquences d'ac-

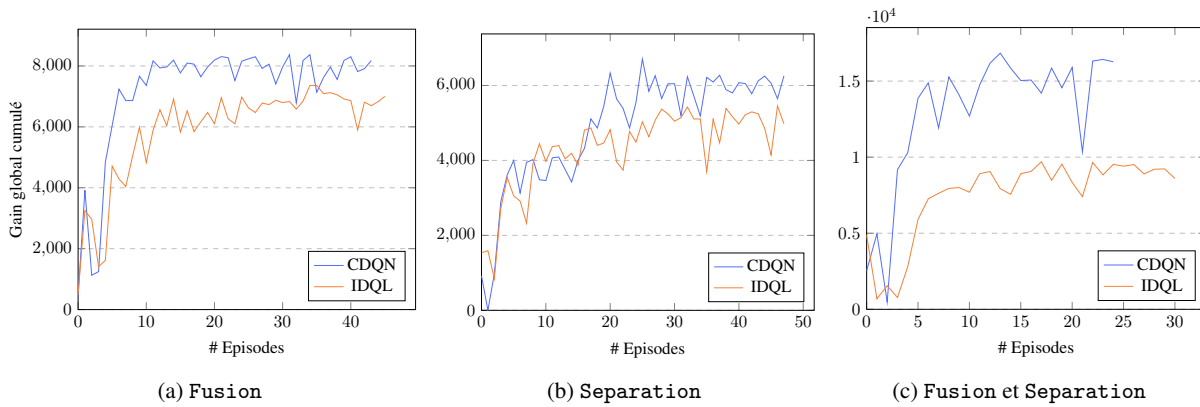


FIGURE 2 – Gain global cumulé pour l'IDQL et le CDQN sur les trois scénarios.

Actions	Contrôleurs	Apprenants	Navetteurs
Fusion	2	8/9	C1
Separation	1	17	C1, C2
Both	2	14/18	C1, C2

TABLE 1 – Apprenants par contrôleur pour chaque scénario et distribution des navetteurs.

tions équivalentes pour fusionner et minimiser efficacement le nombre de clusters. Les résultats attendus pour ce scénario sont un cluster avec le gain cumulé le plus élevé. Le scénario de séparation montre comment l'attribution de scores de confiance est utilisée pour identifier les quartiers et pour appliquer une séparation de cluster appropriée. Les résultats attendus pour ce scénario sont deux clusters avec un gain cumulé global plus élevé qu'avant la séparation du contrôleur.

L'objectif du scénario avec les deux actions est d'identifier les apprenants associés à chaque quartier et d'augmenter ainsi le gain cumulé en les regroupant efficacement. Le résultat attendu est une séparation par contrôleur, puis deux fusions, ou une fusion/séparation.

Le dernier scénario montre comment une combinaison des deux actions de contrôle permet d'augmenter le gain cumulé global et d'assurer une gestion efficace des clusters dans un cadre d'apprentissage par renforcement multi-agent utilisant des DQN indépendants et la réutilisation d'expériences.

Tous les réseaux sont entraînés à l'aide de l'algorithme ADAM [8] et de mini-lots de 64 échantillons. Le facteur de réduction a été défini par  $\gamma = 0.95$ . Nous avons utilisé une règle  $\epsilon$ -greedy lors de la formation, où  $\epsilon$  décroissait de 1,0 à 0,1 chaque fois que les apprenants proposaient une action avec  $\epsilon = \epsilon \times \epsilon_{decay}$  et où  $\epsilon_{decay} = 0.995$ .

## 5 Résultats

La figure 2a représente deux contrôleurs possédant les mêmes types de navetteurs. Seule l'action de fusion est dis-

ponible. Nous observons que le CDQN atteint un meilleur gain cumulé global que l'IDQL avec l'aide de la configuration initiale du cluster. La non-stationnarité est réduite, cela est dû aux apprenants d'un même cluster qui appliquent une action identique, et la vitesse d'apprentissage permet d'atteindre une stratégie individuelle optimale plus rapidement que les IDQL. Après un certain nombre d'itérations, les apprenants proposent des actions qui augmentent les gains. Lorsque les deux contrôleurs ont la même séquence un certain nombre de fois, ils fusionnent à l'épisode 5. La configuration est conservée après l'évaluation montrant que la fusion est efficace.

Dans la figure 2b, nous observons comment le score de confiance est utilisé pour proposer une séparation efficace des clusters. Au début de la simulation, CDQN et IDQL explorent l'environnement. Avant 15 épisodes, un cluster identifie un groupe d'apprenants en fonction de l'évolution du score de confiance. À ce moment, IDQL obtient de meilleurs résultats avec des actions individuelles plus faciles à appliquer. L'impact de la séparation après 15 épisodes est immédiat, car les apprenants proposent des actions plus pertinentes dans un cluster, ce qui permet à CDQN d'obtenir de meilleures performances que IDQL avec la nouvelle configuration des clusters. Certains agents apprennent que les actions appliquées par leur contrôleur  $j$  n'améliorent pas leur gain local, alors que d'autres apprennent que les actions appliquées augmentent leur gain local. L'action ( $\nearrow$ ) améliore le gain et donc le score de confiance des apprenants proposant cette action. Le phénomène s'inverse avec ceux qui proposent l'action ( $\searrow$ ). À l'aide des règles spécifiées, le contrôleur effectue une séparation à l'épisode 15 en conservant les apprenants avec un score de confiance positif et en créant un nouveau contrôleur qui reçoit un groupe d'apprenants dotés d'un score de confiance négatif. Ces scores sont réinitialisés une fois le nouveau cluster constitué. Désormais, les contrôleurs peuvent appliquer différentes actions environnementales sur l'environnement en même temps. Un contrôleur applique donc ( $\nearrow$ ) et l'autre ( $\searrow$ ) et la somme des gains devient plus élevée après la séparation, la configuration est

donc conservée.

Avec les deux scénarios, nous voyons que la fusion et la séparation sont pertinentes pour le contrôle dynamique et le regroupement des apprenants dans l'environnement. L'action de fusion utilise les gains et la séquence d'actions équivalentes, et l'action de séparation repose sur le score de confiance et l'apprentissage indépendant de chaque apprenant. Vers la fin de la simulation, les apprenants proposent des actions qui maintiennent le contrôleur dans un état stable et optimisent les gains à chaque étape. L'utilisation du score de confiance en plus de l'attribution de récompenses aide le système à apprendre des informations locales sur l'environnement.

Le dernier scénario (figure 2c) contient un total de 32 apprenants répartis dans deux groupes. Les contrôleurs fusionnent à l'épisode 4 et le cluster résultant et le cluster résultant obtient de meilleurs résultats. Après un certain temps d'exploration, le contrôleur obtient de meilleurs résultats et identifie les apprenants responsables de cette amélioration. L'écart entre les scores de confiance permet une nouvelle action de contrôle avec une séparation en deux contrôleurs qui, une fois validés, modifient l'environnement pour atteindre un nouveau gain global maximal.

Dans les trois cas, la convergence de la simulation est observée avec l'évolution des scores de confiance et du gain.

## 6 Discussion

Les résultats affichés dans la section précédente montrent que notre approche permet aux agents de se coordonner sans communication ni connaissance des autres agents, avec des résultats plus pertinents que pour IDQL : le CDQN converge vers un état stable avec un gain global plus élevé.

Un point important est que les apprenants de notre modèle n'ont pas à considérer l'influence des actions des autres apprenants avec la gestion des clusters par les contrôleurs. La réorganisation des clusters diminue l'espace d'action joint à prendre en compte dans le processus d'apprentissage. À l'exception des approches IDQL, nous avons, à notre connaissance, aucune approche existante similaire pour comparer nos résultats avec le contexte initial considéré.

Une perspective d'amélioration serait d'utiliser différentes combinaisons de vecteurs d'entrée pour étudier leur impact sur l'apprentissage des apprenants, par exemple avec un niveau moindre de connaissance (ici disposer de moins d'informations sur les personnalités et préférences des navetteurs). Comme chaque apprenant enregistre son modèle de réseau neuronal pendant la simulation, nous pourrions aussi envisager de réutiliser un apprentissage déjà effectué et modifier la configuration de l'environnement pour évaluer la capacité d'adaptation des apprenants avec la réutilisation d'expérience prioritaire [20]. Nous pourrions également essayer différents algorithmes d'apprentissage par renforcement en profond pour observer l'impact sur la vitesse de convergence et les configurations de clusters. En-

fin, une autre piste consisterait à configurer le système initial avec des états non homogènes.

## 7 Conclusion

Dans cet article, nous avons présenté le Clustered Deep Q-Network, une architecture d'apprentissage par renforcement multi-agents pour apprendre des politiques décentralisées dans des environnements partiellement observables sans communication entre les agents. Le modèle proposé distingue différents niveaux d'agents et de récompenses. La récompense de l'apprenant façonne les comportements individuels alors que la configuration des clusters évolue en fonction des scores de confiance des apprenants. Ce travail propose la recherche de configurations de clusters hiérarchiques d'agents avec des actions de fusion et de séparation, et l'apprentissage de leurs Q-valeurs individuelles. Une gestion efficace des agents de contrôle via des clusters d'apprenants contribue à augmenter les gains locaux et globaux. Les résultats sur un problème de politique tarifaire montrent comment la combinaison des scores de confiance et des récompenses individuelles permet un apprentissage efficace et une coordination émergente entre apprenants. Nous comparons ces résultats à un apprentissage profond individuel avec DQN afin de démontrer comment nous pouvons surmonter la non-stationnarité dans un apprentissage décentralisé. Toutes les applications de planification décentralisées ne pourront pas utiliser notre approche, du fait du choix d'une action unique par cluster pour réduire la non-stationnarité. Cependant, elle reste valable pour de nombreux domaines tels que l'élaboration de politiques, urbaines ou autre (santé, etc.), qui ont besoin d'apprendre vite et efficacement. Nous réfléchissons actuellement à la notion d'actions hiérarchiques, impactant la manière dont les apprenants pourraient agir indépendamment des autres apprenants du même groupe sans altérer le mécanisme d'attribution du score de confiance.

## 8 Remerciements

Ce projet est financé par la Région Auvergne-Rhône-Alpes (Contrat ADR ARC7). Il est le fruit d'une collaboration entre le LIRIS (Laboratoire d'InfoRmatique en Image et Systèmes d'information) à Lyon et NLE (NAVER LABS Europe), anciennement XRCE (Xerox ResearchCenter Europe) à Meylan.

## Références

- [1] Lucian Buşoniu, Robert Babuška, and Bart De Schutter. A comprehensive survey of multi-agent reinforcement learning. *IEEE Transactions on Systems Man and Cybernetics*, pages 156–172, 2008.
- [2] Yu-Han Chang, Tracey Ho, and Leslie Pack Kaelbling. All learning is local : Multi-agent learning in global reward games. *Advances in neural information processing systems*, pages 807–814, 2004.

- [3] Jakob N. Foerster, Yannis M. Assael, Nando de Freitas, and Shimon Whiteson. Learning to Communicate to Solve Riddles with Deep Distributed Recurrent Q-Networks. *arXiv preprint arXiv :1602.02672*, 2016.
- [4] Jakob N. Foerster, Gregory Farquhar, Triantafyllos Afouras, Nantas Nardelli, and Shimon Whiteson. Counterfactual multi-agent policy gradient. *arXiv preprint arXiv :1705.08926*, 2017.
- [5] Carlos Guestrin, Daphne Koller, and Ronald Parr. Multiagent planning with factored mdps. *Advances in Neural Information Processing Systems NIPS-14*, 2001.
- [6] Jayesh K. Gupta, Maxim Egorov, and Mykel Kochenderfer. Cooperative Multi-Agent Control Using Deep Reinforcement Learning. *International Conference on Autonomous Agents and Multiagent Systems*, 2017.
- [7] Matthew Hausknecht and Peter Stone. Deep Recurrent Q-Learning for Partially Observable MDPs. *Association for the Advancement of Artificial Intelligence*, 2015.
- [8] Diederik P. Kingma and Jimmy Lei Ba. ADAM : A Method for Stochastic Optimization. *International conference on machine learning*, 2015.
- [9] Guillaume Lample and Devendra Singh Chaplot. Playing FPS Games with Deep Reinforcement Learning. *Association for the Advancement of Artificial Intelligence*, pages 2140–2146, 2017.
- [10] Joel Z. Leibo, Vinicius Zambaldi, Marc Lanctot, Janusz Marecki, and Thore Graepel. Multi-agent reinforcement learning in sequential social dilemmas. *Proceedings of the 16th Conference on Autonomous Agents and MultiAgent Systems*, pages 464–473, 2016.
- [11] Long-Ji Lin. *Reinforcement Learning for Robots Using Neural Networks*. PhD thesis, Carnegie Mellon University, 1993.
- [12] Laetitia Matignon, Guillaume J. Laurent, and Nadine Le Fort-Piat. Independent reinforcement learners in cooperative markov games : a survey regarding coordination problems. *The Knowledge Engineering Review*, 27(1) :1–31, 2012.
- [13] Volodymyr Mnih, Adrià Puigdomènech Badia, Mehdi Mirza, Alex Graves, Tim Harley, Timothy P. Lillicrap, David Silver, and Koray Kavukcuoglu. Asynchronous Methods for Deep Reinforcement Learning. *International conference on machine learning*, pages 1928–1937, 2016.
- [14] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Alex Graves, Ioannis Antonoglou, Daan Wierstra, and Martin Riedmiller. Playing atari with deep reinforcement learning. *arXiv preprint arXiv :1312.5602*, 2013.
- [15] Volodymyr Mnih, Koray Kavukcuoglu, David Silver, Andrei A. Rusu, Joel Veness, Marc G. Bellemare, Alex Graves, Martin Riedmiller, Andreas K. Fidjeland, Georg Ostrovski, Stig Petersen, Charles Beattie, Amir Sadik, Ioannis Antonoglou, Helen King, Dharmashan Kumaran, Daan Wierstra, Shane Legg, and Demis Hassabis. Human-Level Control Through Deep Reinforcement Learning. *Nature*, 518(7540) :529, 2015.
- [16] Frans A. Oliehoek, Matthijs T.J. Spaan, Shimon Whiteson, and Nikos Vlassis. Exploiting locality of interaction in factored dec-pomdps. *Proceedings of the 7th International Conference on Autonomous Agents and Multiagent Systems*, 2008.
- [17] Frans A. Oliehoek, Shimon Whiteson, and Matthijs T.J. Spaan. Approximate solutions for factored dec-pomdps with many agents. *Proceedings of the 12th International Conference on Autonomous Agents and Multiagent Systems*, May 2013.
- [18] Simon Pageaud, Véronique Deslandres, Vassilissa Lehoux, and Salima Hassas. Couplage de simulations multi-agents pour la conception de politiques urbaines. *26èmes Journées Francophones sur les Systèmes Multi-Agents*, 2018.
- [19] Peng Peng, Quan Yuan, Ying Wen, Yaodong Yang, Zhenkun Tang, Haitao Long, and Jun Wang. Multiagent bidirectionally-coordinated nets for learning to play starcraft combat games. *arXiv preprint arXiv :1703.10069*, 2017.
- [20] Tom Schaul, John Quan, Ioannis Antonoglou, and David Silver. Prioritized Experience Replay. *International Conference on Learning Representations*, 2016.
- [21] Yoav Shoham and Kevin Leyton-Brown. *Multiagent Systems : Algorithmic, Game-Theoretic, and Logical Foundations*. Cambridge University Press, 2008.
- [22] Ardi Tampuu, Tambet Matiisen, Dorian Kodelja, Ilya Kuzovkin, Kristjan Korjus, Juhan Aru, Jaan Aru, and Raul Vicente. Multiagent Cooperation and Competition with Deep Reinforcement Learning. *PLoS ONE*, 12(4), April 2017.
- [23] Ming Tan. Multi-agent reinforcement learning : Independent vs. cooperative agents. *Proceedings of the tenth international conference on machine learning*, pages 330–337, 1993.
- [24] Hado van Hasselt, Arthur Guez, Matteo Hessel, Volodymyr Mnih, and David Silver. Learning Values Across many Orders of Magnitude. *Advances in Neural Information Processing Systems*, 2016.
- [25] David H. Wolpert and Kagan Tumer. Optimal payoff functions for members of collectives. *Modeling complexity in economic and social systems*, pages 355–369, 2002.



# Apprentissage de relations floues pour l'annotation sémantique expliquée avec peu de données

Régis Pierrard<sup>1,2</sup>Jean-Philippe Poli<sup>1</sup>Céline Hudelot<sup>2</sup><sup>1</sup> CEA, LIST, 91191 Gif-sur-Yvette cedex, France.<sup>2</sup> Mathématiques et Informatique pour la Complexité et les Systèmes, CentraleSupélec, Université Paris-Saclay, 91190, Gif-sur-Yvette, France.

{regis.pierrard, jean-philippe.poli}@cea.fr, celine.hudelot@centralesupelec.fr

## Résumé

Malgré son récent succès, l'apprentissage profond semble loin d'égaliser certaines capacités humaines comme l'apprentissage à partir de peu d'exemples, le raisonnement ou encore l'explicabilité. Dans cet article, nous mettons l'accent sur l'annotation d'objets dans des images et présentons un cadre de raisonnement s'appuyant sur l'apprentissage de relations floues et permettant de générer des explications. Les relations les plus pertinentes sont induites puis transposées en contraintes afin de définir un problème de satisfaction de contraintes floues. La résolution de ce problème nous permet d'annoter l'image et de générer des explications à l'aide des contraintes utilisées. Nous testons notre approche sur un jeu de données public avec deux objectifs : effectuer une annotation d'organes et fournir une explication à chaque annotation. Nous montrons que le modèle que nous obtenons peut générer des explications tout en atteignant des performances élevées malgré un jeu de données pouvant être constitué de seulement deux exemples.

## Mots Clef

Explicabilité, annotation, logique floue, peu de données.

## Abstract

Despite the recent successes of deep learning, such models are still far from some human abilities like learning from few examples, reasoning and explaining their decisions. In this paper, we focus on object annotation in images and we introduce a reasoning framework that is based on learning fuzzy relations on a dataset for generating explanations. Given a catalogue of relations, it efficiently induces the most relevant relations and combines them for building constraints in order to both solve the target task and generate explanations. We test our approach on a publicly available dataset on which the goal was both to perform multiple organ annotation and to provide explanations. We show that our model can generate explanations and achieve high performance despite being trained on a small dataset containing as few as a couple of examples.

## Keywords

Explainability, annotation, fuzzy logic, few-data learning.

## 1 Introduction

Au cours de ces dernières années, l'explicabilité est devenue un enjeu majeur pour les intelligences artificielles [11, 16]. Cela fait notamment écho à la popularité des réseaux de neurones profonds, qui peuvent atteindre des performances très élevées dans de nombreuses tâches mais ne sont pas adaptés à la génération d'explication [13, 26]. Être capable d'expliquer les résultats produits par les intelligences artificielles est utile non seulement pour comprendre leur raisonnement, mais aussi pour les rendre plus crédibles et dignes de confiance aux yeux des utilisateurs [32]. Dans des domaines en lien direct avec l'être humain tels que l'analyse d'images médicales [24], on ne peut pas prendre de décisions en faisant une confiance aveugle à un modèle car les conséquences peuvent être désastreuses.

Plusieurs définitions d'*interprétabilité* et d'*explicabilité* ont été données dans la littérature [9, 15, 23, 28], mais aucun consensus ne se dégage et ces deux notions sont mêmes parfois utilisées de manière interchangeable. Globalement, il émerge que l'interprétabilité est la capacité à présenter des informations sur le fonctionnement du système en des termes intelligibles, alors que l'explicabilité est la capacité à décrire de manière précise et logique le fonctionnement d'un système. Dans cet article, nous nous focalisons sur l'explicabilité et donc sur la restitution du raisonnement effectué par notre modèle. Pour obtenir des explications, une première famille de méthodes consiste à apprendre une approximation locale et interprétable autour de la prédiction fournie par un modèle de type boîte noire [25, 32]. Ces approches peuvent traiter tout type de modèle et sont donc adaptées aux réseaux de neurones profonds. Cependant, bien qu'elles soient capables d'extraire les caractéristiques principales qui ont servi à générer une prédiction, elles ne peuvent pas exactement reconstituer le raisonnement effectué par le modèle cible. La seconde famille d'approches repose sur l'utilisation de modèles qui sont propices à la gé-

nération d'explications, tels que les arbres de décisions, les règles de décision ou l'obtention d'un modèle explicable à partir d'un modèle de type boîte noire par distillation [17]. Leur avantage principal est qu'ils permettent de facilement reproduire le raisonnement qui a mené à un résultat donné, ce qui est très utile pour la génération d'explications. Cependant, ces modèles risquent de ne pas être aussi performants que des modèles de type boîte noire car l'explicabilité a souvent un coût. En effet, il existe un compromis entre performance et explicabilité [16]. Dans ce papier, nous proposons une approche basée sur cette seconde famille de méthodes qui compense ce compromis en n'apprenant que sur une faible quantité de données étiquetées dont l'acquisition est coûteuse. Cela repose notamment sur des études cognitives sur l'interprétation humaine des images : (1) l'importance des relations spatiales et contextuelles en reconnaissance d'objets et de scènes [1], (2) la capacité des humains à n'avoir besoin que de quelques exemples pour apprendre à accomplir une tâche [22, 35]. Ce second point est notamment abordé par les zero-shot [20], one-shot et few-shot learning [12]. Ils visent à effectuer cela en utilisant les précédents exemples dans la base d'apprentissage ou les représentations disponibles. Cependant, ces méthodes nécessitent toujours d'avoir des informations entre les classes connues et les classes inconnues [37]. Notre objectif est de proposer une nouvelle approche qui permet d'apprendre à raisonner et de générer à la fois des annotations et des explications à partir d'un faible nombre d'exemples d'apprentissage. Elle s'appuie notamment sur l'utilisation de relations floues, qui permettent de prendre en compte à la fois des informations quantitatives et des informations qualitatives. Pour un exemple donné, le système cherche les régions de l'image qui satisfont au mieux les relations entre les objets d'intérêt. Nous modélisons cela sous la forme d'un problème de satisfaction de contraintes. Dans la section 3, nous décrivons la démarche globale qui consiste en trois étapes principales : l'évaluation des relations, l'extraction des relations les plus pertinentes et la génération des contraintes pour résoudre un problème de satisfaction de contraintes et produire des explications. Dans la section 4, l'approche proposée est évaluée sur un exemple d'annotation d'images médicales. Ceci est un bon exemple de raisonnement spatial car la reconnaissance des différents organes repose essentiellement sur leur disposition spatiale. De plus, l'utilisation d'un jeu de données médicales présente plusieurs défis, tels que sa faible taille et l'importance des explications dans ce domaine. Nous avons testé et comparé notre modèle à l'état de l'art et nous avons montré que notre approche est capable d'obtenir de bonnes performances et de générer des explications malgré un faible nombre d'exemples d'apprentissage.

## 2 Rappels

### 2.1 Logique floue

La logique floue et la théorie des ensembles flous ont été introduits par Zadeh [39]. Il s'agit d'une extension de la lo-

gique Booléenne qui permet notamment de gérer l'imprécision des données. Alors qu'une variable est soit vraie soit fausse en logique Booléenne, elle peut prendre n'importe quelle valeur entre 0 (fausse) et 1 (vraie) en logique floue. Cette gamme de valeurs, qu'on appelle degrés, sert à gérer l'imprécision. Dans un univers  $A$ , un ensemble flou  $F$  est défini par une fonction d'appartenance  $\mu_F : A \rightarrow [0; 1]$ . Cette fonction, intitulée *fonction d'appartenance de  $F$* , représente dans quelle mesure chaque  $a \in A$  appartient à  $F$ . Si  $F$  est un ensemble non-flou, alors  $\mu_F(a)$  vaut soit 0 quand  $a$  n'appartient pas à  $F$ , soit 1 si  $a$  appartient à  $F$ .

Il est également possible d'exprimer des relations entre deux univers différents. Pour deux univers  $A$  et  $B$ , une relation floue dyadique  $\mathcal{R}$  est un ensemble flou défini par une fonction d'appartenance  $\mu_{\mathcal{R}} : A \times B \rightarrow [0, 1]$ . Cette fonction attribue à chaque paire  $(a, b) \in A \times B$  un degré d'appartenance  $\mathcal{R}$ . Les relations floues  $n$ -aire sont définies de la même façon. Un ensemble flou peut caractériser une expression du langage naturel lorsqu'il est associé à une description linguistique, ce qui est courant en logique floue.

### 2.2 Problème de satisfaction de contraintes floues

Un problème de satisfaction de contraintes (CSP) est un problème dans lequel l'objectif est d'attribuer des valeurs à un ensemble de variables qui doivent vérifier un ensemble de contraintes. [10] en présente une extension au cadre de la logique floue afin de pouvoir gérer des paramètres et des contraintes flous lors de la résolution de problèmes de satisfaction de contraintes floues (FCSP). Un tel problème est défini par un ensemble de variables  $X$ , un ensemble de domaines  $D$  et un ensemble de contraintes floues  $C$ . Par exemple, dans le cadre de l'annotation d'images,  $X$  correspond aux étiquettes que l'on souhaite attribuer aux objets de l'image,  $D$  représente les différentes régions à annoter et  $C$  est un ensemble de relations floues liant ces objets. Les FCSP nous intéressent tout particulièrement car ils nous permettent d'effectuer l'annotation (en résolvant le problème) ainsi que de générer des explications à partir des contraintes utilisées.

Pour résoudre un tel problème, on utilise un algorithme de backtracking pour parcourir l'ensemble des solutions possibles. Avant chaque assignation d'une valeur à une variable, l'algorithme FAC-3 [10, 36] est utilisé pour maintenir la cohérence de l'ensemble des valeurs possibles. Cette opération de propagation des contraintes permet de réduire la taille de l'espace de recherche. À la fin, la solution retenue est celle qui vérifie au mieux les contraintes.

## 3 Approche proposée

Dans cette partie, nous décrivons notre approche dont l'objectif est d'annoter les régions d'intérêt et de fournir une explication pour chaque annotation. Elle se divise en trois étapes : l'évaluation des relations entre les régions d'intérêt (Sec. 3.1), l'apprentissage des relations les plus pertinentes (Sec. 3.2) et la résolution d'un FCSP pour trouver les ré-

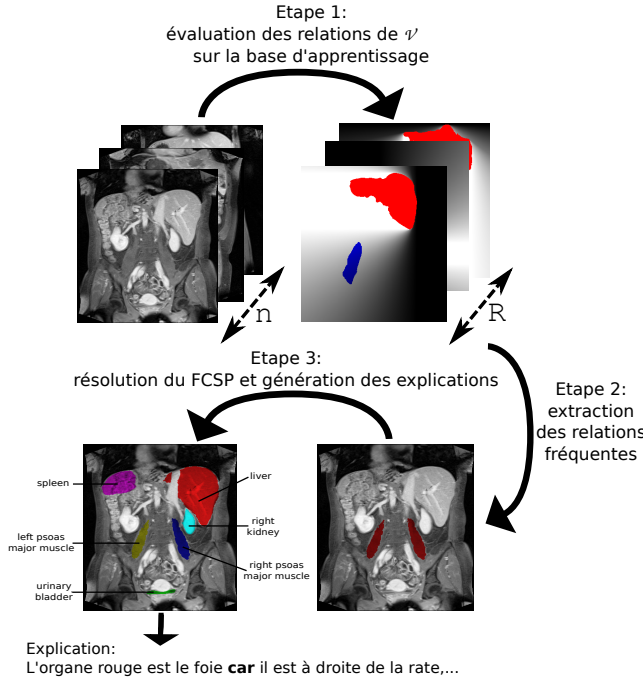


FIGURE 1 – Représentation des trois étapes de l'approche proposée. À l'étape 1, les  $n_r$  relations floues du vocabulaire  $\mathcal{V}$  sont évaluées sur le jeu de données  $\mathcal{T}$  composé de  $n$  images. Ensuite, les relations les plus fréquentes sont extraites sur la base d'apprentissage  $\mathcal{T}_{app}$  et sont utilisées pour déterminer les contraintes de notre problème. Enfin, un FCSP est défini et résolu afin d'étiqueter les objets cibles pour chaque image de  $\mathcal{T}_{test}$ . Une explication basée sur les contraintes utilisées est fournie pour chaque étiquette.

gions qui satisfont au mieux les relations apprises tout en générant des explications à partir des contraintes (Sec. 3.3). L'approche globale est représentée sur la Figure 1.

### 3.1 Étape 1 : évaluation des relations

L'objectif de cette première étape est d'évaluer différentes relations entre les régions d'intérêt (les organes par exemple) afin de pouvoir plus tard (dans l'étape suivante) déterminer celles qui sont les plus pertinentes.

Soit un ensemble d'apprentissage  $\mathcal{T}_{train}$  composé de  $n$  images  $\{\mathbf{i}_1, \dots, \mathbf{i}_n\}$  et un ensemble d'étiquettes  $\mathcal{Y}$  composé de  $N$  éléments  $\{y_1, \dots, y_N\}$  tels que chaque image  $\mathbf{i} \in \mathcal{T}_{train}$  est divisée en un ensemble  $\mathcal{O}$  de  $N$  régions d'intérêt  $\mathcal{O} = \{o_{i,1}, \dots, o_{i,K}\}$  qui sont associées aux étiquettes par la fonction suivante :

$$f : \mathcal{Y} \rightarrow \mathcal{O} \quad (1)$$

$$y_i \mapsto o_j$$

Soit  $\mathcal{V} = \{\mathcal{R}_1, \dots, \mathcal{R}_R\}$  un ensemble de  $R$  relations. Nous appelons cet ensemble un *vocabulaire*. Il est défini par un expert et contient des relations qui sont potentiellement d'intérêt. Par exemple, il peut s'agir de relations directionnelles, comme à *gauche de*, ou de relations de distance,

comme *près de*. Un vocabulaire plus riche devrait rendre le système plus expressif et améliorer les annotations et explications produites. Les relations de notre vocabulaire  $\mathcal{V}$  sont automatiquement évaluées sur les régions d'intérêt de chaque image de  $\mathcal{T}_{train}$ . La manière dont elles sont évaluées dépend de leur définition (cf. Sec. 4.2).

Pour chaque relation  $\mathcal{R} \in \mathcal{V}$ , nous appelons  $\alpha(\mathcal{R})$  son arité.  $\mathcal{R}$  est évaluée pour chaque  $\alpha(\mathcal{R})$ -uplet de régions d'intérêt. Il est important de distinguer le vocabulaire  $\mathcal{V}$  de l'ensemble  $\mathcal{V}_y$  des relations à évaluer entre les régions d'intérêt. Par exemple, pour une relation binaire  $\mathcal{R} \in \mathcal{V}$ , il faut évaluer l'ensemble  $\{\mathcal{R}(f(y), f(y')) \mid y, y' \in \mathcal{Y}, y \neq y'\}$ . Ainsi, le nombre d'évaluations à effectuer est :

$$\text{Card}(\mathcal{V}_y) = \sum_{j=1}^R \frac{N!}{(N - \alpha(\mathcal{R}_j))!} \quad (2)$$

À la fin de cette étape, nous avons donc une base de données  $\mathcal{B} : \mathcal{T}_{train} \times \mathcal{V}_y \rightarrow [0; 1]$  de taille  $n \times \text{Card}(\mathcal{V}_y)$ .

### 3.2 Étape 2 : apprentissage des relations pertinentes

Lors de cette étape, l'objectif est d'extraire les relations les plus pertinentes parmi celles évaluées à l'étape précédente. Pour une étiquette  $y \in \mathcal{Y}$ , nous faisons le postulat que les relations pertinentes impliquant les régions étiquetées comme  $y$  sont les plus fréquentes car elles devraient être vérifiées par la plupart, si ce n'est tous, des exemples associés à cette classe. Ainsi, l'apprentissage des relations les plus pertinentes revient à extraire les relations les plus fréquentes. Cette opération est effectuée classe par classe, il s'agit d'une stratégie *one vs all*. En effet, les relations pertinentes pour une classe d'objet ne le sont pas forcément pour une autre classe. Au sein d'une même classe, les exemples devraient être fortement corrélés entre eux. C'est pourquoi nous utilisons un algorithme d'extraction de motifs fréquents qui tire profit de cette propriété [30].

Un sous-ensemble  $J$  de relations est un ensemble appartenant à  $2^{\mathcal{V}_y}$ . Soit  $J = \{r_1, \dots, r_p \mid r_j \in \mathcal{V}_y, 1 \leq j \leq p \leq \text{Card}(\mathcal{V}_y)\}$ . La fréquence de  $J$  dans la base de données  $\mathcal{B}$  s'appelle le *support* et est défini par :

$$\text{support}(J) = \frac{1}{n} \sum_{k=1}^n \min(\mathcal{B}(\mathbf{i}_k, r_1), \dots, \mathcal{B}(\mathbf{i}_k, r_p)) \quad (3)$$

Étant donné un seuil,  $J$  est dit *fréquent* si son support est supérieur à ce seuil. L'algorithme utilisé s'appuie sur un opérateur de fermeture défini comme  $h : 2^{\mathcal{V}_y} \rightarrow 2^{\mathcal{V}_y}$ . Une première phase consiste à déterminer l'ensemble des sous-ensembles de relations vérifiant  $h(J) = J$ . De tels sous-ensembles de relations sont dits *fermés*. Ensuite, parmi les ensembles fermés, seuls ceux qui sont fréquents sont conservés. Enfin, l'ensemble des sous-ensembles fermés et fréquents permet d'engendrer l'ensemble des sous-ensembles fréquents de relations.

Puisque nous suivons une approche *one-vs-all*, chaque classe d'objets  $y$  a son propre seuil de fréquence  $S_y$  dont la

valeur est déterminée lors d’une phase de validation. La valeur de ce seuil a une influence directe sur le nombre d’ensembles de relations fréquents extraits. Si elle est trop élevée, alors pas ou peu d’ensembles de relations seront considérés comme fréquents. On risque alors de ne pas pouvoir discriminer les différentes classes et on se retrouve donc dans une situation d’*underfitting*. Au contraire, si cette valeur est trop basse, alors des ensembles de relations non pertinents seront appris : il s’agit d’*overfitting*.

À la fin de cette étape, pour chaque étiquette  $y \in \mathcal{Y}$ , nous avons un ensemble  $F_y$  de sous-ensembles fréquents de relations tel que  $F_y \subseteq 2^{2^{\mathcal{V}}}$ .

### 3.3 Étape 3 : résolution du FCSP et génération des explications

Pour un exemple de test  $i$ , on peut obtenir par segmentation un ensemble de régions potentiellement d’intérêt. L’objectif de cette étape est de déterminer les étiquettes des régions qui satisfont au mieux les relations apprises à l’étape précédente. On modélise cela par un FCSP. De plus, comme les relations sont associées à une description linguistique, on peut générer une explication pour chaque annotation.

Pour chaque étiquette  $y \in \mathcal{Y}$ , nous avons obtenu à l’issue de l’étape précédente un ensemble  $F_y$ . Soit  $F_y^{\max}$  un ensemble défini de la façon suivante :

$$F_y^{\max} = \{z \in F_y \mid \text{Card}(z) = \max_{v \in F_y} (\text{Card}(v))\} \quad (4)$$

$F_y^{\max}$  correspond donc à l’ensemble des sous-ensembles fréquents de relations de taille maximale. Chaque relation  $\mathcal{R}(f(y), f(y'))$  présente dans les sous-ensembles de  $\bigcup_{y \in \mathcal{Y}} F_y^{\max}$  est directement convertie en une contrainte floue

$c_{\mathcal{R}}(f(y), f(y'))$ . Nous pouvons donc désormais construire un modèle défini par les contraintes qui ont été apprises ainsi que par les valeurs des seuils de fréquence. Reposant sur l’apprentissage de symboles (les relations) et n’ayant pas recours à la minimisation d’une fonction de perte, notre modèle est adapté à des jeux de données de faible taille.

Après segmentation, l’exemple de test  $i$  est divisé en  $K$  régions potentiellement d’intérêt  $\{o_{i,1}, \dots, o_{i,K}\}$ . Le FCSP que nous cherchons à résoudre est donc le suivant :

$$X = \{x_y \mid y \in \mathcal{Y}\} \quad (5)$$

$$D = \{D_y \mid y \in \mathcal{Y} \text{ and } D_y = \{o_{i,1}, \dots, o_{i,K}\}\} \quad (6)$$

$$C = \{c_{\mathcal{R}}(x_y, x_{y'}) \mid \mathcal{R}(f(y), f(y')) \in U \text{ such as } U \subseteq \bigcup_{y \in \mathcal{Y}} F_y^{\max}\} \quad (7)$$

Chaque élément  $D_y$  de  $D$  est l’ensemble de toutes les régions qui pourraient être associées à la variable  $x_y$ . Au début, tous les  $D_y$  sont identiques. Ils sont ensuite modifiés par l’algorithme FAC-3 qui écarte les solutions incohérentes. Chaque contrainte de  $C$  ayant été évaluée, le

problème peut alors être résolu et les étiquettes correspondant à chaque variable déterminées. Ensuite, pour chaque variable  $x_y \in X$ , une explication est générée en utilisant les contraintes obtenues à partir de  $F_y^{\max}$ . Par exemple, la contrainte  $c_{\mathcal{R}_{\text{à gauche de}}}(x_W, x_Z)$  engendre “ $W$  est à gauche de  $Z$ ”. Ainsi, l’utilisation de contraintes générées à partir de  $F_y^{\max}$  permet d’exprimer une explication sous la forme : “classe  $y$  CAR cause $_1, \dots, \text{cause}_n$ ”. Pour une classe d’objets donnée, la contrainte la moins satisfaite nous donne un facteur de certitude pour modérer l’explication, e.g. “L’annotation de cet organe est *probablement* le foie...”. Les contraintes apprises et ce facteur de certitude sont envoyés à une *surface realiser* comme simpleNLG [14] afin de les agréger en une phrase syntaxiquement correcte.

L’évaluation de la qualité d’une explication est une tâche délicate. Une bonne explication est une explication cohérente [28], simple et générique [31], utile et pertinente [27]. Ces différents critères sont difficiles à évaluer de manière automatique et ils dépendent de la connaissance et des attentes de l’utilisateur. Trois types de solution sont proposés dans [9] : faire évaluer les explications par un expert, poser plusieurs questions simples à un groupe de personnes n’étant pas expert dans le domaine ciblé ou utiliser un modèle proxy dont l’explicabilité a déjà été évaluée pour juger l’explicabilité du modèle en cours d’évaluation.

## 4 Expériences

Dans cette partie, nous décrivons les expériences qui ont été menées sur un jeu d’images médicales. L’objectif est d’effectuer une annotation d’organes justifiée en apprenant un modèle à partir d’un faible nombre de données. Tandis que la détection d’organes est un sujet qui a été régulièrement abordé dans la littérature [34, 8, 29, 21], l’annotation d’organes n’a été traitée que dans [38]. Cette dernière méthode consiste à trouver dans le jeu de données des images qui ont des caractéristiques visuelles similaires à l’image étudiée et à l’étiqueter de la même manière que celles-ci. [21] aborde le problème de la détection d’organes sur des coupes abdominales. Leur méthode s’appuie sur des règles spatiales floues. Cependant, ces règles sont trop spécifiques et ne sont pas adaptées à d’autres jeux de données. La classification d’organes a été traitée par [33] à l’aide d’un réseau de neurones convolutionnel. Leur méthode repose également sur des techniques d’augmentation de données permettant de répondre au problème de la faible taille du jeu de données utilisé.

### 4.1 Jeu de données

Partant du constat qu’il n’existe actuellement pas de jeu de données dédié à l’évaluation de l’explicabilité des modèles, nous avons utilisé un jeu de données initialement créé pour évaluer des méthodes de segmentation supervisée. Il s’intitule *Anatomy3* et a été présenté dans [18]. Il contient 391 images provenant de deux modalités différentes (CT et IRM) ainsi que les segments correspondant à certains organes de ces images. Certaines images repré-

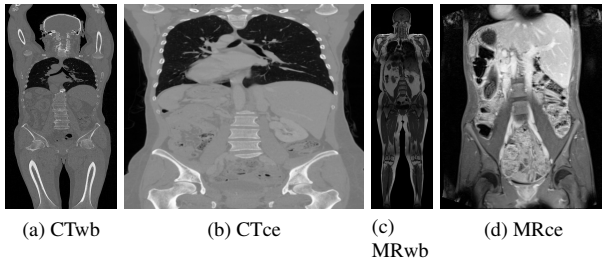


FIGURE 2 – Exemples correspondant aux quatre types d’images composant le jeu de données utilisé.

sentent le corps entier (CTwb et MRwb) et les autres l’abdomen (CTce et MRce). Ce sont des images à trois dimensions obtenues par superposition de tranches en deux dimensions. Puisque nous travaillons sur des données à deux dimensions, nous avons sélectionné une tranche 2D par patient afin de construire un jeu d’images 2D. La figure 2 montre quatre tranches, chacune correspondant à un des quatre types d’images (CTwb, MRwb, CTce et MRce).

L’ensemble  $\mathcal{Y}$  d’organes auquel nous nous intéressons est composé de la *foie*, de la *rate*, de la *vessie*, des *reins gauche et droit*, des *poumons gauche et droit* et des *muscles psoas gauche et droit*. Nous avons conservé toutes les images contenant ces neuf organes et leurs segments correspondants, ce qui donne un jeu de 35 images et 315 segments.

## 4.2 Paramètres expérimentaux

**Apprentissage du modèle.** Le modèle que nous construisons repose sur l’extraction des sous-ensembles de relations fréquents. Il y a autant d’hyperparamètres que d’organes. Ils correspondent tous au seuil utilisé pour juger la fréquence d’une relation. Afin d’optimiser la recherche de ces hyperparamètres, nous avons utilisé une validation croisée imbriquée [5] : (1) une validation croisée externe est réalisée pour obtenir un ensemble d’apprentissage et un ensemble de test à chaque itération, (2) une validation croisée interne est menée sur chaque base d’apprentissage de la validation croisée externe pour obtenir un ensemble d’apprentissage interne et un ensemble de validation permettant d’ajuster la valeur des hyperparamètres. Cela permet d’obtenir une estimation de l’erreur non-biaisée par la construction de l’ensemble de validation.

Dans la boucle interne, l’optimisation des hyperparamètres est effectuée par optimisation bayésienne sur 20 itérations avec un processus Gaussien comme a priori. La fonction d’acquisition utilisée ici est l’*expected improvement* [19].

**Relations.** De nombreuses relations spatiales floues ont été étudiées dans la littérature [4]. Dans nos expériences, nous utilisons des relations directionnelles, de distance et de symétrie. Les relations directionnelles et de distance [2, 3] sont obtenues en générant un paysage flou et en utilisant une approche de type reconnaissance de forme floue [6]. Comme on peut le voir sur la figure 3, le pay-

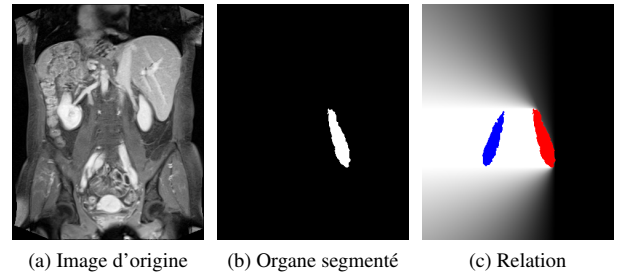


FIGURE 3 – (Affichage optimal en couleur) Exemple d’évaluation d’une relation. L’objectif est de déterminer la relation *l’organe bleu à gauche de l’organe rouge*. Étant donné une image en entrée (3a), un organe segmenté (3b) est considéré comme l’objet de référence. Cet organe est utilisé pour générer un paysage flou (3c) qui représente le degré d’appartenance de chaque pixel à la relation *à gauche de l’organe rouge*. La relation est déterminée en évaluant le degré d’intersection entre le paysage flou et l’organe bleu.

sage flou est obtenu par la dilatation floue d’un objet de référence par un élément structurant dont la définition détermine le type de relation à évaluer. Soit  $S$  l’espace de l’image. Soit  $A$  un objet de référence dans  $S$  et  $\mu_{A,\mathcal{R}}$  la fonction d’appartenance associée au paysage flou représentant la relation  $\mathcal{R}$  dont l’objet de référence est  $A$ . Soit  $\mu_B$  la fonction d’appartenance associée à un objet  $B$  dans  $S$ . La relation  $\mathcal{R}$  entre  $A$  et  $B$  est alors obtenue par évaluation du degré d’intersection flou  $\mu_{int}$  [4] entre  $\mu_{A,\mathcal{R}}$  et  $\mu_B$  :

$$\mu_{int}(\mu_{A,\mathcal{R}}, \mu_B) = \frac{\sum_{x \in S} \min(\mu_{A,\mathcal{R}}(x), \mu_B(x))}{\min\left(\sum_{x \in S} \mu_{A,\mathcal{R}}(x), \sum_{x \in S} \mu_B(x)\right)} \quad (8)$$

Sur la figure 3, la relation  $\mathcal{R}$  est *à gauche de*, l’objet de référence  $A$  est l’organe rouge et l’objet  $B$  est l’organe bleu. Afin d’obtenir un vocabulaire fini, nous n’avons utilisé que certaines valeurs pour les paramètres de ces relations.

La relation de symétrie axiale [7] que nous utilisons repose sur la détermination de l’axe qui maximise une mesure de symétrie entre deux organes. Puisque cette mesure n’est pas différentiable, ce problème d’optimisation est résolu à l’aide d’une méthode telle que le *downhill simplex*.

Nous utilisons également une propriété. Elle peut être vue comme une relation unaire puisqu’elle ne caractérise qu’un seul organe. Cette propriété évalue à quel point la forme d’un organe est allongée. Étant donné un organe segmenté, une analyse en composantes principales est réalisée afin d’en obtenir les deux axes principaux. Ensuite, comme le décrit la figure 4, l’organe est projeté sur chacun de ces deux axes. Le rapport entre ces deux projections est alors utilisé pour déterminer le degré d’appartenance à cette propriété. Cependant, cette définition ne permet pas de bien juger l’allongement des objets concaves.

Notre vocabulaire de relations  $\mathcal{V}$  contient les relations sui-

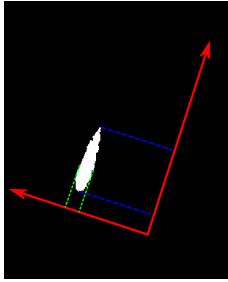


FIGURE 4 – (Affichage optimal en couleur) Illustration décrivant l'évaluation de la propriété *allongé*. Une analyse en composantes principales est effectuée afin d'obtenir les deux axes principaux. Ensuite, l'organe étudié est projeté sur ces deux axes. Soit  $p_g$  la projection verte (la plus petite) et  $p_b$  la projection bleue (la plus grande). Le degré d'appartenance est déterminé par  $1 - \frac{p_g}{p_b}$ .

vantes : à gauche de, à droite de, en dessous de, au-dessus de, près de, symétrique à et est allongé.  $\mathcal{V}$  comporte donc six relations binaires et une relation unaire. Puisque  $\text{Card}(\mathcal{Y}) = 9$ , le nombre de relations à évaluer pour chaque image est  $\text{Card}(\mathcal{V}_y) = 441$ . Cela contribue à rendre notre modèle expressif, mais le temps de calcul nécessaire à ces évaluations peut être important.

### 4.3 Initialisation du problème

Comme nous l'avons montré dans la section 3, la méthode complète est composée de trois étapes. Notre jeu de données contient pour chaque image les segments associés à chaque organe. Ces segments ne sont pas flous, mais la méthode resterait inchangée s'ils l'étaient.

L'objectif intermédiaire est de générer des contraintes pour définir un FCSP pour chaque exemple à évaluer. Une fois ce problème résolu, nous obtenons les segments correspondant à chaque organe et les contraintes peuvent désormais être utilisées pour générer des explications.

L'ensemble de variables  $X$  et l'ensemble de domaines  $D$  de notre FCSP sont les suivants :

$$X = \{x_{\text{foie}}, x_{\text{rate}}, x_{\text{vessie}}, x_{\text{rein\_d}}, x_{\text{rein\_g}}, x_{\text{poumon\_d}}, x_{\text{poumon\_g}}, x_{\text{psoas\_d}}, x_{\text{psoas\_g}}\}$$

$$D = \{D_{\text{foie}}, D_{\text{rate}}, D_{\text{vessie}}, D_{\text{rein\_d}}, D_{\text{rein\_g}}, D_{\text{poumon\_d}}, D_{\text{poumon\_g}}, D_{\text{psoas\_d}}, D_{\text{psoas\_g}}\}$$

avec  $D_i$  l'ensemble des segments possibles pour la variable  $x_i$ . Il s'agit des segments fournis dans le jeu de données. Pour chaque organe  $y \in \mathcal{Y}$ , les contraintes floues sont générées à partir de l'ensemble des sous-ensembles fréquents de relations de taille maximale  $F_y^{\text{max}}$ . De plus, comme chaque organe est unique, il ne peut pas y avoir deux annotations identiques sur la même image. Il faut donc compléter  $C$  avec des contraintes imposant la différence entre chaque variable deux à deux. C'est pourquoi nous ajoutons la contrainte globale *AllDifferent*.

Organe	Valeur du seuil
Foie	0.96
Rate	0.86
Vessie	0.80
Rein droit	0.92
Rein gauche	0.89
Poumon droit	0.98
Poumon gauche	0.97
Muscle psoas droit	0.92
Muscle psoas gauche	0.88

TABLE 1 – Valeurs des seuils correspondant à chaque organe. Ce sont les hyperparamètres de notre modèle.

La définition du FCSP est donc effectuée automatiquement. Ensuite, pour un exemple donné, le problème est résolu comme décrit dans la section 2.2.

### 4.4 Résultats

Un exemple est présenté sur la figure 5. Il y a neuf organes à annoter et donc neuf explications à fournir.

La mesure d'évaluation de notre modèle est la *précision*. Il s'agit du rapport entre le nombre d'annotations correctes pour tous les organes et le nombre total d'annotations effectuées. Nous obtenons une précision de 100% avec un modèle ne contenant que des relations directionnelles. La validation croisée externe contient trois plis (23/24 exemples dans la base d'apprentissage et 12/11 dans la base de test à chaque itération). La validation croisée interne comporte quatre plis. Comme il y a neuf organes à annoter, il y a neuf hyperparamètres à déterminer pour extraire les ensembles de relations fréquents. Les valeurs obtenues pour ces hyperparamètres sont affichées dans la table 1. Des contraintes pourraient être ajoutées sur les valeurs que peuvent prendre ces hyperparamètres pour spécifier la longueur des explications.

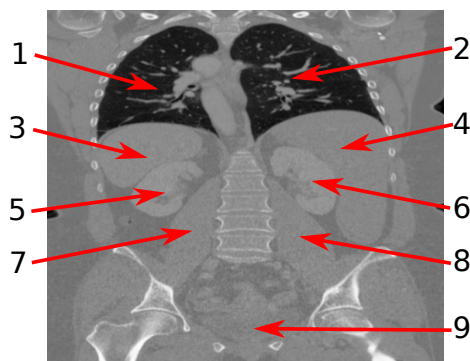
Nous observons que les explications reprennent bien les relations qui ont été extraites puis traduites en contraintes. Pour l'exemple proposé sur la figure 5, l'ensemble des contraintes associées au rein droit est :

$$C_{\text{rein\_d}} = \{(x_{\text{rein\_d}}, x_{\text{rein\_g}}, R_{\text{symétrique à}}), (x_{\text{poumon\_d}}, x_{\text{rein\_d}}, R_{\text{au-dessus de}}), (x_{\text{rein\_d}}, x_{\text{foie}}, R_{\text{à gauche de}}), (x_{\text{vessie}}, x_{\text{rein\_d}}, R_{\text{en dessous de}}), (x_{\text{rein\_d}}, x_{\text{rein\_g}}, R_{\text{à droite de}}), (x_{\text{rein\_g}}, x_{\text{rein\_d}}, R_{\text{à gauche de}})\}.$$

Certaines de ces contraintes peuvent sembler redondantes, comme les deux dernières par exemple. Ce genre de situation peut arriver car le résultat d'une dilatation floue dépend de la forme de l'objet de référence. Comme deux organes différents ne peuvent pas avoir exactement la même forme, il peut y avoir de légères différences entre deux contraintes réciproques. Chaque organe est lié à un tel ensemble de contraintes.  $C$  est l'union de tous ces ensembles.

Bien que le modèle entraîné avec uniquement des relations directionnelles n'ait pas besoin d'autres types de relation pour obtenir une précision de 100%, ajouter de nouvelles





L'annotation de l'organe 1 est très probablement le poumon gauche **car** il est à gauche du poumon droit (organe 2), il est *symétrique* au poumon droit et il est *au-dessus* de la rate (3).

L'annotation de l'organe 9 est probablement la vessie **car** elle est *allongée*, elle est *en dessous* du rein droit (6) et *en dessous* du rein gauche (5).

L'annotation de l'organe 4 est le foie **car**...

FIGURE 5 – Exemple d'annotations justifiées.

relations peut permettre de construire de meilleures explications. Comme on peut le voir sur la figure 5, la relation *symétrique* à est utilisée dans l'explication de l'annotation du poumon gauche. Cette relation n'est pas nécessaire puisque nous atteignons 100% de précision sans elle, mais elle rend l'explication plus convaincante.

Nous avons aussi étudié le nombre d'exemples nécessaires dans la base d'apprentissage pour obtenir de bonnes performances. Pour une base d'apprentissage comportant uniquement deux images (et 33 exemples dans la base de test, par validation croisée imbriquée), nous obtenons au moins 99% de précision. Dans le cas extrême d'une base d'apprentissage contenant un seul exemple, les relations fréquentes seront celles dont le degré d'appartenance est plus grand que les seuils considérés (cf. section 3.2). Ainsi, tout exemple qui n'est pas une aberration devrait permettre d'obtenir de bonnes performances.

Enfin, nous observons que notre modèle obtient de meilleures performances que le réseau de neurones convolutionnel présenté dans [33]. Ce modèle a été entraîné sur un jeu de données plus grand que le nôtre et ne fournit pas d'explications. La méthode la plus proche de la nôtre, présentée dans [38], ne donne pas de mesure de précision comme référence. Son désavantage est qu'elle peut ne pas assigner certaines classes d'objet, ce qui arrive en moyenne une fois tous les cinq exemples. Avec notre approche, ce genre de situation est impossible car chaque variable de notre problème doit être associée à une région de l'image.

## 5 Conclusion et perspectives

Dans cet article, nous présentons une nouvelle méthode d'apprentissage et de raisonnement visuel dont le but est d'annoter des objets dans des images et d'expliquer les

résultats obtenus. Cette tâche est formalisée comme un problème de satisfaction de contraintes floues. Cette méthode est basée sur l'apprentissage de relations spatiales floues sur un ensemble d'objets annotés afin de générer contraintes. Nous avons validé notre approche sur un jeu d'images médicales et avons montré qu'elle tire profit de l'apprentissage symbolique et de sa capacité à raisonner pour expliquer les résultats qu'elle produit et atteindre une précision de 99% pour une base d'apprentissage de taille 2. Nous envisageons de développer une stratégie pour réaliser la première étape de manière plus rapide. Une première idée est de déterminer une structure hiérarchique des relations spatiales afin de pouvoir utiliser un tri topologique. Une deuxième idée est de mettre à jour l'ordre d'évaluation des relations à partir des exemples précédemment étudiés. De plus, puisque l'utilisation de la logique floue nous permet de gérer des segments imprécis, nous souhaitons utiliser une méthode de segmentation non-supervisée sur les images de tests pour initialiser les domaines de  $D$ . Plus globalement, il s'agit d'une première étape vers une combinaison des apprentissages statistique et symbolique afin de créer une intelligence artificielle explicable.

## Références

- [1] I. Biederman. *On the Semantics of a Glance at a Scene*. 1981.
- [2] I. Bloch. Fuzzy relative position between objects in image processing : a morphological approach. *IEEE transactions on pattern analysis and machine intelligence*, 21(7) :657–664, 1999.
- [3] I. Bloch. On fuzzy distances and their use in image processing under imprecision. *Pattern Recognition*, 32(11) :1873–1895, 1999.
- [4] I. Bloch. Fuzzy spatial relationships for image processing and interpretation : a review. *Image and Vision Computing*, 23(2) :89–110, 2005.
- [5] G. C. Cawley and N. L. C. Talbot. On over-fitting in model selection and subsequent selection bias in performance evaluation. *Journal of Machine Learning Research*, 11(Jul) :2079–2107, 2010.
- [6] M. Cayrol, H. Farreny, and H. Prade. Fuzzy pattern matching. *Kybernetes*, 11(2) :103–116, 1982.
- [7] O. Colliot, A. V. Tuzikov, R. M. Cesar, and I. Bloch. Approximate reflectional symmetries of fuzzy objects with an application in model-based object recognition. *Fuzzy Sets and Systems*, 147(1) :141 – 163, 2004.
- [8] A. Criminisi, J. Shotton, D. Robertson, and E. Konukoglu. Regression forests for efficient anatomy detection and localization in ct studies. In *Medical Computer Vision. Recognition Techniques and Applications in Medical Imaging*, pages 106–117, 2011.
- [9] F. Doshi-Velez and B. Kim. Towards a rigorous science of interpretable machine learning. In *eprint arXiv :1702.08608*, 2017.
- [10] D. Dubois, H. Fargier, and H. Prade. Possibility theory in constraint satisfaction problems : Handling priority, preference and uncertainty. *Applied Intelligence*, 6(4) :287–309, 1996.

- [11] European Council. The general data protection regulation, 2016.
- [12] L. Fei-Fei, R. Fergus, and P. Perona. One-shot learning of object categories. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 28(4) :594–611, April 2006.
- [13] M. Garnelo and M. Shanahan. Reconciling deep learning with symbolic artificial intelligence : representing objects and relations. *Current Opinion in Behavioral Sciences*, 29 :17 – 23, 2019.
- [14] A. Gatt and E. Reiter. Simplenlg : A realisation engine for practical applications. In *Proceedings of the 12th European Workshop on Natural Language Generation*, ENLG '09, pages 90–93, 2009.
- [15] L. H. Gilpin, D. Bau, B. Z. Yuan, A. Bajwa, M. Specter, and L. Kagal. Explaining Explanations : An Approach to Evaluating Interpretability of Machine Learning. *ArXiv e-prints*, 2018.
- [16] D. Gunning. Explainable artificial intelligence (xai). 2017.
- [17] G. Hinton and N. Frosst. Distilling a neural network into a soft decision tree. 2017.
- [18] O. Jimenez-del Toro, H. Müller, M. Krenn, K. Gruenberg, A. A. Taha, M. Winterstein, I. Eggel, A. Foncubierta-Rodríguez, O. Goksel, A. Jakab, et al. Cloud-based evaluation of anatomical structure segmentation and landmark detection algorithms : Visceral anatomy benchmarks. *IEEE transactions on medical imaging*, 35(11) :2459–2475, 2016.
- [19] D. R. Jones, M. Schonlau, and W. J. Welch. Efficient global optimization of expensive black-box functions. *Journal of Global Optimization*, 13(4) :455–492, 1998.
- [20] H. Larochelle, D. Erhan, and Y. Bengio. Zero-data learning of new tasks. In *Proceedings of the 23rd National Conference on Artificial Intelligence - Volume 2*, pages 646–651, 2008.
- [21] C-C. Lee, P-C. Chung, and H-M. Tsai. Identifying multiple abdominal organs from ct image series using a multi-module contextual neural network and spatial fuzzy rules. *IEEE Transactions on Information Technology in Biomedicine*, 7(3) :208–217, Sep. 2003.
- [22] F-F. Li, R. VanRullen, C. Koch, and P. Perona. Rapid natural scene categorization in the near absence of attention. *Proceedings of the National Academy of Sciences*, 99(14) :9596–9601, 2002.
- [23] Z.C. Lipton. The myths of model interpretability. *Queue*, 16(3) :30 :31–30 :57, June 2018.
- [24] G. Litjens, T. Kooi, B. E. Bejnordi, A. A. Adiyoso Setio, F. Ciompi, M. Ghafoorian, J. A.W.M. van der Laak, B. van Ginneken, and C. I. Sánchez. A survey on deep learning in medical image analysis. *Medical Image Analysis*, 42 :60 – 88, 2017.
- [25] S.M. Lundberg and S-I. Lee. A unified approach to interpreting model predictions. In *Advances in Neural Information Processing Systems 30*, pages 4765–4774. 2017.
- [26] G. Marcus. Deep learning : A critical appraisal. *CoRR*, abs/1801.00631, 2018.
- [27] J. McClure. Goal-based explanations of actions and outcomes. *European review of social psychology*, 12(1) :201–235, 2002.
- [28] T. Miller. Explanation in artificial intelligence : Insights from the social sciences. *Artificial Intelligence*, 267 :1–38, February 2019.
- [29] O. Pauly, B. Glocker, A. Criminisi, D. Mateus, A.M. Möller, S. Nekolla, and N. Navab. Fast multiple organ detection and localization in whole-body mr dixon sequences. In *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2011*, pages 239–247, 2011.
- [30] R. Pierrard, J-P. Poli, and C. Hudelot. A fuzzy close algorithm for mining fuzzy association rules. In *Information Processing and Management of Uncertainty in Knowledge-Based Systems. Theory and Foundations*, pages 88–99, 2018.
- [31] S. J. Read and A. Marcus-Newhall. Explanatory coherence in social explanations : A parallel distributed processing account. *Journal of Personality and Social Psychology*, 65(3) :429, 1993.
- [32] M.T. Ribeiro, S. Singh, and C. Guestrin. Why should i trust you? : Explaining the predictions of any classifier. In *Proceedings of the 22nd ACM SIGKDD international conference on knowledge discovery and data mining*, pages 1135–1144, 2016.
- [33] H.R. Roth, C.T. Lee, H-C. Shin, A. Seff, L. Kim, J. Yao, L. Lu, and R.M. Summers. Anatomy-specific classification of medical images using deep convolutional nets. *arXiv preprint arXiv :1504.04003*, 2015.
- [34] H. Shin, M. R. Orton, D. J. Collins, S. J. Doran, and M. O. Leach. Stacked autoencoders for unsupervised feature learning and multiple organ detection in a pilot study using 4d patient data. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 35(8) :1930–1943, Aug 2013.
- [35] S. Thorpe, D. Fize, and C. Marlot. Speed of processing in the human visual system. *Nature*, 381(6582) :520, 1996.
- [36] M. C. Vanegas, I. Bloch, and J. Inglada. Fuzzy constraint satisfaction problem for model-based image interpretation. *Fuzzy Sets and Systems*, 286 :1 – 29, 2016.
- [37] Y. Xian, C. H. Lampert, B. Schiele, and Z. Akata. Zero-shot Learning – a Comprehensive Evaluation of the Good, the Bad and the Ugly. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2018.
- [38] Z. Xue, S. Antani, L. R. Long, and G. R. Thoma. Automatic multi-label annotation of abdominal ct images using cbir. In *Medical Imaging 2017 : Imaging Informatics for Healthcare, Research, and Applications*, volume 10138, page 1013807, 2017.
- [39] L. A. Zadeh. Fuzzy sets. *Information and Control*, 8(3) :338 – 353, 1965.



# Pourquoi dois-je croire ta prédiction ?

## Comment expliquer les résultats d'une classification automatique de sentiments ?

Waleed Ragheb<sup>1,2</sup>  
Sandra Bringay<sup>1,3</sup>

Jérôme Azé<sup>1,2</sup>  
Maximilien Servajean<sup>1,3</sup>

<sup>1</sup> LIRMM UMR 5506, CNRS, University of Montpellier

<sup>2</sup> IUT DE BÉZIERS, University of Montpellier, Béziers

<sup>3</sup> AMIS, Paul Valéry University - Montpellier 3

first.last@lirmm.fr

### Résumé

*Dans le cadre d'un problème classique de classification de sentiments, nous proposons un modèle qui utilise 1) l'apprentissage par transfert plutôt que les méthodes classiques de word embedding et 2) des mécanismes d'attention permettant de se concentrer sur les parties du texte importantes pour la tâche de classification étudiée. Notre modèle a été évalué sur plusieurs jeux de données et montre des résultats très compétitifs. Or, si ces méthodes d'apprentissage en profondeur s'avèrent très efficaces, elles sont souvent considérées comme des boîtes noires, difficiles à interpréter. Nous évaluons ici l'impact des mécanismes d'attention traduits sous la forme de nuages de mots-clés pour aider les utilisateurs à interpréter les résultats de la classification.*

### Mots Clef

Mécanisme d'attention et interprétabilité.

### Abstract

*Text sentiment classification is one of the important problems in natural language processing. However good the performance of most deep learning models, they often lack interpretability. To this end, this paper addresses the problem of sentiment analysis by proposing a model that makes use of 1) transfer learning, rather than the classical shallow methods of word embedding and 2) self-attention mechanisms to focus on the most important parts of the text that have high influence on sentiments. Our model was evaluated on several datasets and shows very competitive results. Moreover, we evaluate the impact of attention mechanisms that enables users to have an efficient and meaningful interpretation.*

### Keywords

Attention mechanism and interpretability.

## 1 Introduction

L'analyse de sentiments à partir de textes est un domaine de recherche très actif dans la communauté de l'apprentissage automatique [12] avec de nombreuses applications comme le e-commerce, la gestion de la réputation, le support client, la politique, etc. De nombreux outils sont désormais disponibles pour détecter des sentiments subjectifs, tels que la polarité (positif ou négatif) et les émotions (peur, joie, etc.). Si ces outils sont de plus en plus efficaces, ils sont généralement critiqués pour deux raisons.

Tout d'abord, la plupart des approches existantes ne sont efficaces que pour des domaines d'application spécifiques ou des types de textes particuliers tels que l'analyse de sentiments des sites Web liés à la finance ou des avis clients relatifs à des ordinateurs portables ou des restaurants. Dans ce contexte, l'apprentissage par transfert et l'adaptation au domaine sont largement utilisés, en particulier combinés à des réseaux de neurones profonds, pour aider à réutiliser les modèles développés pour une tâche source vers une autre tâche cible. L'apprentissage par transfert fonctionne particulièrement bien quand les caractéristiques apprises pour la tâche source sont générales et peuvent être réutilisées pour les tâches cibles. Ce type d'approche a fait ses preuves dans le domaine de la vision par ordinateur où l'extraction des caractéristiques se fait à partir de modèles pré-entraînés de type AlexNet, ResNet, MS-COCO [22]. Dans les modèles de traitement du langage naturel, cette approche n'a connu de réel succès que très récemment grâce au modèle de langage universel (ULMFiT) proposé par [6] qui servira de baseline à ces travaux.

Par ailleurs, en traitement automatique de la langue, la plupart des modèles de transduction compétitifs ont une structure de type codeur-décodeur [20]. Une limite de ces architectures est qu'elle code la séquence d'entrée dans une représentation interne de longueur fixe. Cela entraîne une dégradation des résultats lorsque la longueur de la séquence augmente. Dans ce contexte, les mécanismes d'attention [26] ont récemment été utilisés pour résoudre ce problème.

Inspirés des mécanismes d'attention visuelle que l'on retrouve chez l'homme, ils focalisent l'analyse sur certaines régions d'une image avec une "haute résolution" tout en percevant le reste de l'image en "basse résolution". L'attention guide le réseau pour qu'il sache où accorder son attention sur la séquence d'entrée. Les premières applications des mécanismes d'attention ont été naturellement réalisées dans le domaine de la vision par ordinateur [1] et plus récemment sur les textes pour des applications de traduction automatique [2] et d'analyse des sentiments [10]. Dans ce travail, nous évaluerons la combinaison des mécanismes d'attention à une architecture de type ULMFiT sur des jeux de données réelles de la littérature.

Par ailleurs, nous nous posons la question de l'interprétabilité. Les modèles d'apprentissage sont souvent décriés car perçus comme des boîtes noires. Or, pour un utilisateur qui va baser des décisions et des actions à partir de prédictions, il peut être fondamental d'interpréter les raisons sous-jacentes aux prédictions. En effet, il y a des cas où l'erreur a peu d'importance, par exemple quand on recommande un film, mais il y a des cas où l'erreur porte à conséquence, par exemple pour le diagnostic médical. Selon l'importance de l'erreur pour l'activité de l'utilisateur, celui-ci aura besoin de comprendre comment le modèle fonctionne en général et comment il est arrivé à proposer une prédiction particulière.

Dans ce travail, nous nous intéressons à la sortie des mécanismes d'attention qui est la visualisation des parties du texte ayant impacté la prédiction du sentiment. Cette visualisation donne des justifications complémentaires aux étiquettes prédites associées aux textes. À notre connaissance, aucune étude n'a montré l'apport de la visualisation des mécanismes d'attention sur l'interprétation des utilisateurs pour l'analyse de sentiments à partir de textes.

Finalement, l'objectif de cet article est double : 1) montrer l'impact en terme d'exactitude des mécanismes d'attention pour l'analyse des sentiments lorsqu'ils sont ajoutés à l'architecture de référence ULMFiT qui intègre un apprentissage par transfert ; 2) évaluer dans quelle mesure une visualisation basée sur le mécanisme d'attention facilite la prise de décision en apportant un élément explicatif supplémentaire à l'étiquette proposée.

Le reste de cet article est organisé comme suit. Dans la section 2, nous décrivons les travaux connexes de la littérature puis le modèle proposé dans la section 3. Dans la section 4, nous présentons les expérimentations sur les performances du modèle, les jeux de données et les résultats obtenus. Dans la section 5, nous détaillons les expérimentations réalisées sur l'interprétabilité du modèle. Nous concluons et donnons des perspectives dans la section 6.

## 2 État de l'art

Afin d'améliorer l'efficacité des algorithmes d'analyse de sentiments et leur explicabilité, nous proposons une méthode basée sur l'apprentissage par transfert de modèles de langue ainsi que sur les mécanismes d'attention.

Les modèles de langue (LM) visent à prédire un mot à partir des mots le précédant. Ces modèles sont utilisés dans de nombreuses applications de traitement automatique de la langue naturelle car ils permettent de capturer des dépendances éloignées ainsi que la structure hiérarchique du texte. L'apprentissage des modèles de langue est non supervisé, car ne nécessitant pas de corpus de texte préalablement étiqueté. Or, ces modèles sont peu adaptés aux petits ensembles de données et peuvent donner un mauvais rappel pour certaines tâches de classification. Récemment, [6] ont proposé la méthode ULMFiT, basée sur une représentation des mots peu profonde, qui combinée à un apprentissage par transfert s'est avérée être très efficace pour différentes tâches dont l'analyse de sentiments [13]. Cette méthode servira de baseline à nos travaux. Nous avons choisi d'utiliser ULMFiT pour ses performances en *fine tuning* et sa taille raisonnable en comparaison à d'autres approches très récentes comme BERT [3] et ELMO [17]. De plus, pour des tâches de classification de sentiments, ULMFiT est utilisé comme l'état de l'art sur la plupart des jeux de données de la littérature, dont ceux sur lesquels nous avons travaillé.

Dans le domaine du traitement automatique de la langue naturelle, les mécanismes d'attention ont été récemment étudiés. L'auto-attention met en relation différentes positions d'une séquence afin d'en calculer une représentation [9]. Chaque partie de la séquence d'entrée est associée à un score de probabilité. L'attention a été appliquée avec succès pour des tâches comme la compréhension de la lecture, le résumé, la traduction automatique [18].

Outre le fait que les mécanismes d'attention améliorent l'efficacité de la classification, ce qui nous intéresse dans cet article est le fait que le résultat de la couche d'attention peut être utilisé en entrée d'une visualisation visant l'explication d'une prédiction [24].

Une explication est une réponse à une question de type « Pourquoi ? ». Dans notre cas, pourquoi le système prédit une polarité pour un texte ? Il a été démontré par [5] qu'associer des explications aux prédictions améliore l'acceptation de ces prédictions dans le cas de la recommandation de films. En effet, les explications influencent le ressenti de celui à qui l'on fournit les explications et donc les actions qu'il peut être amené à réaliser. Pour expliquer une prédiction, la plupart des méthodes utilisent des artefacts visuels qui fournissent une compréhension qualitative des liens existants entre les instances (des mots, des parties d'images...) et les prédictions.

Il existe différentes méthodes d'interprétabilité. On distingue les méthodes d'apprentissage intrinsèquement interprétables lorsque la sélection et l'entraînement du modèle d'apprentissage sont par eux-même interprétables (e.g. les arbres de décision) et les méthodes d'interprétabilité post-hoc qui donnent des explications *a posteriori* et qui s'appliquent sur des méthodes d'apprentissage de type boîte noire après la sélection et l'entraînement du modèle, pour expliquer les prédictions réalisées. La plupart des modèles d'apprentissage de l'état de l'art dont les modèles

de type réseau de neurones étudiés dans cet article, n'étant pas interprétables, nous nous sommes focalisés sur cette deuxième catégorie de méthodes d'interprétabilité.

À notre connaissance, il n'y a pas eu d'étude qualitative de l'impact de la visualisation de l'attention sur l'interprétation que les utilisateurs peuvent faire de cette nouvelle information dans le cas de la classification de données textuelles. Cette étude nous paraît importante dans le contexte du besoin de confiance manifesté par les utilisateurs des méthodes d'apprentissage notamment dans des domaines comme la santé. Ces utilisateurs ne sont pas uniquement en attente de résultats performants mais ils recherchent des explications qui, associées aux prédictions, vont améliorer la prise de décision [21].

### 3 Architecture et apprentissage

Notre architecture est composée de trois composants entraînés selon une procédure basée sur le transfert d'apprentissage. Ces éléments sont décrits dans les sections suivantes.

#### 3.1 Auto-Attention basée sur un encodeur AWD-LSTM

Un LSTM traditionnel possède un portail d'entrée  $i_t$ , d'oubli  $f_t$ , de sortie  $o_t$  et une cellule mémoire  $c_t$ . Ce sont des vecteurs de  $R^d$  qui correspondent à la représentation vectorielle d'une dimension  $d$ . Les équations de transition de LSTM sont les suivantes :

$$\begin{aligned} f_t &= \sigma(W_f x_t + U_f h_{t-1}) \\ i_t &= \sigma(W_i x_t + U_i h_{t-1}) \\ o_t &= \sigma(W_o x_t + U_o h_{t-1}) \\ \tilde{c}_t &= \tanh(W_c x_t + U_c h_{t-1}) \\ c_t &= f_t \odot c_{t-1} + i_t \odot \tilde{c}_t \\ h_t &= o_t \odot \tanh(c_t) \end{aligned} \quad (1)$$

où  $x_t$  est l'entrée au pas de temps actuel,  $\sigma$  est la fonction sigmoïde ;  $\odot$  l'opération de multiplication par élément,  $W_{\{i,f,o,c\}}$ ,  $U_{\{i,f,o,c\}}$  sont des ensembles de poids appris. Dans notre modèle, nous utilisons le vecteur d'état caché à chaque pas de temps comme représentation du mot correspondant dans une phrase. Afin d'éviter le sur-apprentissage lors de l'entraînement du LSTM, [13] ont proposé AWD-LSTM. basée sur DropConnect [23] pour pondérer les matrices  $U_{\{i,f,o,c\}}$ . Nous avons utilisé les trois mêmes couches liées de LSTM et avons également appliqué une auto-attention sur les vecteurs d'état cachés à chaque pas de temps. La séquence d'états cachés en entrée  $H^{i-1} = \{h_1^{i-1}, h_2^{i-1}, \dots, h_N^{i-1}\}$ , où  $N$  est la longueur de la séquence, est passée aux états de la couche LSTM. Les états de sortie sont de la forme de  $H^i = \{h_1^i, h_2^i, \dots, h_N^i\}$ . La couche d'attention prend la séquence en entrée encodée et calcule les scores d'attention  $S^i = \{s_1^i, s_2^i, \dots, s_N^i\}$ . La

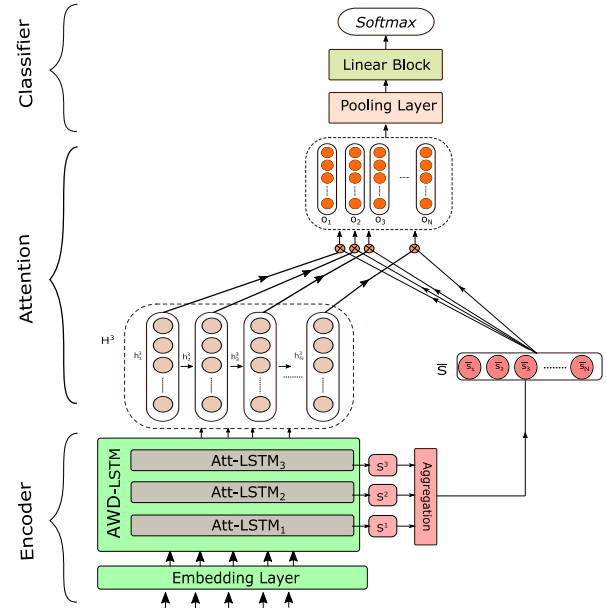


FIGURE 1 – Notre architecture

couche d'attention est une couche linéaire sans biais.

$$\begin{aligned} \alpha^i &= \{V^i \cdot H^i\} \\ S^i &= \exp(\alpha^i) / \sum_{j=1}^N \exp(\alpha_j^i) \end{aligned} \quad (2)$$

Où  $V^i$  est le poids de la couche d'attention  $i^{th}$ .

#### 3.2 Agrégation sur plusieurs niveaux de l'auto-attention

L'architecture proposée utilise un empilement de trois Att-LSTM superposés exactement, de la même manière que l'architecture AWD-LSTM classique. Nous n'utilisons pas d'architecture Bi-LSTM ici, car notre modèle correspond à l'ensemble des directions avant et arrière. La figure 1 montre notre modèle. À chaque couche, les scores d'attention sont obtenus selon un niveau spécifique de codage de la séquence, puis agrégés pour obtenir les scores d'attention globaux  $\bar{S}$ . La fonction d'agrégation est la moyenne logarithmique des 3 niveaux de scores d'attention.

$$\bar{S} = \log \sum_{i=1}^3 S^i / 3 \quad (3)$$

Les scores globaux d'attention  $\bar{S}$  sont utilisés pour calculer la séquence  $O = \{o_1, o_2, \dots, o_N\}$  où  $o_i$  est le produit du score d'attention et de la sortie de la couche Att-LSTM :

$$o_i = \bar{s}_i \otimes h_i^3 \quad (4)$$

### 3.3 Couches de classification

Après avoir agrégé les informations provenant de l'attention multi-niveaux et de la sortie de l'encodeur, nous convertissons les représentations résultantes de toutes les positions de  $O$  en un vecteur à longueur fixe avec une opération de pooling. Nous avons utilisé trois fonctions de pooling. Nous appliquons une attention  $X_{att}$  telle que :

$$X_{att} = \sum_{i=1}^N \exp(\bar{s}_i) \otimes h_i^3 \quad (5)$$

Nous appliquons un pooling maximum  $X_{max}$  et moyen  $X_{avg}$  à  $O$  afin d'obtenir les représentations finales du texte saisi après encodage et application de l'attention :

$$X_{in} = [X_{att} \oplus X_{max} \oplus X_{avg}] \quad (6)$$

Puis, nous alimentons le bloc linéaire du classifieur. Ce bloc est constitué de deux couches entièrement connectées, de tailles différentes, suivies d'un indicateur *softmax* pour déterminer la classe de sentiment en sortie.

### 3.4 Entraînement du modèle

L'entraînement se fait en trois étapes :

- Le modèle de langue est initialisé de manière aléatoire, puis entraîné en empilant un décodeur linéaire au-dessus de l'encodeur. Le modèle de langue est appris à partir d'un corpus du domaine général. Cela aide le modèle à apprendre des caractéristiques générales du langage.
- Le même modèle de langue après entraînement est utilisé pour l'initialisation puis ajusté à partir des données de la tâche cible, dans notre cas : différents jeux de données d'analyse de sentiments. Ici, nous limitons le vocabulaire du modèle de langue aux mots fréquents (répétés plus de deux fois).
- Nous conservons l'encodeur et remplaçons le décodeur par le classifieur et les paramètres de ces deux étapes sont réglés avec précision pour la tâche cible.

Lors de la première étape et pour l'apprentissage du modèle de langage, nous avons utilisé le jeu de données Wikitext-103 [14]. Avec plus de 28 000 articles Wikipédia et 103 millions de mots, le modèle détermine la structure principale et la hiérarchie du langage par modélisation séquence par séquence. Pour l'entraînement du modèle de classification, à partir du jeu étiqueté, nous optimisons tous les paramètres afin de minimiser autant que possible la fonction objectif. Dans notre travail, nous prenons  $y_i$  la polarité du sentiment correcte et  $\hat{y}_i$  qui désigne la polarité de sentiment prédite. Nous considérons l'entropie croisée comme la fonction de perte, dont la formule est la suivante :

$$loss = - \sum_{\langle T \rangle} y_i \log(\hat{y}_i) + \lambda \|\theta\|^2 \quad (7)$$

Où  $\lambda$  est le facteur de régularisation,  $\theta$  contient tous les paramètres du modèle et  $T$  est l'ensemble des exemples

d'entraînement. L'entraînement de l'architecture est effectué à l'aide de taux d'apprentissage triangulaires inclinés (STLR) qui modifient le taux d'apprentissage pour chaque itération de manière triangulaire. Nous avons utilisé un seul cycle comme recommandé par [6]. Le modèle a été entraîné en utilisant un taux d'apprentissage différent pour chaque groupe de couches.

Nous entraînons le modèle sur les modèles de langue avant et arrière pour les données du domaine général et spécifiques à la tâche. Ces deux modèles de langue sont utilisés pour construire deux versions de la même architecture. La décision finale est l'ensemble des deux. Nous avons utilisé Pytorch<sup>1</sup> pour construire l'ensemble du modèle et les bibliothèques Fastai<sup>2</sup> pour l'apprentissage et les modèles de langue. Pour le prétraitement du texte, celui-ci a tout d'abord été normalisé et tokenisé. Des tokens spéciaux ont été ajoutés pour les mots en majuscules et ceux répétés. Nous conservons les symboles de ponctuation et de sentiments dans le texte. Nous avons utilisé Spacy<sup>3</sup> et FastText<sup>4</sup> pour ces prétraitements. Les modèles sont appris et testés sur 4 GPU Nvidia GEFORCE GTX 1080 ti.

## 4 Expérimentations sur les performances du modèle

Dans cette section, nous discutons plus en détail de l'efficacité de la méthode proposée.

### 4.1 Jeux de données

Nous appliquons le modèle à différents jeux de données de classification de sentiments. Le tableau 1 présente des statistiques sur ces jeux de données. Le jeu de données IMDB est un ensemble de données pour la classification des sentiments binaires de critiques de films [11]. Nous avons également utilisé les versions binaires et complètes des jeux de données d'avis d'utilisateurs de Yelp et d'Amazon [27]. Pour les jeux de données binaires (IMDB, Yelp-bi et Amazon-bi), les classes à prédire sont positif ou négatif. Pour les autres, il s'agit d'un nombre d'étoiles : de négatif (1 étoile) à positif (5 étoiles). Ces jeux de données sont tous équilibrés.

TABLE 1 – Jeux de données de sentiments et nombre d'exemples d'apprentissage et de test

Dataset	#Exemples d'apprentissage	#Exemples de test	#classes
IMDB	25K	25K	2
Yelp-bi	560K	38K	2
Yelp-Full	650K	50K	5
Amazon-bi	3.6M	400K	2
Amazon-Full	3M	650K	5

1. <https://pytorch.org/>

2. <http://www.fast.ai/>

3. <https://spacy.io/>

4. <https://fasttext.cc/>

## 4.2 Baselines et Résultats

Nous comparons notre modèle à plusieurs baselines compétitives de l'état de l'art qui utilisent le mécanisme d'attention pour la classification de sentiments :

- HN-ATT [25] reflète la structure hiérarchique des documents à travers deux niveaux d'attention dans les mots et les phrases.
- DCCNN-ATT [24] est un réseau de neurones convolutifs avec des connexions denses et des fonctionnalités multi-échelles.
- SANet [8] utilise l'attention pour modéliser les interactions entre toutes les paires de mots d'entrée.
- SA-Embedding [9] est basé sur l'extraction d'une représentation interprétable de la phrase donnée en entrée via un mécanisme d'attention.
- CSC [16] utilise un réseau de neurones hiérarchique basé sur l'attention qui intègre les préférences de l'utilisateur et les caractéristiques du produit dans les tâches de classification de sentiments.
- CRAN [4] associe les attentions basées sur la convolution et celles basées sur la récurrence.
- IRAM [19] construit de manière récursive des représentations d'entrée par réutilisation des résultats précédemment calculés.
- ULMFiT [6] est la méthode de référence actuelle.

Le tableau 2 montre l'erreur obtenue lors du test du modèle proposé et de toutes les baselines sur les jeux de tests. Nous présentons les résultats tels que rapportés dans la publication d'origine. Notre modèle surpasse tous les modèles basés sur l'attention avec une marge significative et reste compétitif par rapport à ULMFiT. Notre modèle proposé obtient de meilleurs résultats qu'ULMFiT pour le jeu IDBM et la version complète de Yelp.

Dans la suite, nous allons nous demander comment améliorer l'interprétabilité de ce modèle et montrer comment les mécanismes d'attention peuvent être utilisés pour expliquer les prédictions.

## 5 Expérimentations sur l'interprétabilité du modèle

Dans cette section, nous discutons plus en détail de l'impact de l'application de l'attention sur l'interprétabilité.

### 5.1 Visualisation de l'attention

L'utilisateur doit avoir confiance dans un modèle et pour cela, il voudra vérifier que celui-ci fonctionne bien sur des données réelles selon une métrique d'intérêt, comme le taux d'erreur utilisé dans la section 4. Ce type d'évaluation peut s'avérer inefficace notamment quand les données évoluent au cours du temps car elle revient à évaluer entre autre l'écart entre les données réelles et celles d'entraînement. Un humain peut alors élaborer différentes stratégies pour sélectionner un modèle parmi plusieurs modèles et notamment préférer un modèle explicable mais moins performant, qui lui laisse la responsabilité du choix final.

Dans ce contexte, l'un des résultats les plus intéressants du mécanisme d'attention est sa capacité à traiter toutes les séquences d'entrée avec différents poids d'attention. Le modèle accorde une plus grande attention aux éléments qui influencent la décision du réseau. La figure 2 montre des exemples d'avis positifs et négatifs correctement classés pour les avis de restaurants (Yelp-bi) et de critiques de film (IDBM). Les scores d'attention sont utilisés pour colorer le texte. Cette information peut être alors présentée à l'utilisateur avec la prédiction comme explication.

Dans notre contexte, une explication est un ensemble de mots-clés associés à un score d'attention. Dans la suite, nous évaluons la qualité des explications fournies via la visualisation de l'attention selon une approche centrée sur le système (le niveau d'interprétabilité est basé sur l'analyse des sorties du système) puis selon une approche centrée sur l'humain (le niveau d'interprétabilité est basé sur une tâche expérimentale d'inférence de la polarité).

### 5.2 Évaluation centrée système

Nous avons mis en place une évaluation centrée système, sans intervention humaine. Pour cela, nous avons mesuré l'intersection entre le nombre de mots appartenant à des lexiques ayant fait leur preuve pour les tâches de classification de sentiments et les mots repérés par les mécanismes d'attention. Nous avons utilisé pour cela EmoLex proposé par [15] qui comporte plus de 10 000 entrées.

Le tableau 3 montre les résultats de la mise en correspondance des principaux mots repérés avec l'attention et retrouvés dans EmoLex pour les deux jeux de tests. Par exemple, nous pouvons dire que 88,11% des exemples du jeu de données IMDB contiennent des sentiments dans les 5% supérieurs des mots identifiés par le processus d'attention. Cela reflète la précision du mécanisme d'attention qui se concentre sur ces mots et expressions.

### 5.3 Évaluation centrée humain

Même si les modèles d'attention se concentrent sur certains mots porteurs d'émotions, ils se repèrent aussi d'autres parties du texte, ayant un impact important sur la décision du réseau. Sur la base d'exemples tels que ceux de la figure 2, nous pourrions facilement trouver des parties du texte surlignées qui ne sont pas liées aux lexiques de sentiments.

Afin de déterminer dans quelle mesure une explication basée sur le simple fait de prêter attention aux mots importants, malgré une sélection imparfaite par l'algorithme, pourrait induire une décision de polarité meilleure qu'une explication basée sur les termes d'un lexique, nous avons mis en place une évaluation centrée humain impliquant des utilisateurs non experts. Pour cela, nous leur avons présenté les deux types d'explications générées sous la forme de nuages de mots. Les nuages de mots sont générées de trois manières différentes : en utilisant 1) uniquement le lexique, 2) uniquement l'attention, ou 3) les deux. Dans le cas des nuages de mots générés à partir du lexique, la taille des mots est déterminée par les fréquences des mots et leur ordre dans le texte. Dans le cas des mécanismes d'atten-

TABLE 2 – Taux d'erreur (%) de notre modèle et des baselines

Models	IMDB	Yelp-bi	Yelp-Full	Amazon-bi	Amazon-Full
HN-ATT	-	-	-	-	36.40
DCCNN-ATT	-	2.64	30.58	<b>3.32</b>	34.81
SANet	-	4.77	36.03	4.52	38.67
SA-Embedding	-	5.10	36.60	-	40.20
CSC	-	6.90	35.97	4.90	39.89
CRAN	7.90	-	-	-	-
IRAM	8.80	-	-	-	-
ULMFiT	4.60	<b>2.16</b>	29.98	-	-
Ours	<b>4.51</b>	2.25	<b>29.76</b>	3.43	<b>34.78</b>

● I love the idea of this place but I bought agroupon and you have to sign in on line within 30 days or if wo n't let you and they never answer the phone or return phone calls or email and when you go by no one is there I do n't know how they keep running specials I suggest do n't by a group on and the instructors are n't very pleasant to be around good luck I had to contactgroupon to get my money back to purchase another if this happens to you group on is wonderful they will do what it is you want they will even contact tough lotus if you want .

FIGURE 2 – Exemples de visualisation de l'attention sur une critique de restaurant et de film

TABLE 3 – Sentiments et émotions dans les top 5, 10 ou 20% des scores d'attention calculés

Dataset	Top 5%	Top 10%	Top 20%
IMDB	88.11%	97.30%	99.65%
Yelp-bi	51.86%	78.81%	94.37%
Yelp-Full	64.11%	84.65%	95.47%
Amazon-bi	55.71%	80.18%	94.65%
Amazon-Full	57.67%	81.02%	94.84%

tion, la taille et la position des mots sont déterminées par le score d'attention donné dans la section 3.1 . Pour les nuages de mots mixtes, nous combinons les deux informations. Les couleurs des mots sont choisies pour augmenter le contraste entre les mots de tailles différentes pour une meilleure visualisation. Ces couleurs ne représentent aucune information supplémentaire.

Chaque participant a répondu à 30 questions, 10 questions pour chaque modalité. La question posée a consisté à deviner la polarité des critiques (positives ou négatives) en se basant sur les nuages de mots. Nous informons le participant si le nuage de mots provient d'un commentaire de film, de restaurant ou d'un autre produit. La sélection des questions est randomisée pour chaque type de question et est différente pour chaque participant. Le nombre total de participants est de 85. Il s'agit d'étudiants inscrits dans des cursus en informatique. Pour chaque type de question, nous calculons l'index moyen de facilité  $F_i$  comme suit :

$$F_i = 100 * \frac{\bar{x}_i - x_{min}^i}{x_{max}^i - x_{min}^i} \quad (8)$$

où  $\bar{x}_i$  est le score moyen par question de type  $i$ ,  $x_{min}^i$  et  $x_{max}^i$  sont les scores minimum et maximum pour une question de type  $i$ . Le score par question donne une in-

dication sur la facilité avec laquelle la question est posée. Nous calculons également l'écart-type de tous les scores pour chaque question. De plus, nous calculons l'indice de discrimination  $D_i$  comme :

$$D_i = 100 * corr(x_i, T) \quad (9)$$

qui correspond au pourcentage de corrélation entre les scores de chaque type de question et le score total par participant. Le tableau 4 présente les résultats du sondage sur les nuages de mots pour les trois types de questions.

TABLE 4 – Résultats de l'enquête sur les nuages de mots

Types de Question	Index de facilité	Déviati on Standard	Index de discrimination
Attention	72.12%	43.90%	7.70%
Lexicon	58.90%	49.86%	-16.10%
Mixte	68.27%	49.19%	6.70%

Ces résultats restent préliminaires. Ils montrent que deviner le sentiment sans le texte complet est une tâche difficile. Cependant, les questions basées sur les nuages de mots construits par les mécanismes d'attention sont plus faciles que les autres avec un indice de facilité moyen de 72.12% et un écart-type de 43.90%. Par ailleurs, nous notons une corrélation positive entre l'accord et l'indice de discrimination. Nous pouvons en conclure que la tâche devient plus difficile en utilisant uniquement le lexique. Toutefois, les nuages construits à partir des scores d'attention captent des arguments qui ne sont pas centraux à la tâche et ne mettent pas en avant les expressions de négation et adversatives qu'il serait intéressant de repérer et de mettre en relief dans cette visualisation. D'autres formats de visualisation pourraient également être explorés.

## 5.4 Discussions

La méthode d'interprétabilité proposée, basée sur les mécanismes d'attention, permet de visualiser un résumé des

statistiques (taille et ordre) des attributs (les mots) et de leur impact sur les prédictions du modèle (intensité de la couleur). Cette méthode d'explication est spécifique à l'algorithme d'apprentissage utilisé qui fournit l'information de l'attention et ne rentre donc pas dans la catégorie des modèles d'interprétation *a posteriori*, qui peuvent être utilisés avec n'importe quel autre algorithme d'apprentissage. Un avantage de notre méthode est qu'elle est expressive car la représentation visuelle retenue permet de structurer les explications. Comme les auteurs de la méthode LIME [7], nous avons pu remarquer que les utilisateurs apprécient les explications qu'ils sélectionnent par eux-mêmes. En effet, ils ne s'attendent pas à une liste finie d'explications de la prédiction mais plutôt au choix d'une ou plusieurs explications dans une liste de choix, ce qui est possible en piochant des mots-clés dans le nuage. Toutefois, une limite est que l'interprétation proposée est locale et se limite à une prédiction unique alors que d'autres approches dites globales vont concerner tout le modèle d'apprentissage et notamment le choix des paramètres.

## 6 Conclusions et perspectives

Nous pouvons conclure de cette étude que l'idée de l'apprentissage par transfert est très efficace pour une application de classification de sentiments en terme de taux d'erreur. Elle fonctionne mieux que les modèles classiques d'apprentissage basés sur les word embeddings. De plus, l'ajout d'un mécanisme d'auto-attention a un impact direct sur les performances de ces modèles. Le modèle proposé a été évalué sur cinq jeux de données de la littérature. Nos expériences montrent des résultats compétitifs par rapport aux modèles basés sur l'attention à la pointe de l'état de l'art dont ULMFit. Pour finir, même si la visualisation proposée peut facilement être améliorée et que les expérimentations restent préliminaires, nous avons démontré que la visualisation des scores d'attention présentés sous la forme de nuages de mots, impacte positivement les perceptions liées à l'interprétabilité des utilisateurs.

Dans les travaux futurs, une nouvelle expérimentation, cette fois comparative, devra être menée pour savoir si l'ajout d'une information comme la visualisation proposée permet à un utilisateur d'avoir confiance dans le modèle et si cela impacte sa prise de décision. Nous allons chercher à améliorer l'interprétation locale d'une prédiction. Par exemple, il a été démontré que les humains apprécient les explications contrastives, qui permettent de comprendre pourquoi une prédiction a été faite à la place d'une autre. Pour un texte pour lequel nous cherchons à déterminer la polarité, faire évoluer notre visualisation pour colorer d'une couleur les mots-clés ayant contribué au choix du sentiment prédit et d'une autre couleur les mots-clés ayant contribué au sentiment inverse. Nous pourrions également travailler sur les techniques de résumé automatique, apparié avec des méthodes d'analyse de sentiments basées sur les facettes afin d'améliorer la visualisation.

Nous prévoyons de travailler sur une interprétation globale du modèle. Tout d'abord, nous pourrions définir la "représentativité" des explications. L'intuition est la suivante. Les bonnes explications sont souvent générales. La représentativité d'une explication correspondrait au nombre de textes couverts par cette explication. Nous pourrions également détecter les explications fréquemment associées via les mots-clés se retrouvant dans les nuages de mots. Ces deux informations pourrait aider l'utilisateur à comprendre sur quoi le classifieur s'appuie généralement pour prendre sa décision. Par opposition, nous pourrions repérer les explications "anormales" (peu fréquentes). Si l'une des caractéristiques d'entrée d'une prédiction est anormale (un groupe de mots-clés rares), elle devrait être présentée dans l'explication du modèle afin d'identifier les cas particuliers. Pour conclure, un outil permettant de naviguer dans les explications basé sur la recherche d'explications représentatives, fréquemment associées et anormales permettrait d'améliorer l'interprétabilité globale du modèle.

## Remerciement

Nous tenons à remercier la Région Occitanie et la Communauté d'Agglomération de Béziers Méditerranée pour le financement de la thèse de Waleed Ragheb.

## Références

- [1] Anderson, P., He, X., Buehler, C., Teney, D., Johnson, M., Gould, S., Zhang, L. : Bottom-up and top-down attention for image captioning and VQA. CoRR **abs/1707.07998**, 6077–6086 (2017), <http://arxiv.org/abs/1707.07998>
- [2] Bahdanau, D., Cho, K., Bengio, Y. : Neural machine translation by jointly learning to align and translate. In : International Conference on Learning Representations (ICLR) (2014)
- [3] Devlin, J., Chang, M., Lee, K., Toutanova, K. : BERT : pre-training of deep bidirectional transformers for language understanding. CoRR **abs/1810.04805** (2018)
- [4] Du, J., Gui, L., He, Y., Xu, R. : A convolutional attentional neural network for sentiment classification. In : International Conference on Security, Pattern Analysis, and Cybernetics (SPAC). pp. 445–450 (2017)
- [5] Herlocker, J.L., Konstan, J.A., Riedl, J. : Explaining collaborative filtering recommendations. In : ACM Conference on Computer Supported Cooperative Work. pp. 241–250. CSCW '00, ACM, New York, NY, USA (2000). <https://doi.org/10.1145/358916.358995>, <http://doi.acm.org/10.1145/358916.358995>
- [6] Howard, J., Ruder, S. : Universal language model fine-tuning for text classification. In : 56th Annual Meeting of the Association for Computational Linguistics (Volume 1 : Long Papers). pp. 328–339 (2018)

- [7] Hu, L., Chen, J., Nair, V.N., Sudjianto, A. : Locally interpretable models and effects based on supervised partitioning (LIME-SUP). *CoRR* **abs/1806.00663** (2018)
- [8] Letarte, G., Paradis, F., Giguère, P., Laviolette, F. : Importance of self-attention for sentiment analysis. In : *EMNLP Workshop BlackboxNLP : Analyzing and Interpreting Neural Networks for NLP*. pp. 267–275 (2018)
- [9] Lin, Z., Feng, M., dos Santos, C.N., Yu, M., Xiang, B., Zhou, B., Bengio, Y. : A structured self-attentive sentence embedding. In : *International Conference on Learning Representations (ICLR)* (2017)
- [10] Ma, Y., Peng, H., Cambria, E. : Targeted aspect-based sentiment analysis via embedding common-sense knowledge into an attentive lstm. In : *AAAI Conference on Artificial Intelligence*. pp. 5876–5883 (2018)
- [11] Maas, A.L., Daly, R.E., Pham, P.T., Huang, D., Ng, A.Y., Potts, C. : Learning word vectors for sentiment analysis. In : *49th Annual Meeting of the Association for Computational Linguistics : Human Language Technologies - Volume 1*. pp. 142–150. *HLT '11*, Association for Computational Linguistics, Stroudsburg, PA, USA (2011)
- [12] Mäntylä, M.V., Graziotin, D., Kuuttila, M. : The evolution of sentiment analysis - a review of research topics, venues, and top cited papers. *Computer Science Review* **27**, 16–32 (2018)
- [13] Merity, S., Keskar, N.S., Socher, R. : Regularizing and optimizing LSTM language models. In : *International Conference on Learning Representations (ICLR)* (2018)
- [14] Merity, S., Xiong, C., Bradbury, J., Socher, R. : Pointer sentinel mixture models. *CoRR* **abs/1609.07843** (2016), <http://arxiv.org/abs/1609.07843>
- [15] Mohammad, S.M. : Word affect intensities. In : *11th Edition of the Language Resources and Evaluation Conference (LREC-2018)*. Miyazaki, Japan (2018)
- [16] Mokhtari, S., Li, T., Xie, N. : Context-sensitive neural sentiment classification. In : *2018 IEEE International Conference on Information Reuse and Integration (IRI)*. pp. 293–299 (2018)
- [17] Peters, M., Neumann, M., Iyyer, M., Gardner, M., Clark, C., Lee, K., Zettlemoyer, L. : Deep contextualized word representations. In : *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics : Human Language Technologies, Volume 1 (Long Papers)*. pp. 2227–2237. Association for Computational Linguistics, New Orleans, Louisiana (Jun 2018). <https://doi.org/10.18653/v1/N18-1202>, <https://www.aclweb.org/anthology/N18-1202>
- [18] Su, J., Zeng, J., Xiong, D., Liu, Y., Wang, M., Xie, J. : A hierarchy-to-sequence attentional neural machine translation model. In : *IEEE/ACM Trans. Audio, Speech & Language Processing*. vol. 26, pp. 623–632 (2018)
- [19] Tutek, M., Šnajder, J. : Iterative recursive attention model for interpretable sequence classification. In : *EMNLP Workshop BlackboxNLP : Analyzing and Interpreting Neural Networks for NLP*. pp. 249–257 (2018)
- [20] Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A.N., Kaiser, L.u., Polosukhin, I. : Attention is all you need. In : *Advances in Neural Information Processing Systems 30*, pp. 5998–6008. Curran Associates, Inc. (2017)
- [21] Vellido, A., Martín-Guerrero, J.D., Lisboa, P.J.G. : Making machine learning models interpretable. In : *European Symposium on Artificial Neural networks, computational intelligence and machine learning* (2012)
- [22] Voulodimos, A., Doulamis, N., Doulamis, A., Proto-papadakis, E. : Deep learning for computer vision : A brief review. In : *Computational Intelligence and Neuroscience*. vol. 2018, pp. 1–13 (02 2018)
- [23] Wan, L., Zeiler, M., Zhang, S., Cun, Y.L., Fergus, R. : Regularization of neural networks using dropconnect. In : *30th International Conference on Machine Learning. Proceedings of Machine Learning Research*, vol. 28, pp. 1058–1066. PMLR, Atlanta, Georgia, USA (17–19 Jun 2013), <http://proceedings.mlr.press/v28/wan13.html>
- [24] Wang, S., Huang, M., Deng, Z. : Densely connected cnn with multi-scale feature attention for text classification. In : *IJCAI* (2018)
- [25] Yang, Z., Yang, D., Dyer, C., He, X., Smola, A., Hovy, E. : Hierarchical attention networks for document classification. In : *Conference of the North American Chapter of the Association for Computational Linguistics : Human Language Technologies*. pp. 1480–1489. Association for Computational Linguistics (2016)
- [26] Young, T., Hazarika, D., Poria, S., Cambria, E. : Recent trends in deep learning based natural language processing [review article]. In : *IEEE Computational Intelligence Magazine*. vol. 13, pp. 55–75 (2018)
- [27] Zhang, X., Zhao, J., LeCun, Y. : Character-level convolutional networks for text classification. In : *Advances in Neural Information Processing Systems 28*. pp. 649–657 (2015)



# Méthodes de représentation de la langue pour l'analyse syntaxique multilingue

M. Scholivet

Aix-Marseille Univ, Université de Toulon, CNRS, LIS, Marseille  
manon.scholivet@lis-lab.fr

## Résumé

*L'existence de modèles universels pour décrire la syntaxe des langues a longtemps été débattue. L'apparition de ressources tel le World Atlas of Language Structures et les corpus des Universal Dependencies rend possible l'étude d'une grammaire universelle pour l'analyse syntaxique en dépendances. Notre travail se concentre sur l'étude de différentes représentations des langues dans des systèmes multilingues appris sur des corpus arborés de 37 langues. Nos tests d'analyse syntaxique montrent que représenter la langue dont est issu chaque mot permet d'obtenir de meilleurs résultats qu'en cas d'un apprentissage sur une simple concaténation des langues. En revanche, l'utilisation d'un vecteur pour représenter la langue ne permet pas une amélioration évidente des résultats dans le cas d'une langue n'ayant pas du tout de données d'apprentissage.*

## Mots Clef

Analyse syntaxique multilingue, Représentations de la langue, Traits typologiques, WALs, UD.

## Abstract

*The existence of universal models to describe the syntax of languages has been debated for decades. The availability of resources such as the World Atlas of Language Structures and the Universal Dependencies treebanks makes it possible to study the plausibility of universal grammar from the perspective of dependency parsing. Our work consists in studying different language representations in multilingual systems learned on treebanks in 37 languages. We show that representing the language of each word leads to better results compared to simply concatenate all the languages. However, using a vector to represent the language does not increase obviously the results in zero-shot learning cases.*

## Keywords

Multilingual parsing, Language representations, Typological features, WALs, UD.

## 1 Introduction

L'analyse syntaxique en dépendances requiert, comme de nombreuses autres tâches en TAL, de grandes quantités de données pour apprendre la syntaxe de la langue. Cer-

taines langues comme l'anglais bénéficient d'une grande quantité de données annotées. En revanche, pour de nombreuses langues, il n'en existe actuellement pas. L'analyse syntaxique multilingue est une solution possible pour tenter de tirer profit des langues bien dotées pour apprendre des informations communes entre les langues. Deux problèmes importants se posent alors : comment représenter les données, et comment représenter la langue ?

Avec l'apparition des Universal Dependencies (UD)<sup>1</sup>, l'utilisation d'annotations universelles pour la syntaxe permet en partie de répondre au premier problème. En effet, UD propose un ensemble de relations de dépendances, de parties du discours et de traits morphologiques qui se veulent communs à toutes les langues [12]. Les corpus arborés de UD sont disponibles pour de nombreuses langues, et sont basés sur un guide d'annotation commun.

Bien que les UD permettent un certain partage entre les langues au niveau des traits universels proposés, il n'existe pas de lexique commun à toutes les langues. Le choix de l'utilisation d'un analyseur delexicalisé [19] a été fait pour ce travail. Cette technique consiste à ignorer le lexique lors de l'entraînement de l'analyseur. Cet appauvrissement des données conduit à des analyseurs moins précis, mais offre une solution simple au problème du lexique<sup>2</sup>.

Le problème de représentation de la langue est le cœur de cet article. Trois méthodes de représentations seront présentées. En plus du simple identifiant de la langue, nous verrons comment tenter d'extraire les paramètres syntaxiques, premièrement à partir des données, ensuite en utilisant le World Atlas of Language Structures (WALS)<sup>3</sup> [4]. Nos contributions consistent à comparer ces différentes méthodes et à évaluer leurs effets dans un cadre d'entraînement multilingue, où le corpus d'entraînement est composé de corpus arborés de 37 langues, puis dans un cadre de *zero-shot*, c'est-à-dire pour une langue n'ayant pas de données d'entraînement [1, 6].

Après la présentation de l'état de l'art (Sec. 2), nous décrirons les UD (Sec. 3), nos représentations des langues (Sec. 4), en particulier nous présenterons en détail le WALs (Sec. 4), puis présenterons l'analyseur utilisé (Sec. 5). Le

1. <http://universaldependencies.org>

2. Une alternative non explorée dans ce travail et que nous comptons explorer à l'avenir est l'utilisation de plongements de mots multilingues [2]

3. <https://wals.info/>

cadre expérimental (Sec. 6) précédera nos résultats (Sec. 7) et nos conclusions (Sec. 8).

## 2 État de l'art

Ce travail est à l'intersection de trois tendances dans la littérature sur l'analyse syntaxique en dépendances multilingue. La première est le *transfer parsing*, qui consiste à apprendre un analyseur sur une langue (ou un ensemble de langues) puis à le tester sur une autre. La seconde est l'*analyse syntaxique delexicalisée*, qui a pour but de supprimer le lexique. La troisième et dernière tendance est l'utilisation de *ressources typologiques* telles que le WALs.

Le *transfer parsing* est une bonne solution lorsqu'il faut traiter des langues avec peu de ressources. [8] décrivent deux types de transfert : la première se base sur des corpus parallèles dont une des deux langues n'a pas, ou pas assez, de données d'entraînement pour apprendre un analyseur, alors que l'autre langue en a. La deuxième, les approches par transfert direct, reposent sur les similarités entre les langues et ne nécessitent pas de corpus parallèles. Par exemple, [7] proposent un analyseur pour l'irlandais entraîné d'abord sur une autre langue, puis appliqué à l'irlandais. Étonnamment, l'indonésien est la langue donnant les meilleurs résultats, bien qu'elles n'appartiennent pas à la même famille de langues. L'hypothèse des auteurs serait que les dépendances très distantes sont mieux représentées dans l'indonésien que dans les autres langues testées.

Les langues pauvrement dotées peuvent avoir une petite quantité de données d'entraînement disponible. En concaténant les corpus d'entraînement de deux langues, on peut obtenir un analyseur bilingue pour vérifier si une amélioration est possible comparé à un analyseur monolingue [16]. Les méthodes de transfert direct et les analyseurs bilingues sont proches de cet article, puisqu'on retrouve cette idée de concaténation des corpus d'entraînement. Cependant, nous combinons les corpus de plus de langues (environ 40) et incluons des méthodes de représentation des langues. La combinaison de corpus de multiples langues pour l'entraînement d'un analyseur est facilitée par les récentes avancées sur les standards multilingues et les ressources disponibles, en particulier grâce aux Universal Dependencies pour la syntaxe en dépendances [12]. La recherche sur l'analyse syntaxique multilingue est fortement stimulée par des initiatives comme les campagnes d'évaluation CoNLL 2017 et 2018, sur l'analyse syntaxique en dépendances fortement multilingue à partir de textes bruts [18, 17].

Les analyseurs delexicalisés ignorent la forme superficielle et les lemmes des mots lors de l'analyse d'une phrase, utilisant des traits plus abstraits comme les étiquettes de parties de discours (POS). L'utilisation d'analyseurs delexicalisés est particulièrement pertinente pour l'apprentissage d'analyseurs multilingues, puisque les langues partagent généralement peu de leurs lexiques. L'approche proposée par [19] consiste à adapter un analyseur pour une nouvelle langue en utilisant soit un corpus parallèle, soit un analyseur delexicalisé. Cette méthode peut être utilisée pour construire

rapidement un analyseur si la langue source et la langue cible sont suffisamment proches.

De plus, les traits typologiques comme ceux présents dans le WALs donnent une information sur la structure des langues [4]. Ces traits peuvent permettre à l'analyseur multilingue d'apprendre des caractéristiques communes à plusieurs langues. [9] et [20] utilisent l'ensemble des traits issus de la catégorie de l'Ordre des mots du WALs qui sont disponibles pour leurs langues. [14] utilisent plutôt les traits qu'ils jugent pertinents dans les diverses catégories (pas uniquement celle concernant l'ordre des mots).

[15] utilisent une méthode de transfert multilingue delexicalisée, montrant en quoi le partage de paramètres basé sur des traits typologiques et l'appartenance à une famille de langues peut être utilisé dans un analyseur en dépendances discriminant. Les traits typologiques choisis sont basés sur ceux utilisés par [9], en retirant deux traits qu'ils ne considéraient pas utiles.

Les travaux les plus proches du notre concatènt des corpus pour entraîner un analyseur multilingue [1]. Les auteurs utilisent un analyseur à transitions S-LSTM similaire au notre (en dehors de l'utilisation de récurrence) entraîné sur un ensemble de traits lexicaux incluant des plongement de mots multilingues, des clusters de Brown, et des étiquettes POS<sup>4</sup> détaillées, alors que nous utilisons seulement des POS plus grossières et des traits morphologiques dans un cadre delexicalisé. Ils utilisent également un vecteur de représentation de la langue encodé en one-hot, l'ensemble de 6 traits issus du travail de [9], et testent en plus l'utilisation de la matrice complète du WALs. Nous avons testé ces deux premières représentations, ainsi qu'un autre vecteur de dimension 22 issu du WALs également. Leurs expériences ont été réalisées sur sept langues richement dotées alors que nous avons mené les nôtres sur un échantillon beaucoup plus grand de 40 langues. Bien que [1] ont montré que, dans un cadre lexicalisé, la concaténation de corpus peut donner des résultats similaires à des analyseurs syntaxiques monolingues, les origines et les limites de ces gains restent floues. Nous explorons entre autres des pistes pour évaluer les avantages des caractéristiques typologiques dans un analyseur delexicalisé.

## 3 Universal Dependencies

La cohérence de l'annotation des données entre les langues est un problème majeur sur les tâches d'analyse multilingue, la plupart des corpus étant annotés en utilisant différents guides d'annotation et jeux d'étiquettes. L'initiative des Universal Dependencies (UD) a pour but de créer des corpus arborés consistants entre les langues, facilitant ainsi les analyses lors de travaux sur la langue.

Nous utilisons la version 2.0 de UD pour nos corpus d'entraînement de développement<sup>5</sup>, et les corpus de tests de la campagne d'évaluation CoNLL 2017<sup>6</sup>. 64 corpus arbo-

4. Partie du discours

5. <http://hdl.handle.net/11234/1-1983>

6. <http://hdl.handle.net/11234/1-2184>

rés de UD en 45 langues sont disponibles pour l’entraînement et le développement. Cependant, ces corpus sont de taille très variable (kazakh : 529 mots, tchéque : 1 1842 867 mots). Les corpus de tests contiennent au minimum 10 000 mots par langue et sont disponibles pour 49 langues. 4 langues n’ont pas de corpus d’entraînement correspondant. Les analyseurs delexicalisés, entraînés à partir des corpus de UD, prennent en entrée les parties du discours universelles (UPOS) et les traits morphologiques (FEAT) et prédisent les étiquettes de l’arbre de dépendances (incluant des sous-relations syntaxiques pouvant être spécifique au langage (p. ex., *acl :relcl*)). Les traits morphologiques sont renseignés pour presque tous les corpus, mais présentent de grandes différences. Par conséquent, nous avons fait le choix de ne garder que les traits les plus fréquents, qui apparaissent dans au moins 28 langues. De plus, les traits morphologiques sont représentés comme une liste de paires (*clef, valeur*) que nous avons séparé, afin que chaque paire soit considérée indépendamment, produisant ainsi un ensemble de 16 traits morphologiques par mot.

## 4 Les différentes représentations de la langue

Dans ce travail, nous utilisons trois méthodes de représentation de la langue. La première (appelée ID) consiste en un simple identifiant de la langue, sous la forme d’un vecteur one-hot de dimension 37 en entrée de notre analyseur. Les deux autres méthodes sont décrites ci-dessous.

**Apprentissage à partir des données** Une méthode pour représenter la langue consiste à apprendre un vecteur à partir de nos données d’entraînement, utilisant des statistiques sur les dépendances vues dans les corpus d’entraînement. Deux vecteurs seront appris de cette façon.

Le premier, appelé  $W_d$ , consiste à encoder, pour chaque étiquette des dépendances syntaxiques, la distance moyenne qui sépare les deux éléments de la relation dans une langue. Chaque dépendance  $d$  sera encodée sur deux composantes du vecteur  $W_d$  : une pour le cas où  $d$  est une dépendance droite, et l’autre pour les dépendances gauches.

La deuxième représentation, appelée  $W_{df}$ , ajoute à  $W_d$  l’information sur la fréquence de chacune de ces dépendances. Ainsi, si il existe  $n$  étiquettes de dépendance,  $W_d$  sera de taille  $2 * n$  et  $W_{df}$  de taille  $4 * n$ . Un exemple de  $W_d$  et  $W_{df}$  est donné, dans la Table 1.

Ces vecteurs sont appris sur notre corpus d’entraînement (Sec. 6) et ne sont donc pas disponibles pour les langues n’ayant pas de corpus d’entraînement. Ils ne conviennent donc pas aux tests de type zero-shot, contrairement aux vecteurs issus du WALS présentés ci-dessous.

**World Atlas of Language Structures** Le World Atlas of Language Structures (WALS) est une base de données de propriétés structurelles (phonologiques, grammaticales et lexicales) réunie par 55 auteurs à partir de matériel descriptif comme par exemple des grammaires de référence. Cette base de données nous a permis d’associer à chaque

6,00	3,36	12,80	12,59	...	
L_acl	R_acl	L_advcl	R_advcl	...	
6,00	3,36	0,00005	0,02	12,80	...
L_acl	R_acl	freq_L_acl	freq_R_acl	L_advcl	...

TABLE 1 – Exemple de  $W_d$  (en haut) et  $W_{df}$  (en bas).  $L\_acl$  représente les dépendances gauches étiquetées *acl*, dont la distance moyenne entre les éléments est de 6.  $R\_acl$  représente la même chose, mais pour une dépendance droite.  $freq\_L\_acl$  est la fréquence des dépendances *acl* gauche dans la langue, et  $freq\_R\_acl$  est la fréquence des dépendances *acl* droite.

langue des corpus de UD un ensemble de traits décrivant des propriétés pertinentes pour l’analyse syntaxique.

Le WALS décrit 2 676 langues grâce à un ensemble de 192 traits, répartis entre 11 familles (p.ex. Phonologie, Ordre des mots...). Il peut ainsi être représenté via une matrice  $W$  de 2 676 lignes et 192 colonnes, où chaque cellule  $W(l, f)$  donne la valeur du trait  $f$  pour la langue  $l$ . Chaque ligne  $W(l)$  est le vecteur de trait de la langue  $l$ .

Cette matrice a été épurée puis complétée pour correspondre à nos conditions expérimentales. Pour commencer, nous n’avons gardé que les lignes correspondant aux 49 langues de notre corpus de test. En revanche, 4 langues de UD (le Vieux Slave (cu), le Gothique (got), le Grec Ancien (grc), et le Latin (la)) n’apparaissent pas dans le WALS et ont donc été mises de côté. On obtient alors une version réduite de  $W$  contenant 45 lignes.

Deux de nos représentations de la langue ont été extraites à partir du WALS. La première, appelée  $W_N$ , est basée sur les travaux de [9], qui ont sélectionné les 6 traits<sup>7</sup> de la famille de l’Ordre des mots qui étaient entièrement renseignés pour leurs 17 langues cibles. Ces traits couvrent des phénomènes tels que l’ordre verbe-objet ou adjectif-nom, et ont été largement discutés dans la littérature [15, 20, 1]. La matrice qui en résulte a 45 lignes (langues) pour 6 colonnes (traits). Cependant, le WALS est une matrice creuse, puisque certains traits ne sont pas renseignés pour certaines langues. C’est pourquoi nous avons fait le choix de ne garder que les langues pour lesquelles au plus la moitié du vecteur n’est pas renseignée, éliminant ainsi 5 langues de plus : le galicien (gl), le haut sorabe (hsb), le kazakh (kk), le slovaque (sk), et le ouïghour (ug). Nos expériences sont effectuées sur cet ensemble de 40 langues.

La deuxième représentation de la langue extraite du WALS, que l’on appellera  $W_{80}$ , est une version plus étendue de  $W_N$ . Nous nous sommes demandé s’il n’était pas dommage de se restreindre aux traits de [9], qui ne gardent que 6 traits. Nous incluons alors dans  $W_{80}$  tous les traits renseignés pour au moins 80% de nos 40 langues.

En plus des traits de la famille de l’Ordre des mots, nous avons également inclus ceux de la famille des Propositions simples<sup>8</sup>. Il en résulte une matrice de 40 lignes et 22 co-

7. Ces traits sont ceux identifiés par les codes 81A, 85A, 86A, 87A, 88A, 89A dans le WALS.

8. Nous avons également considéré la famille des Phrases Complexes,

	Romane	Germanique	Slave	Aléatoire
$W_N$	0.33	1.33	0.67	2.41
$W_{80}$	4.13	4.47	4.19	10.15

TABLE 2 – MID de chaque famille de langues comparée à l’aléatoire. Aléatoire est la moyenne de 50 000 familles de 6 langues.

lonnes, correspondant à 3 traits de la famille des Propositions simples (101A, 112A, 116A) et 19 traits de celle de l’Ordre des mots (81A, 82A, 83A, 85A, 86A, 87A, 88A, 89A, 90A, 92A, 94A, 95A, 96A, 97A, 144A, 143A, 143E, 143F, 143G).

Les matrices  $W_N$  et  $W_{80}$  obtenues ne sont cependant pas complètes : elles contiennent respectivement 4 et 35 valeurs non-remplées, que nous avons rempli automatiquement. Chaque matrice  $W$  (abréviation pour  $W_N$  et  $W_{80}$ ) permet de comparer deux langues  $l_1$  et  $l_2$  de manière simple en utilisant la distance de Hamming<sup>9</sup> entre leur vecteur  $W(l_1)$  et  $W(l_2)$ , noté  $d(l_1, l_2)$ . Pour remplacer les valeurs manquantes, nous avons sélectionné, pour chaque langue  $l_1$  contenant au moins une valeur non-remplée (‘?’), la valeur correspondante dans le vecteur de la langue  $l_2$  la plus proche qui soit entièrement renseignée, où  $l_2 = \arg \min_{l_i \mid \text{“?”} \notin W(l_i)} d(l_1, l_i)$ .

Les matrices  $W_N$  et  $W_{80}$  ne fournissent qu’une description partielle des langues, fortement biaisée en faveur de l’analyse syntaxique et ignorant les autres aspects (p.ex. la phonologie). Néanmoins, il est tentant de comparer les distances de langues appartenant aux mêmes familles typologiques dans ce mode de représentation. Pour cela, nous nous sommes intéressés à 3 familles présentes dans notre ensemble de 40 langues : les langues Romanes (6 langues), les langues Germaniques (6 langues) et les langues Slaves (7 langues). Nous avons alors calculé la proximité des vecteurs de ces langues. Nous définissons la distance interne moyenne (MID) d’un ensemble de langues  $L = \{l_1, \dots, l_n\}$ , comme la moyenne des distances de chaque paire dans  $L$  :

$$MID(L) = \frac{1}{n^2 - n} \sum_{\substack{(l_i, l_j) \in L \times L \\ i \neq j}} d(l_i, l_j)$$

Nous avons calculé le MID de chaque famille de langues, et l’avons comparé au MID d’ensembles aléatoires de 6 langues (ce qui correspond au nombre de langues des familles Germanique et Romane). Les résultats de la Table 2 montrent clairement que les vecteurs du WALS permettent de capturer des similarités au sein d’une famille de langues, puisque le MID des vecteurs de langues d’une même famille est nettement inférieur au simple hasard.

mais aucun trait ne dépassait le seuil des 80%.

9. Le nombre de dimensions pour lesquelles les valeurs diffèrent.

## 5 Analyseur

L’analyseur utilisé dans nos expériences est un analyseur par transition de type *arc-eager* [11], entraîné avec un oracle dynamique [5]. La prédiction des transitions est faite par un perceptron multi-couches (MLP) similaire au système de [3], consistant en une couche d’entrée, une couche cachée et une couche de sortie. Deux ensembles de traits entièrement delexicalisés ont été définis pour la prédiction : BASIQUE et ÉTENDU. BASIQUE est un ensemble classique composé de 9 traits liés au POS, 7 traits syntaxiques, 32 traits morphologiques et un trait de distance (la distance entre la tête et le dépendant).<sup>10</sup> L’ensemble ÉTENDU ajoute à BASIQUE de nouveaux traits correspondant aux vecteurs du WALS  $W_N$  et  $W_{80}$ , et/ou l’identifiant de la langue. Chaque trait est associé à un plongement de taille 3, initialisé à zéro. L’identifiant de la langue (issu de la représentation ID) est un vecteur one-hot de dimension 37 (correspondant aux 37 langues uniques). La couche d’entrée du MLP correspond à la concaténation des plongements des différents traits, dont les dimensions varient de 396 à 465 selon la configuration (avec ou sans vecteurs de langue  $W_N$  et  $W_{80}$ , ou de l’identifiant de la langue ID). La couche de sortie est composée de 263 neurones, correspondant au nombre de transitions que l’analyseur peut prédire. La couche cachée est de taille 1 000, avec un *dropout* durant l’entraînement de 0.4, le nombre d’itérations est égal à 10, la fonction d’activation est la fonction ReLU, la fonction objectif est un softmax de vraisemblance négative, et l’algorithme d’apprentissage est AMSgrad, utilisant les paramètres par défaut de Dynet [10].<sup>11</sup>

À chaque étape du processus d’analyse syntaxique, l’analyseur prédit une action à réaliser, pouvant aboutir à la création d’une nouvelle dépendance entre deux mots de la phrase. La prédiction des actions est basée sur la valeur des traits donnés au MLP. Dans la configuration BASIQUE, ces traits décrivent différents aspects du gouverneur, du dépendant, et du contexte. Par exemple, si la tête est un verbe et que le dépendant est un nom situé avant le verbe, une dépendance sujet aura une forte probabilité d’être prédite pour les langues utilisant majoritairement l’ordre sujet-verbe (SV). Avec l’ensemble de traits ÉTENDU, l’utilisation de l’ordre SV pour une langue est explicite. Le MLP a donc la possibilité de combiner une *configuration de phrasique* (p.ex., un nom avant un verbe) avec une *configuration de la langue* (p.ex., la langue est SV) quand il prédit une action. Les langues partageant un trait dans  $W$  auront la possibilité de générer la même prédiction pour une configuration phrasique correspondant à ce trait (p.ex. le nom précédant le verbe et la langue est de type SV).<sup>12</sup>

10. Nos corpus, le code source de l’analyseur, les fichiers de configuration et les matrices issues du WALS seront disponibles.

11. Les valeurs des hyperparamètres ont été définies dans des conditions similaires à  $\Sigma W_N$ , cf. Section 6).

12. Notre analyseur ne peut pas prédire d’arbre non-projectif. La présence d’une dépendance non-projective génère systématiquement une erreur d’analyse lors de la phase de test. Le taux moyen de projectivité du corpus de test est égal à 1%, avec un écart-type de 1% pour les 40 langues.

## 6 Cadre expérimental

**Corpus** Nos expériences ont été réalisées sur les données de la campagne d'évaluation CoNLL 2017 [18], en utilisant la tokenisation de référence et en ignorant les contractions (p.ex. *du=de+la*). Nos modèles sont évalués individuellement sur chacune des 40 langues pour lesquelles nous avons un vecteur  $W(l)$  (Sec. 4), en utilisant les corpus de test de la campagne d'évaluation pour faciliter la comparaison avec les travaux similaires. Le corpus de test pour chaque langue est obtenu en faisant la concaténation de tous les corpus de test disponibles pour cette langue.

Trois langues n'ont pas de corpus d'entraînement ou de développement (bxr, kmr, sme). L'entraînement et le développement se font sur des corpus multilingues (ML) dérivés des 37 langues de UD restantes, que l'on nommera TRAIN-ML et DEV-ML. La taille des corpus de UD peut fortement varier d'une langue à l'autre, allant de 529 mots pour le Kazakh (kk) à 1 842 867 pour le Tchèque (cs). Ainsi, concaténer simplement tous les corpus pour constituer TRAIN-ML et DEV-ML sur-représenterait certaines langues et introduirait un biais en leur faveur. C'est pourquoi nous avons décidé d'équilibrer le nombre de tokens de TRAIN-ML et DEV-ML entre les langues.

La construction des corpus de développement et d'entraînement se fait en deux étapes. Tout d'abord la totalité des données des corpus disponibles (entraînement et développement) est divisée en deux à hauteur de 10% pour un sous-corpus de développement et le reste pour l'entraînement. Ensuite, pour chacun des sous-corpus précédents, on sélectionne de façon aléatoire des phrases pour chaque langue jusqu'à ce que les corpus finaux atteignent une certaine limite en terme de tokens (respectivement 2 000 pour le DEV et 20 000 pour le TRAIN). Enfin les données ainsi sélectionnées sont mélangées (pour chacun des corpus) afin d'éviter de conserver l'ordre des langues. Avec cette procédure, la même phrase peut apparaître plusieurs fois. Néanmoins, cette approche garantit une représentation équilibrée de chaque langue dans TRAIN-ML et DEV-ML.

**Métrique** La qualité des arbres prédits est évaluée par une mesure standard pour l'analyse syntaxique en dépendance : le score d'attachement de l'étiquette (LAS).<sup>13, 14</sup> On donne le LAS par langue, ainsi que le MACRO-LAS, qui est la macro-moyenne du LAS de toutes les langues qui ont un corpus d'entraînement. Cette mesure est indépendante de la taille du corpus de test de chaque langue, et n'est pas biaisée en faveur des langues sur-représentées dans l'ensemble de test.

<sup>14</sup> corpus (pour 20 langues) ont un taux de non-projectivité inférieur à 1%, et le taux maximum est de 8% pour le corpus du néerlandais (corpus Lassysmal). Nous avons fait des tests utilisant une transformation en arbre pseudo-projectif [13], mais l'impact sur les résultats étant négligeable, nous avons choisi de garder l'algorithme projectif original.

<sup>13</sup>. Nous avons omis le score UAS car le UAS et le LAS sont étroitement corrélés ( $r = 0.98$ ).

<sup>14</sup>. Nous avons utilisé le script d'évaluation de la campagne d'évaluation CoNLL 2017.

**Configuration d'entraînement** Nos expériences sur plusieurs paires (corpus d'entraînement, représentation de la langue) sont désignées par les codes suivants :

$L$  : Corpus monolingue. Le corpus d'entraînement de la langue  $l$  consiste simplement à prendre toutes les phrases de la langue  $l$  dans TRAIN-ML. 37 analyseurs avec la configuration BASIQUE ont été entraînés, un pour chaque langue. Cette configuration correspond à la situation standard des expériences en analyse syntaxique : entraînement et test sur une même langue.

$\Sigma$  : Corpus multilingue. Un analyseur est entraîné sur l'intégralité de TRAIN-ML, sans indication sur la langue. Le modèle est delexicalisé, le corpus ne contient donc que les étiquettes POS de référence, les traits morphologiques de référence, et les relations syntaxiques à apprendre.

$\Sigma$  ID : Corpus multilingue + ID de la langue. Un analyseur entraîné avec la configuration ÉTENDU sur TRAIN-ML, en utilisant, attaché à chaque mot, l'identifiant de la langue.

$\Sigma W_N, \Sigma W_{80}$  : Corpus + WALS. Deux analyseurs entraînés avec l'ensemble de traits ÉTENDU, sur TRAIN-ML. L'entraînement de  $\Sigma W_N$  (resp.  $W_{80}$ ) ajoute à chaque mot de  $\Sigma$  un vecteur  $W(l)$  du WALS, qui correspond à la langue du mot. Un modèle entraîné pour  $\Sigma W_N$  (resp.  $W_{80}$ ).

$\Sigma W_d, \Sigma W_{df}$  : Corpus + vecteurs de données. Deux analyseurs entraînés avec l'ensemble de traits ÉTENDU, sur TRAIN-ML, avec  $W_d$  (respectivement  $W_{df}$ ) attaché à chaque mot. Cet ajout est fait de la même façon que pour les tests  $\Sigma W_N$  et  $\Sigma W_{80}$ .

$\bar{L}$  : Zero-shot. Dans cette configuration, 37 corpus d'entraînements sont dérivés de TRAIN-ML : pour chaque langue  $l$ , un corpus d'entraînement est construit à partir de toutes les phrases de TRAIN-ML, sauf celles appartenant à la langue  $l$ . Cette configuration représente la situation où un analyseur dans la configuration BASIQUE est entraîné pour une langue pour laquelle aucune donnée d'entraînement n'est disponible, comme dans les méthodes de transfert direct.

$\bar{L} W_{80}$  :  $\bar{L} W_{80}$  ajoute à  $\bar{L}$  l'indication de la langue de chaque mot à travers son vecteur du WALS, de la même manière que  $\Sigma W_{80}$  faisait avec  $\Sigma$ .

## 7 Résultats et analyse

Nos expériences ont été réalisées dans les configurations décrites ci-dessus. Le LAS est donné pour chaque langue, ainsi que la macro-moyenne du LAS (MACRO), dans la Table 3. Nous commentons ci-dessous les résultats pour  $L$ , et comparons les résultats de certaines expériences (voir Table 4).

$L$  : Les résultats de l'expérience  $L$  montrent une importante variation des performances selon les langues. Le LAS varie de 46,78 pour le turc à 81,44 pour l'italien. Plus d'expériences seraient nécessaires pour expliquer les raisons d'une telle variabilité, mais n'entrent pas dans le cadre de ce travail. Nous pouvons tout de même émettre certaines hypothèses. Tout d'abord, certaines spécificité des

$L$	$\Sigma$	$\Sigma ID$	$\Sigma W_N$	$\Sigma W_{80}$	$\Sigma W_d$	$\Sigma W_{df}$	$\bar{L}$	$\bar{L} W_{80}$	Lang.
65,89	60,59	64,04	63,15	64,38	64,08	63,29	27,50	34,34	ar
78,59	74,32	79,25	76,26	77,47	78,28	78,94	67,05	63,73	bg S
77,18	72,76	76,63	73,03	76,27	76,90	77,17	70,48	68,88	ca R
68,92	68,01	69,41	68,72	69,61	69,75	69,67	62,08	59,47	cs S
73,62	67,38	70,26	70,19	70,25	70,57	72,19	61,56	63,87	da G
71,07	63,76	70,36	69,18	69,22	70,37	70,44	59,65	60,51	de G
77,11	71,26	76,16	73,29	75,84	76,11	76,90	63,74	65,96	el
70,05	66,02	71,17	69,91	70,19	70,78	71,21	60,87	62,11	en G
71,47	71,98	73,83	72,29	73,22	73,89	74,12	71,40	70,76	es R
66,98	63,76	68,41	65,75	67,79	67,94	69,08	58,52	58,77	et
63,26	55,76	60,11	60,22	59,39	60,00	60,31	33,50	31,68	eu
72,85	66,02	69,50	69,63	70,00	69,67	68,02	31,25	34,27	fa
60,97	56,29	59,65	57,37	59,28	59,99	59,92	50,69	48,72	fi
75,74	74,25	76,16	74,79	75,82	75,80	76,50	72,68	71,97	fr R
66,55	60,41	66,86	64,68	65,96	66,02	66,30	43,51	42,75	ga
70,21	63,03	67,97	66,09	67,45	67,19	68,33	51,36	52,98	he
78,91	73,86	74,86	75,77	74,45	76,66	76,52	54,23	57,33	hi
71,03	67,49	71,37	70,00	70,40	71,42	71,28	62,88	65,71	hr S
67,08	62,55	66,84	67,19	67,51	67,79	68,14	48,24	49,62	hu
68,64	58,38	63,98	62,61	64,57	64,31	64,57	45,57	43,03	id
81,44	76,45	80,56	76,97	79,83	80,45	81,73	75,82	77,04	it R
78,26	68,22	75,81	74,85	75,56	76,59	76,89	7,64	28,53	ja
47,68	37,17	38,61	38,07	39,66	41,11	39,36	17,99	23,18	ko
59,89	54,11	59,98	58,23	60,17	59,78	60,86	44,41	47,38	lv
62,56	57,21	59,28	58,13	58,59	59,11	60,00	51,11	48,98	nl G
74,59	73,19	76,95	73,51	75,93	77,67	76,75	53,09	55,49	no G
81,24	74,24	80,52	74,78	79,02	80,44	80,06	69,15	70,72	pl S
72,00	65,74	69,83	68,74	69,86	69,88	69,96	62,85	65,86	pt R
70,99	67,72	71,47	70,38	70,61	71,26	71,19	55,84	61,01	ro R
74,06	61,35	74,72	68,65	74,45	74,92	75,06	55,09	55,50	ru S
67,10	64,12	66,44	64,75	66,36	66,63	66,69	60,52	63,26	sl S
72,05	69,97	72,55	70,71	71,88	73,21	73,25	64,80	67,05	sv G
46,78	41,01	43,16	43,26	41,62	45,73	43,57	29,46	30,33	tr
71,60	69,40	75,30	69,77	72,81	74,05	74,56	67,08	64,92	uk S
74,35	69,15	70,93	70,71	70,76	72,41	71,47	58,93	59,41	ur
54,40	42,42	53,75	51,94	51,72	52,81	52,42	25,86	41,07	vi
59,83	45,46	58,48	53,42	54,87	56,95	59,19	22,24	24,48	zh
69,32	63,64	68,25	66,41	67,64	68,39	68,54	51,86	53,80	MACRO
-	33,32	34,10	30,37	28,49	-	-	-	-	bxr
-	40,34	37,20	41,41	44,04	-	-	-	-	kmr
-	47,34	45,60	47,63	42,38	-	-	-	-	sme

TABLE 3 – LAS pour chaque langue, et MACRO LAS, pour les 9 configurations. Les langues suivies d’un S appartiennent à la famille des langues Slave, G appartiennent à la famille des langues Germaniques et R appartiennent à la famille des langues Romanes.

langues, comme l’équilibre entre les marqueurs morphologiques et syntaxiques (p.ex. les langues morphologiquement riches sont probablement favorisées dans notre configuration, l’analyse morphologique étant donnée en entrée de l’analyseur). D’autres sont spécifiques au genre textuel. Bien que la delexicalisation permette de neutraliser certains biais de genre, le genre peut aussi influencer la syntaxe, notamment la longueur des phrases (or, les phrases longues sont généralement plus dures à analyser), ou encore la proportion de constructions difficiles, comme les prépositions ou les coordinations ambiguës. Enfin, l’hétérogénéité de la qualité des annotations selon les langues peut également expliquer la variabilité du LAS.

**L vs  $\Sigma$**  : Une chute attendue des performances est observée lorsque l’on passe de  $L$  à  $\Sigma$ . Le MACRO LAS chute de 5,68 points. L’hypothèse principale pour expliquer cette chute est le bruit qui est introduit par le mélange des langues. Concaténer toutes les phrases de plusieurs langues sans préciser l’identité de la langue introduit du bruit dans l’ana-

$X$	$Y$	$\bar{X} - \bar{Y}$	$\sigma$	min	max
$L$	$\Sigma$	5,68	3,32	-0,51 es	14,37 zh
$\Sigma W_{80}$	$\Sigma$	4,00	2,58	0,59 hi	13,10 ru
$\Sigma ID$	$\Sigma$	4,60	2,91	1,00 hi	13,37 ru
$\Sigma W_d$	$\Sigma$	4,75	2,57	1,55 fr	13,57 ru
$\Sigma W_{80}$	$\Sigma W_N$	1,24	1,45	-1,64 tr	5,80 ru
$\Sigma ID$	$\Sigma W_{80}$	0,61	0,91	-1,05 ko	3,61 zh
$\Sigma W_d$	$\Sigma ID$	0,14	0,89	-1,53 zh	2,57 tr
$\Sigma W_{df}$	$\Sigma W_d$	0,15	0,89	-2,16 tr	2,24 zh
$L$	$\bar{L}$	17,47	13,57	0,07 es	70,62 ja
$\bar{L} W_{80}$	$\bar{L}$	1,95	4,60	-3,32 bg	20,89 ja

TABLE 4 – Différence entre les configurations  $X$  et  $Y$  : moyenne ( $\bar{X} - \bar{Y}$ ), écart type ( $\sigma$ ), minimum et maximum avec la langue correspondante.

lyseur. Par exemple, la configuration phrastique associée à une dépendance sujet dans une langue SV ou VS sera très différente, et l’analyseur n’est pas capable de faire la distinction entre ces langues et voit donc des contradic-

tions. La chute du LAS est cependant très variable selon les langues, allant même jusqu'à une augmentation plutôt qu'une chute dans le cas de l'espagnol (+0,51 points). Nous n'avons pas actuellement d'explication pour ce résultat, tout au plus une intuition qui serait que  $\Sigma$  apprendrait implicitement une langue moyenne (bruitée) qui serait plus proche de l'espagnol que du chinois par exemple (dont le LAS chute de 14,37 points), les langues composant  $\Sigma$  étant plus proches de l'espagnol que du chinois en moyenne.

$\Sigma$  vs  $\Sigma W_{80}$ ,  $\Sigma W_d$ , ID :<sup>15</sup> Nous obtenons ici le premier résultat important de ce travail : lors de l'ajout d'une représentation de la langue à l'analyseur, la MACRO LAS augmente de 4 à 4,75 points comparé à  $\Sigma$ . Le LAS augmente pour toutes les langues, pour les trois expériences. Ajouter l'information ID à  $\Sigma$  permet une augmentation des résultats MACRO de 4,60 points. Cette augmentation était attendue puisque dans ce test, les configurations phrastiques sont associées à l'ID de la langue, ce qui aide à diminuer le bruit dans les données. Deux interprétations sont possibles pour  $\Sigma W_N$  : celle optimiste serait que les vecteurs de représentations de la langue aident à diminuer le bruit introduit par le mélange des langues dans  $\Sigma$  en "expliquant" certaines informations contradictoires dans les données grâce à l'utilisation des traits linguistiques encodés dans le WALS pour  $W_{80}$ , et grâce aux informations de distance contenues dans  $W_d$ . L'interprétation pessimiste consiste à dire que ces vecteurs sont un simple encodage arbitraire des langues. Dans ce cas, le MLP de l'analyseur apprendrait à associer les configurations phrastiques à certaines langues spécifiquement, et apprendrait alors différents modèles pour différentes langues, ce qui reviendrait à l'expérience ID. Plus d'expériences sont prévues pour comprendre comment le MLP utilise les vecteurs  $W$ .

$\Sigma W_N$  vs  $\Sigma W_{80}$  : Les vecteurs  $W_N$  et  $W_{80}$  n'ont pas le même impact lorsqu'on les ajoute à  $\Sigma$ . L'ajout de  $W_{80}$  augmente de 4 points la mesure MACRO tandis que l'ajout de  $W_N$  augmente le score MACRO par rapport au MACRO obtenu avec  $\Sigma$  de 2,77 points seulement. L'analyseur est donc capable de tirer profit d'une description des langues plus riche lors de son apprentissage. Ce résultat pourrait indiquer que les résultats décevants sur l'analyse syntaxique rapportés par [1], qui utilisaient le vecteur  $W_N$ , pourraient venir des traits extraits du WALS qui n'étaient pas suffisamment riches pour expliquer à l'analyseur les différences syntaxiques importantes entre les langues.

$\Sigma$  ID vs  $\Sigma W_{80}$  : Malgré de meilleurs résultats obtenus en utilisant  $W_{80}$  plutôt que  $W_N$ , ce nouveau vecteur n'est toujours pas capable de surpasser les performances de l'utilisation d'un simple identifiant de la langue. Mais l'augmentation de la mesure MACRO entre  $W_N$  et  $W_{80}$  peut laisser supposer que le choix des traits du WALS influence fortement les résultats. Plus d'expériences sont nécessaires pour confirmer cette hypothèse. On pourra également se poser la question de l'influence de la méthode de remplacement des valeurs inconnues.

15.  $W_N$  et  $W_{df}$  sont analysés plus tard.

$\Sigma$  ID vs  $\Sigma W_d$  : Le résultat de ces expériences représentent le deuxième résultat majeur de ce travail : le vecteur  $W_d$  permet de battre les résultats de  $\Sigma$  ID. Même si l'hypothèse qu'une partie du vecteur  $W_d$  sert à représenter la langue s'avère juste, une partie du modèle a été capable d'apprendre des informations supplémentaires, partagées entre les différentes langues. Ce résultat est tout de même à nuancer, puisque le LAS de  $\Sigma W_d$  n'est supérieur que de 0,14 points à celui de  $\Sigma$  ID.

$\Sigma W_d$  vs  $\Sigma W_{df}$  : Les résultats de  $\Sigma W_d$  et de  $\Sigma W_{df}$  sont assez proches (+0,15 points pour la mesure MACRO de  $W_{df}$ ). L'information sur la fréquence des dépendances permet donc l'apprentissage de connaissances supplémentaires pour l'analyseur. Cependant, le temps d'apprentissage de  $\Sigma W_{df}$  étant assez conséquent, nous avons fait le choix de nous concentrer sur les résultats de  $\Sigma W_d$ . Le vecteur  $W_{df}$  étant deux fois plus gros que  $W_d$ , on peut également supposer que la taille de la couche cachée du MLP de l'analyseur n'est pas suffisante pour utiliser toutes les informations du vecteur  $W_{df}$ .

$L$  vs  $\bar{L}$  :  $\bar{L}$  correspond à des conditions extrêmes mais également plus réalistes. La situation simulée est celle où aucune donnée d'entraînement n'est disponible pour la langue  $l$ . La chute des performances comparée à  $L$  était prévisible, mais n'en reste pas moins dramatique : la MACRO LAS chute de 17,47 points. Cette chute varie énormément selon les langues (cela se ressent sur l'écart type qui atteint 13,57 points). Certaines langues ne sont presque pas affectées, comme l'espagnol qui ne perd que 0,07 point. En revanche, le japonais perd 70,62 points de LAS, la plus grosse chute toutes expériences et langues confondues. Les langues les plus isolées sont celles qui souffrent le plus du passage à  $\bar{L}$ . Les familles de langues sont moins affectées, avec une chute de "seulement" 6,65 points en moyenne pour les langues romanes par exemple.

$\bar{L}$  vs  $\bar{L} W_{80}$  : Cette comparaison constitue notre troisième et dernier résultat majeur. Cette expérience consiste à observer si l'ajout d'un vecteur issu du WALS permet de limiter la chute des score LAS de l'expérience  $\bar{L}$ . Comme on peut le voir dans les Tables 3 et 4, l'ajout de  $W_{80}$  permet d'augmenter la MACRO LAS de 1,95 points, ce qui pourrait laisser penser que  $W_{80}$  permet bien d'aider le modèle à gérer les langues inconnues. Cependant, on remarque que l'écart type est assez élevé (4,60 points). En effet, lorsque l'on regarde le détail langue par langue, 11 de nos 37 langues ont de moins bons résultats avec le modèle  $\bar{L} W_{80}$  qu'avec  $\bar{L}$ . Le bulgare perd même 3,32 points de LAS. Les langues ayant des résultats extrêmement bas dans  $\bar{L}$  ont réussi à tirer profit du vecteur  $W_{80}$ , comme le japonais qui, après son importante chute au passage à  $\bar{L}$ , remonte de 20,89 points avec  $\bar{L} W_{80}$ . Le vecteur  $W_{80}$  reste une description extrêmement partielle de la langue, et il est probable qu'il ne soit pas suffisant pour compenser l'absence de corpus d'entraînement. La question précédemment soulevée se pose à nouveau : le vecteur sert-il simplement d'identifiant de la langue ? Si l'analyseur uti-

lise, ne serait-ce qu'en partie,  $W_{80}$  pour détecter la langue de la phrase analysée, l'absence de cette langue dans le corpus d'entraînement rend la tâche impossible.

## 8 Conclusions et travaux futurs

Dans ce travail, nous avons analysé des représentations possibles d'une langue dans le cadre d'une analyse en dépendances multilingue. Les meilleurs résultats sont obtenus avec les représentations apprises sur les données  $(W_d, W_{df})$ , bien qu'ils dépassaient de peu la représentation par un simple identifiant (ID). Mais ces représentations avaient l'inconvénient de ne pas être utilisables dans le cadre d'une langue pour laquelle il n'existe pas de données d'apprentissage, contrairement aux représentations utilisant le WALS. L'utilisation de ce vecteur ne permettant cependant pas une amélioration nette des résultats pour toutes les langues lors d'un d'apprentissage de type *zero-shot*.

Pour les travaux futurs, nous prévoyons de faire une analyse des poids du réseau afin de tester l'hypothèse selon laquelle  $W$  est utilisé comme un simple identifiant de la langue. Nous aimerions également tester d'autres méthodes de remplacement des valeurs non-enseignées du WALS, ainsi que la création d'un  $W_{de}$ , qui utiliserait l'écart type plutôt que la fréquence. Une autre étape sera l'introduction de plongements de mots multilingues pour donner des informations lexicales à l'analyseur.

## Références

- [1] W. Ammar, G. Mulcaire, M. Ballesteros, C. Dyer, and N. A. Smith. Many languages, one parser. *TACL*, 2016.
- [2] W. Ammar, G. Mulcaire, Y. Tsvetkov, G. Lample, C. Dyer, and N. A. Smith. Massively multilingual word embeddings. *CoRR*, 2016.
- [3] D. Chen and C. Manning. A fast and accurate dependency parser using neural networks. In *Proceedings of the 2014 Conference on EMNLP*, 2014. ACL.
- [4] M. S. Dryer and M. Haspelmath. Wals online. 2013.
- [5] Y. Goldberg and J. Nivre. A dynamic oracle for arc-eager dependency parsing. In *Proceedings of COLING 2012*, 2012. The COLING 2012 Organizing Committee.
- [6] J. Guo, W. Che, D. Yarowsky, H. Wang, and T. Liu. Cross-lingual dependency parsing based on distributed representations. In *Proceedings of the 53rd Annual Meeting of ACL and the 7th International Joint Conference on NLP*, 2015. ACL.
- [7] T. Lynn, J. Foster, M. Dras, and L. Tounsi. Cross-lingual transfer parsing for low-resourced languages : An irish case study. In *Proceedings of the First Celtic Language Technology Workshop*, 2014. ACL and Dublin City University.
- [8] R. McDonald, S. Petrov, and K. Hall. Multi-source transfer of delexicalized dependency parsers. In *Proceedings of the 2011 Conference EMNLP*, 2011. ACL.
- [9] T. Naseem, R. Barzilay, and A. Globerson. Selective sharing for multilingual dependency parsing. In *Proceedings of the 50th Annual Meeting of ACL (Volume 1 : Long Papers)*, 2012. ACL.
- [10] G. Neubig, C. Dyer, Y. Goldberg, A. Matthews, W. Ammar, A. Anastasopoulos, M. Ballesteros, D. Chiang, D. Clothiaux, T. Cohn, K. Duh, M. Faruqui, C. Gan, D. Garrette, Y. Ji, L. Kong, A. Kunz, G. Kumar, C. Malaviya, P. Michel, Y. Oda, M. Richardson, N. Saphra, S. Swayamdipta, and P. Yin. DyNet : The Dynamic Neural Network Toolkit. *CoRR*, 2017.
- [11] J. Nivre. Algorithms for deterministic incremental dependency parsing. *Computational Linguistics*, 2008.
- [12] J. Nivre, M.-C. de Marneffe, F. Ginter, Y. Goldberg, J. Hajič, C. D. Manning, R. McDonald, S. Petrov, S. Pyysalo, N. Silveira, R. Tsarfaty, and D. Zeman. Universal dependencies v1 : A multilingual treebank collection. In *Proceedings of the Tenth International Conference LREC*, 2016. ELRA.
- [13] J. Nivre and J. Nilsson. Pseudo-projective dependency parsing. In *Proceedings of the 43rd Annual Meeting of ACL*, 2005. ACL.
- [14] E. M. Ponti, R. Reichart, A. Korhonen, and I. Vulić. Isomorphic transfer of syntactic structures in cross-lingual NLP. In *Proceedings of the 56th Annual Meeting of ACL (Volume 1 : Long Papers)*, 2018. ACL.
- [15] O. Täckström, R. McDonald, and J. Nivre. Target language adaptation of discriminative transfer parsers. In *Proceedings of NAACL*, 2013. ACL.
- [16] D. Vilares, C. Gómez-Rodríguez, and M. A. Alonso. One model, two languages : training bilingual parsers with harmonized treebanks. In *Proceedings of the 54th Annual Meeting of ACL (Volume 2 : Short Papers)*, 2016. ACL.
- [17] D. Zeman, J. Hajič, M. Popel, M. Potthast, M. Straka, F. Ginter, J. Nivre, and S. Petrov. CoNLL 2018 shared task : Multilingual parsing from raw text to universal dependencies. In *Proceedings of the CoNLL 2018 Shared Task*, 2018. ACL.
- [18] D. Zeman, M. Popel, M. Straka, J. Hajic, J. Nivre, F. Ginter, J. Luotolahti, S. Pyysalo, S. Petrov, and M. Potthast. CoNLL 2017 shared task : multilingual parsing from raw text to universal dependencies. *Proceedings of the CoNLL 2017 Shared Task*, 2017.
- [19] D. Zeman and P. Resnik. Cross-Language Parser Adaptation between Related Languages. 2008. ACL.
- [20] Y. Zhang and R. Barzilay. Hierarchical low-rank tensors for multilingual transfer parsing. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, 2015. ACL.



## Positionnement sur le Web Sémantique des Objets

Nicolas Seydoux<sup>1,2,3</sup>

Maxime Lefrançois<sup>4</sup>

Lionel Médini<sup>5</sup>

<sup>1</sup> LAAS-CNRS

7 avenue du colonel Roche, F-31400 Toulouse, France

<sup>2</sup> IRIT Toulouse, France

<sup>3</sup> UNIV. DE TOULOUSE,

INSA, LAAS, F-31400 Toulouse, France

<sup>4</sup> Mines Saint-Étienne, Univ Lyon, Univ Jean Monnet, IOGS, CNRS, UMR 5516, LHC,  
Institut Henri Fayol, F - 42023 Saint-Étienne France

<sup>5</sup> Univ. Lyon, Université Lyon 1 LIRIS, CNRS UMR5205,  
F-69622, Villeurbanne, France

### Résumé

*L'Internet des Objets (IoT en anglais) est un domaine riche en enjeux scientifiques, technologiques et sociétaux. Afin de faire face aux problématiques d'interopérabilité faisant obstacle au développement de l'IoT, les principes et les technologies du Web Sémantique sont intégrées aux réseaux d'objets connectés. Cette intégration est désignée sous le nom de Web Sémantique des Objets (SWoT en anglais). L'émergence de ce domaine repose sur l'adaptation des technologies du Web Sémantique aux contraintes de l'IoT. Ce papier de positionnement vise à décrire les vocabulaires principaux utilisés dans le SWoT, ainsi que l'architecture dans lequel la pile technologique du Web Sémantique se déploie, afin d'ensuite identifier les problématiques émergentes dans le domaine du SWoT. Ces problématiques sont des domaines d'intérêts potentiels pour les futurs travaux de la communauté Ingénierie des Connaissances (IC).*

### 1 Introduction

Les objets connectés sont désormais une réalité du quotidien, et depuis le premier emploi du terme en 1999, les enjeux scientifiques, techniques et sociétaux autour de la notion d'Internet of Things (IoT) n'ont fait que croître (Ashton, 2009). En particulier, le déploiement de réseaux dits IoT a des enjeux dans des domaines divers, tels que la ville intelligente, la télé-santé, l'agriculture numérique, ou l'industrie du futur. La grande hétérogénéité de ces domaines d'application a mené à l'émergence de "silos", avec l'intégration verticale de l'ensemble de la chaîne de traitement, depuis l'objet collectant la donnée jusqu'au programme l'intégrant dans son processus de collecte et de décision, sans oublier l'interface de l'application finalement consultée par l'utilisateur. Cette fracturation verticale est une source de problèmes d'interopérabilité et limite donc l'émergence d'applications complexes et multi-domaines, ainsi que l'intégration horizontale d'objets ou de services. Depuis sa création il y a 30 ans, le Web vise à masquer l'hétérogénéité des plateformes matérielles et logicielles, et

à permettre l'utilisation d'applications aussi spécifiques et performantes que possibles, en standardisant les langages et protocoles destinés à chaque type d'application. Dans cet esprit, le World Wide Web Consortium (W3C) s'est intéressé aux silos créés par l'IoT et est actuellement en train de finaliser plusieurs standards concernant le Web des Objets (WoT)<sup>1</sup>. Depuis leurs origines (Berners-Lee et al., 2001), les principes et technologies du Semantic Web (SW) ont été dévolus à la communication entre agents complexes, et ont promu l'interopérabilité. L'utilisation de modèles riches et expressifs semble particulièrement indiquée dans le cas des objets connectés, car on se situe dans le cas d'une communication destinée principalement à être interprétée par une machine, et ce n'est qu'une fois transformée qu'elle sera éventuellement transmise à un utilisateur humain. La convergence entre les domaines de l'IoT, du WoT et celui du SW ont mené à l'apparition d'un nouveau domaine de recherche, le Semantic Web of Things (SWoT). D'abord représenté par les réseaux de capteurs sémantisés (Sheth et al., 2008), le SWoT s'est ensuite diversifié pour aussi intégrer des actionneurs (Wang et al., 2015). Cependant, l'intégration des technologies du SW dans les réseaux d'objets connectés n'est pas triviale. En effet, ce sont des technologies qui sont en général demandeuses en ressources, autant en termes de puissance de calcul pour le raisonnement qu'en bande passante pour le transfert de données ou qu'en mémoire pour le stockage. Cette consommation importante de ressources est en opposition directe à la nature contrainte des réseaux IoT, dans lesquels sont déployés de nombreux objets aux ressources limitées. En effet, afin de permettre la dissémination d'un grand nombre d'objets dans l'environnement, il est nécessaire de miniaturiser ces objets, ainsi que de potentiellement les rendre autonomes en énergie. La combinaison de ces deux contraintes amène à une limitation nécessaire de la puissance de calcul, de l'espace mémoire, et des capacités de communication de l'objet. De plus, d'après des estimations récentes, le nombre

1. <https://www.w3.org/WoT/>

d'objets connectés va continuer à croître rapidement, pour passer d'environ 23 milliards en 2018 à une estimation de 75 milliards en 2025. Cette augmentation importante du nombre d'objets s'accompagne nécessairement de l'augmentation du volume de données qu'ils collectent, amenant un problème de passage à l'échelle pour les technologies du SW.

Partant de ce constat, nous questionnons dans ce papier une possible convergence entre les domaines du SWoT et des objets contraints. En particulier, nous nous appuyons sur des travaux existants dans la communauté IC dans ces deux domaines, en termes d'intérêt et de faisabilité, pour mettre en lumière certaines questions qu'une telle convergence fait émerger. Après une présentation des ontologies principales utilisées pour produire des messages compréhensibles par les machines dans le SWoT, une architecture de référence permettant de situer les contributions du Web Sémantique dans l'IoT est détaillée. Les défis représentés par le déploiement du SWoT, ainsi que des contributions y faisant face, sont ensuite décrits, afin d'identifier de futurs domaines d'intérêt pour la communauté.

## 2 Interopérabilité entre objets

L'un des moteurs poussant au développement du SWoT est le besoin d'interopérabilité de l'IoT, et en particulier d'interopérabilité sémantique (Murdock et al, 2016). En effet, les données et les systèmes qui les collectent sont extrêmement divers, ce qui rend complexe l'intégration de multiples sources de donnée à grande échelle. Afin de décrire le domaine de l'IoT, de nombreuses ontologies ont été proposées, c'est pourquoi nous avons identifié celles ayant un intérêt particulier. Les ontologies que nous avons listées ici sont celles se rapportant directement aux réseaux d'objets connectés, permettant de décrire les dispositifs, la manière dont ils interagissent avec le monde réel, ou la manière dont leurs services peuvent être découverts ou sollicités. Les domaines d'application de l'IoT sont nombreux, et décrire les ontologies qui s'y rapportent sort du périmètre de ce papier.

### 2.1 Les ontologies standardisées

Dans les domaines technologiques, le premier recours au manque d'interopérabilité est la standardisation. Avec l'émergence du SWoT, les standards incluent des ontologies, de façon à promouvoir l'utilisation de quelques ontologies de référence, plutôt que d'observer un morcellement et la création de nouvelles ontologies avec chaque nouveau projet. Les organismes de standardisation qui contribuent aux ontologies du SWoT sont notamment : (1) Le W3C, organisme international de standardisation du Web et du Web sémantique. (2) L'European Telecommunication Standards Institute (ETSI), organisme européen de standardisation pour les télécommunications. (3) oneM2M, consortium international rassemblant des organismes de standardisation (dont l'ETSI), des organismes de recherche et des industriels autour d'un standard pour l'IoT. Le groupe d'incubation Semantic Sensor Network du W3C a publié en 2011 un rapport analysant les différents modèles conceptuels existants pour décrire les capteurs et leurs observations, ainsi qu'une proposition d'ontologie SSNX (Lefort et al., 2011). L'ontologie SSNX a été massivement réutilisée par

d'autres ontologies et jeux de données. Le groupe de travail Spatial Data on the Web Working Group commun aux organismes de standardisation Open Geospatial Consortium (OGC) et W3C ont récemment publié une mise à jour de cette ontologie, nommée **SOSA/SSN** (Haller et al., 2017; Janowicz et al., 2018; Haller et al., 2018). Celle-ci spécifie la sémantique des capteurs et actionneurs, entre autres. Elle permet de décrire notamment les capteurs, les propriétés des choses d'intérêt qu'ils observent, les observations qu'ils font et le résultat de ces observations. De manière analogue, elle permet de décrire les actionneurs, les propriétés des choses d'intérêt sur lesquelles ils peuvent agir, et les actionnements qu'ils font et le résultat de ces actionnements. L'ontologie SOSA/SSN est modulaire. Par exemple : (1) SOSA (Sensor, Observation, Sampler, and Actuator) est un module qui comprend seulement quelques termes et peu d'axiomatisation ; (2) SSN (Semantic Sensor Network) importe SOSA, introduit d'autres termes, et enrichit l'axiomatisation de SOSA ; (3) SSN-System (SSN System Capability module) importe SSN (et SOSA par transitivité), et définit des classes supplémentaires pour modéliser les capacités des systèmes, leurs domaine d'opération ou leurs conditions nominales de déploiement.

Le W3C comprend un autre groupe de travail qui contribue directement au SWoT : le W3C Web of Things working group<sup>2</sup>. Entre autre, ce groupe a pour tâche la conception d'une ontologie, la **Thing Description ontology**, ainsi que son association à un ensemble de métadonnées décrivant les données exposées aux applications par les objets, leurs règles de sécurité, ou les informations de connexion à leurs points d'accès. La spécification de l'ontologie associée au modèle Thing Description est disponible dans (Kaebisch et al., 2019), encore en cours de développement. Une Thing est définie dans (Kovatsch et al., 2019) comme l'abstraction d'une entité (virtuelle ou physique) devant être manipulée par une application IoT, *e.g.* un objet, un service, ou une entité logique telle qu'une pièce ou un bâtiment. L'ontologie Smart Appliances Reference (**SAREF**) (Daniele et al., 2015), dont le développement est soutenu par l'ETSI, est centrée sur le concept de Device en tant qu'objet physique remplissant une ou plusieurs fonctions. Initialement conçue pour décrire le domaine de la maison intelligente, SAREF a été étendue par des modules spécialisés pour l'énergie, les bâtiments, l'environnement ou l'usine du futur. (Poveda Villalón et al., 2018) a analysé l'alignement possible entre SAREF et SSN, notamment pour l'étendre à la description d'un scénario d'agriculture numérique.

La **oneM2M Base Ontology**<sup>3</sup> est documentée dans l'une des spécifications techniques<sup>4</sup> du standard international oneM2M par le consortium éponyme. La oneM2M Base Ontology est une ontologie de type "cœur de domaine" : elle ne décrit qu'un ensemble réduit de concepts de haut niveau qui vise à être étendu par les implémentations du standard. Ces concepts, par exemple Device, Service ou Variable, font référence à

2. W3C WoT Charter - <https://www.w3.org/2016/12/wot-wg-2016.html>

3. [https://git.onem2m.org/MAS/BaseOntology/raw/master/base\\_ontology.owl](https://git.onem2m.org/MAS/BaseOntology/raw/master/base_ontology.owl)

4. TS-0012 : <http://www.onem2m.org/technical/published-drafts>

des éléments du standard. Un alignement entre SAREF et la oneM2M base Ontology est aussi distribué par oneM2M<sup>5</sup>. La prochaine version de SAREF devrait également comporter un alignement avec SOSA/SSN.

Il subsiste des questions ouvertes sur la modélisation et l'utilisation de ces ontologies, voici deux points de départ. (1) La section (Haller et al., 2017, §7) discute de questions ouvertes de modélisation avec SOSA/SSN. Les différentes options présentées n'ont pas encore fait l'objet d'articles de recherche. (2) L'ontologie TD étant maintenant stabilisée, il reste à ré-étudier l'alignement et l'interopérabilité possible avec SOSA/SSN.

## 2.2 Autres ontologies importantes du SWoT

Ces ontologies ne sont pas proposées par des organismes de standardisation, mais sont elle-même construites sur des ontologies de référence, issues d'organismes comme le W3C, ou de standards *de facto* dont le statut est acquis par l'usage.

L'ontologie **SEAS**<sup>6</sup> (Lefrançois, 2017) est une ontologie modulaire dont tous les termes sont contenus dans le même espace de noms. SEAS étend SOSA/SSN, et propose un cœur de quatre patrons ontologiques décrivant les systèmes physiques et leurs connexions, les valeurs associées à leurs propriétés, et les processus par lesquels ces associations de valeurs sont faites. Ces patrons sont ensuite instanciés dans chaque module pour un domaine particulier. Dans le contexte d'un projet ETSI Specialist Task Force 556<sup>7</sup> ces patrons ontologiques sont partiellement incorporés dans SAREF.

L'ontologie **S3N - Smart Semantic Sensor Networks**<sup>8</sup> (Sagar et al., 2018) est une ontologie modulaire qui étend SSN pour décrire des capteurs intelligents, c'est à dire des plateformes hébergeant au moins un capteur, un dispositif communicant, et un micro-contrôleur.

**IoT-O**<sup>9</sup> est une ontologie modulaire pour l'IoT, construite selon un ensemble de bonnes pratiques visant à permettre sa réutilisation (Seydoux et al., 2015, 2016a). Elle est notamment basée sur SSN initialement, et dispose d'une version mise à jour basée sur SOSA. Les différents modules de IoT-O, e.g. Capteurs et Observation ou Actionneurs et Actions, couvrent chacun un domaine de l'IoT, et sont indépendants les uns des autres.

schema.org est une ontologie légère décrivant un large panel de concepts se rapportant aux ressources publiées en ligne (e.g. blogs, boutiques en ligne ou contenu multimédia). Ce cœur a été ensuite étendu avec des vocabulaires se rapportant à un domaine spécifique présent sur le Web, comme les références bibliographiques<sup>10</sup>, et en particulier l'IoT, avec **iot.schema.org**. Un point de départ pour suivre le développement de cette extension ou contribuer est le groupe collaboratif GitHub <https://github.com/iot-schema-collab/>.

5. <https://git.onem2m.org/MAS/BaseOntology>

6. <https://w3id.org/seas/>

7. <https://portal.etsi.org/STF/STFs/STFHomePages/STF556>

8. <https://w3id.org/s3n/>

9. <http://irit.fr/recherches/MELODI/ontologies/IoT-O>

10. <https://bib.schema.org/>

## 3 Une architecture de référence pour le SWoT

La maturité du domaine du Web Sémantique, et le développement rapide du domaine de l'IoT, amènent l'émergence de problématiques spécifiques au SWoT. Ces problématiques s'articulent autour d'une dichotomie entre les **ressources nécessaires** au déploiement des technologies du Web Sémantique et les **capacités limitées** des nœuds présents dans les réseaux IoT. Ces contraintes sont caractérisées dans la suite de la présente section par l'introduction d'un patron architectural de référence pour le SWoT, utilisé pour situer les problématiques émergentes. Ce patron est ensuite instancié par une architecture orientée avatar permettant d'envisager la construction d'applications WoT et SWoT.

### 3.1 Que sont le Cloud et le Fog computing ?

Le Cloud computing (Mell and Grance, 2011) est un paradigme dans lequel les traitements sont effectués à distance sur des machines dotées d'importantes ressources : puissance de calcul, mémoire, stockage, bande passante, voire systèmes et applications. L'**accessibilité via le Web** et l'**élasticité** des machines du Cloud en font des relais intéressants pour les applications du SWoT. Dans ce modèle, les données sont collectées par les objets connectés, puis concentrées sur les nœuds Cloud où elles sont stockées, traitées, et intégrées à des applications.

Cependant, les réseaux d'objets connectés contiennent par nature des objets contraints, et très disséminés dans leur environnement. Leurs capacités de communication limitées peuvent être un obstacle à la communication systématique avec des nœuds distants. De plus, la concentration des données dans un serveur Cloud peut introduire un délai important, dépendant du volume de données qui transitent et de la qualité du lien réseau. Pour faire face à cette fracture entre les infrastructures Cloud et les sources de données de l'IoT, un nouveau paradigme est apparu, le Fog Computing (Bonomi et al., 2012), basé sur l'utilisation des ressources de calcul et de stockage disponibles **en bordure de réseau**. Les nœuds Fog sont donc **massivement distribués, hétérogènes**, et disposent de **capacités de calcul limitées**, et se situent entre les objets connectés et les machines du Cloud. Cette définition englobe notamment les passerelles que l'on retrouve dans de nombreux déploiement d'objets connectés, et qui assurent la connexion entre les protocoles de l'IoT et Internet, permettant ainsi de transporter les données des objets jusqu'aux serveurs distants du Cloud.

Le Cloud comme le Fog computing visent à fournir une abstraction des ressources de calcul pour permettre une gestion plus souple des tâches. Cependant, il existe des différences fondamentales entre ces paradigmes. **(I)** Les nœuds du Fog sont situés à **proximité** des objets connectés, fournissant ainsi des capacités de calcul plus proche des sources de données dans le cas de l'IoT. **(II)** Le traitement distribué supporté par le Fog computing a des propriétés de résilience, par l'absence de point critique sur lequel repose l'intégralité du calcul. **(III)** Cloud et Fog offrent des approches différentes pour le passage à l'échelle. Le Cloud s'appuie sur deux axes : l'attribution de machines supplémentaires à la tâche en cours, ou l'attribution de ressources

supplémentaires aux machines qui y sont déjà dédiées. Dans le Fog computing, les ressources des nœuds n'étant pas flexibles comme dans le Cloud, seule l'augmentation du nombre de machines est possible. Cependant, cette augmentation se faisant localement par rapport aux sources de données, et s'appuyant sur des machines faisant potentiellement déjà partie du déploiement d'objet, elle s'intègre naturellement au développement massif des grands réseaux d'objets, amenant des ressources complémentaires à celles des serveurs Cloud (Dastjerdi and Buyya, 2016). **(IV)** La nature distribuée des nœuds constituant l'infrastructure Fog rend aussi plus fluide la répartition de la puissance de calcul face à la mobilité des besoins, ce qui est en particulier adapté au domaine de l'IoT (Dastjerdi and Buyya, 2016), où les objets sont disséminés dans un espace géographiquement étalé dans lequel les utilisateurs et certains objets se déplacent. Le Fog computing n'a pas vocation à remplacer le Cloud computing, mais bien à le compléter. La capacité de calcul limitée des nœuds Fog, leur mobilité, et la localité de leur déploiement ne sont pas adaptés à la prise en charge de cas d'utilisation du Cloud computing. Le rapprochement de ces deux paradigmes est un vecteur intéressant pour fournir des infrastructures soutenant le développement de l'IoT en général, et du SWoT en particulier.

### 3.2 Structurer une architecture de référence du SWoT

Tous les nœuds de l'IoT ne sont pas équivalents, et ce domaine se distingue par l'hétérogénéité des capacités des machines impliquées, depuis des objets très contraints jusqu'aux serveurs Cloud dotés d'une très grande puissance de calcul. Les principales caractéristiques distinguant les nœuds sont : la puissance de calcul, la mémoire, le stockage, les capacités de communications, et le type de source d'énergie.

Trois classes de nœuds se détachent par l'analyse de la littérature via ces critères :

1. Les **nœuds Cloud**, caractérisés par leur puissance de calcul importante, leur large capacité de communication (nombreux protocoles, connectés en permanence), ainsi que leur mémoire et leur capacité de stockage étendue. Cette classe de nœuds inclut les serveurs déployés dans le Cloud, et peut aussi être étendue aux nœuds dont la puissance est importante relativement aux autres machines du même déploiement, pouvant contextuellement s'appliquer à une station de travail classique. De tels nœuds sont présents dans les architectures mises en jeu dans des travaux tels que (Ben-Alaya et al., 2015; Szilagy et al., 2016; Su et al., 2018), avec des appellations variables.
2. Les **nœuds Fog**, caractérisés en premier lieu par leurs importantes capacités de communication, et leur rôle de pourvoyeur de contenu entre objets connectés et nœuds Cloud en assurant la continuité des protocoles. Ce rôle fait que ces nœuds sont souvent désignés par le terme de "passerelle" dans la littérature (Compton et al., 2009; Ben-Alaya et al., 2015).
3. Les **objets contraints**, caractérisés par leurs capacités de traitement et de communication limitée, leur autonomie en énergie, et leur très faible capacité de stockage. Ces nœuds sont par définition présents dans tous les réseaux IoT, dont ils

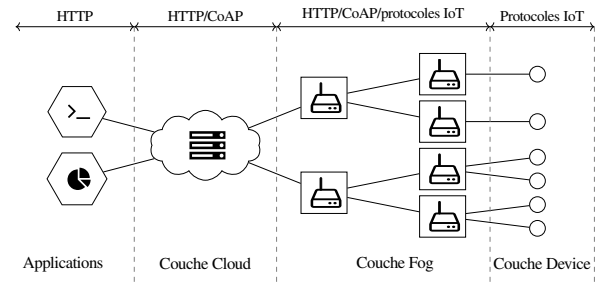


FIGURE 1 – Patron architectural de référence pour le SWoT

sont la source principale d'informations sur l'environnement.

### 3.3 Un patron architectural de référence pour le SWoT

Les classes de nœuds précédemment définies constituent un patron architectural récurrent dans la littérature, découpé en trois couches : **Cloud-Fog-Devices**, représenté sur la Fig. 1. Dans ce patron architectural, les services sont exposés aux utilisateurs par un nœud Cloud via des protocoles Web, lequel collecte des données produites par des objets via l'entremise de nœuds Fog. Ce patron architectural, et les différentes couches qu'il définit, seront utilisés comme référence dans la suite de ce papier.

### 3.4 Composition de services métier orientée-avatars

Depuis les automates programmables et les robots, des éléments de programmation ont été associés aux objets. Avec l'avènement des communications en réseau (sans fil), ils ont évolué vers des systèmes de contrôle distribués, des systèmes embarqués et, plus récemment, l'intelligence ambiante et la robotique distribuée (Galloway and Hancke, 2013). L'IoT est une conséquence directe de cette évolution et vise à exploiter les capacités de communication d'Internet, la puissance de calcul quasi illimitée des infrastructures de la couche Cloud et des interfaces utilisateur modernes pour fournir aux utilisateurs finaux des applications utiles. Le WoT s'appuie sur l'IoT et encourage l'utilisation des standards du Web.

Mais dès que des objets physiques entrent en jeu, la programmation de telles applications devient plus complexe et moins déterministe. Les données des capteurs souffrent d'imprécisions, les actionneurs ont besoin de boucles de rétro-action, les processus critiques et de synchronisation nécessitent une attention constante et les communications réseau peuvent perdre des données ou être interrompues. Les défenseurs des Systèmes Cyber-Physiques (CPS) affirment que de telles difficultés proviennent de phénomènes physiques et doivent être modélisées avec des modèles et des processus informatiques (Lee, 2008). Parmi les nombreuses architectures CPS qui ont été conçues, bon nombre d'entre elles sont à base de composants et incluent des entités logicielles qui modélisent ces préoccupations.

Dans la communauté WoT, nous avons proposé (Jamont et al., 2014; Mrissa et al., 2015) la notion d'**avatar** pour désigner les artefacts logiciels attachés à un objet et agréger le code né-

cessaire pour exécuter des applications WoT. Comme détaillé ci-dessous, les avatars reposent sur une architecture interne à base de composants, de sorte que toutes les préoccupations nécessaires du point de vue des CPS puissent être décrites. Plus récemment, le groupe de travail WoT du W3C a proposé la notion de « servient » actuellement définie pour normaliser les objets logiciels dans les applications WoT. Les servients sont très proches des avatars : ils fournissent un accès aux objets, peuvent être exécutés sur ceux-ci dans la couche Device, mais aussi dans les couches Fog et Cloud, et peuvent interagir avec d'autres serveurs. Les deux s'appuient également sur des technologies sémantiques pour échanger des données compréhensibles par une machine. Comme les standards du WoT doivent faire face à une variété de cas d'utilisation et de plateformes, l'architecture cliente ne spécifie que des blocs de construction (« environnement d'exécution », « modèle de ressource », etc.). Les avatars peuvent être considérés comme une spécialisation de servients, davantage centrés sur le déploiement et l'exécution d'applications WoT, et s'appuyant sur une architecture à base de composants pour tirer parti des avancées liées aux préoccupations et aux exigences spécifiques dans divers domaines.

### 3.5 Avatars

Certains composants de l'architecture de l'avatar (Figure 2) sont dédiés au contrôle des objets et d'autres implémentent le comportement autonome, adaptatif et collaboratif des avatars. La configuration physique est découplée de son architecture logique : un avatar peut adapter de manière dynamique la distribution de ses composants à travers le patron architectural de référence pour améliorer leur efficacité. Nous avons regroupé les composants d'avatar en 8 modules fonctionnels.

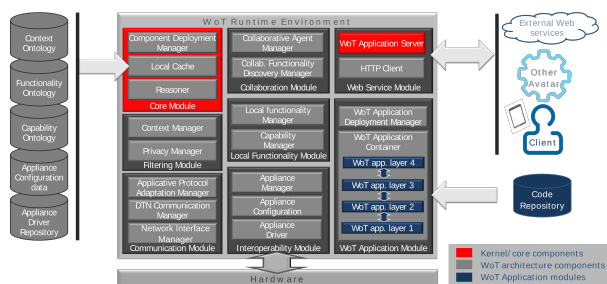


FIGURE 2 – Architecture logicielle de l'avatar

Le **module Core** comprend des composants utilisés à plusieurs étapes du cycle de vie d'un avatar. Le gestionnaire de déploiement de composants définit quels composants d'avatar seront instanciés par rapport aux fonctionnalités de l'objet, et où<sup>11</sup>. Chaque avatar intègre un moteur d'inférences, utilisé par d'autres composants pour traiter les informations sémantiques relatives aux capacités, aux fonctionnalités et au contexte. Il en

11. Les composants d'un avatar peuvent être situés : sur l'objet si celui-ci dispose de suffisamment de capacités de traitement ou pour des modules de code possédant des contraintes temporelles ; dans la couche Fog pour les processus impliquant une communication inter-avatars localisée ; ou dans la couche Cloud pour les processus gourmands en calculs.

va de même pour le cache local, qui stocke des informations sémantiques provenant de sources diverses (objet, référentiels, contexte externe) et reflète l'état actuel de l'avatar.

Le **module d'interopérabilité** fournit aux autres modules de l'avatar une interface uniforme pour interagir avec l'objet auquel il est attaché. Cette interface consiste en un ensemble de *capacités* représentant l'API de l'objet. Il charge les pilotes depuis un dépôt de la plateforme WoT (voir plus loin) et les utilise pour identifier les schémas de communication compris par l'objet. Finalement, il télécharge sur l'objet la configuration appropriée.

Le **module de filtrage** limite l'exposition des fonctionnalités et les échanges de données. Si, pour des raisons de confidentialité ou de sécurité, certaines fonctionnalités ne sont pas réalisées par l'avatar, elles seront filtrées par le gestionnaire de la confidentialité. Le gestionnaire de contexte a un rôle plus complexe, qui permet notamment de décider quelles fonctionnalités peuvent être exposées, à partir de quelles capacités de l'objet les réaliser, et avec quels autres avatars collaborer pour composer des fonctionnalités complexes (Terdjimi et al., 2016).

Le **module de communication** assure une communication fiable avec l'objet. Il sélectionne l'interface réseau appropriée (Ethernet, Wi-Fi, Zigbee...) et les protocoles (CoAP, HTTP...) en fonction des objectifs de la communication et des besoins en performances (débit / consommation d'énergie). Il prend également en charge les interruptions de connectivité (Mahéo et al., 2012).

Le **module de services Web** permet aux avatars de communiquer entre eux et avec le monde externe dans le respect des standards du Web. Ainsi, les avatars peuvent : interagir avec la plateforme WoT pour interroger des dépôts de code, répondre aux demandes des clients concernant les fonctionnalités qu'ils exposent en tant que ressources RESTful, échanger des données avec d'autres avatars pour réaliser des fonctionnalités collaboratives et interroger des services Web externes pour enrichir leurs propres données.

Le **module de fonctionnalités locales** gère les *fonctionnalités* de haut niveau réalisables à l'aide des capacités atomiques de l'objet et de leur composition. Il s'appuie sur des technologies sémantiques pour lier la couche physique (capacités) avec la couche applicative (fonctionnalités) de manière déclarative et faiblement couplée, en garantissant l'interopérabilité des applications pour les objets (Mrissa et al., 2014). Lorsque l'avatar est créé, le CapabilityManager interroge le module d'interopérabilité pour connaître les fonctionnalités de l'objet et l'ontologie des fonctionnalités de la plateforme pour obtenir leurs descriptions sémantiques. Il est interrogé par le LocalFunctionalityManager, qui charge également les descriptions de fonctionnalités et utilise le raisonneur pour déduire les fonctionnalités locales de l'avatar. Pour chaque fonctionnalité inférée, le LocalFunctionalityManager interroge le gestionnaire de contexte pour déterminer si elle doit être exposée aux clients. Les fonctionnalités exposées sont liées à un registre, afin de permettre leur découverte.

Le **module de Collaboration** gère les fonctionnalités nécessitant une collaboration entre plusieurs avatars. Le CollaborativeFunctionalityDiscoveryManager interroge le raisonneur

comme décrit ci-dessus pour identifier, à partir des fonctionnalités locales, celles auxquelles il pourrait participer aux niveaux supérieurs. Ensuite, il interroge le répertoire des fonctionnalités de la plateforme pour rechercher les fonctionnalités manquantes localement. Si de telles fonctionnalités sont disponibles auprès d'autres avatars, il charge le CollaborativeAgentManager, qui gère la négociation avec ces autres avatars, de mettre en oeuvre ces fonctionnalités collaboratives (Jamont, 2016).

Le **module d'Application WoT** fournit et contrôle des « conteneurs d'applications WoT » qui exécutent des modules de code implémentant les différents aspects d'une application (voir ci-dessous). De tels conteneurs peuvent être répliqués sur l'objet, sur la passerelle et sur l'infrastructure cloud grâce au gestionnaire de déploiement, afin que les modules soient exécutés à l'emplacement approprié.

### 3.6 Plateforme orientée-avatar

Les solutions IoT existantes peuvent être utilisées en tant que couche de support pour connecter les objets <sup>12</sup>. Les plateformes WoT sont déployées en tant que « serveurs d'applications WoT » au-dessus de ces solutions. Pour permettre le déploiement et l'exécution d'applications WoT sur des environnements informatiques ubiquitaires, les plateformes WoT doivent fournir un accès à des installations de stockage d'informations et de connaissances et à une puissance de calcul supplémentaire dans la couche Cloud. Dans la mesure où nous promouvons ici les plateformes WoT basées sur des avatars, elles devraient également prendre en charge la gestion, l'exécution et la (dé)serialisation des avatars. Le mécanisme de sérialisation permet un passage à l'échelle horizontale en répliquant les plateformes et en déplaçant les avatars entre elles à travers les couches Fog et Cloud. La figure 3 décrit l'infrastructure de telles plateformes, où on retrouve instancié le patron architectural décrit en Section 3.3 <sup>13</sup>.

Les principaux éléments de cette infrastructure sont le contrôleur d'infrastructure et l'environnement d'exécution WoT. Le premier interagit avec la plateforme IoT et est chargé de décider de créer <sup>14</sup>, de mettre à jour <sup>15</sup> ou de supprimer des avatars, au fur et à mesure que les objets sont connectés et déconnectés de la couche IoT. Le second est le conteneur qui isole les avatars et gère leur cycle de vie. Pour des raisons de performances, il est également connecté à la passerelle, de sorte que les avatars puissent directement interagir avec les objets via Fog computing, sans traverser la pile de la plateforme IoT à chaque requête. À l'intérieur du conteneur, les avatars peuvent également : accéder aux ressources Web externes via le proxy Web, partager des informations sur les fonctionnalités exposées par chaque avatar à l'aide du registre de fonctionnalités, interroger les différents dépôts pour accéder aux descriptions sémantiques dont ils ont besoin pour fonctionner, et récupérer

12. Nous avons développé une *Couche d'Interopérabilité Matériels-Applications* (CIMA) sur ce principe, à partir de la plateforme IoT OM2M (Alaya et al., 2014) : <https://github.com/ucbl/CIMA>

13. La Gateway est une passerelle, élément commun de la couche Fog.

14. Chaque avatar est construit de manière à pouvoir accéder à l'objet auquel il se rapporte à travers la passerelle. Les capacités de l'objet sont injectées et les composants de l'avatar instanciés.

15. En requêtant périodiquement la plateforme IoT.

le code des modules applicatifs dans le dépôt de code. Les programmes *Device Installer* et *Application Installer* sont chargés d'alimenter les différents référentiels et sont indépendants de la notion d'avatar. Les utilisateurs interagissent en toute sécurité avec la plateforme pour télécharger des pilotes d'objets, installer des applications WoT à partir d'un « marketplace » en ligne ou les exécuter à l'aide de leur navigateur Web, via le proxy qui bloque les demandes entrantes non identifiées.

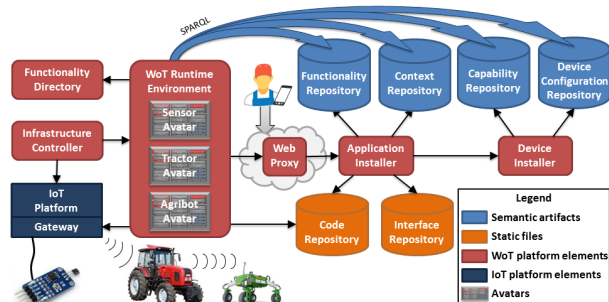


FIGURE 3 – Architecture d'une plateforme orientée-avatar pour le WoT.

### 3.7 Applications WoT

Afin de faciliter la conception des applications WoT et de conserver une indépendance vis-à-vis des caractéristiques des objets disponibles, une application WoT ne traite que la couche fonctionnalité. Elle décrit principalement une hiérarchie de fonctionnalités, dont les nœuds terminaux sont des fonctionnalités terminales (c-à-d. devant être implémentées par une capacité de l'objet <sup>16</sup>), et tous les autres nœuds sont des fonctionnalités composées (*i.e.* qui nécessitent des sous-fonctionnalités et les interrogent à l'aide de modules applicatifs). Certaines de ces fonctionnalités composées peuvent nécessiter d'utiliser les capacités de plusieurs objets et, par conséquent, une collaboration entre plusieurs avatars. La fonctionnalité de niveau supérieur correspond alors à l'application que l'utilisateur final souhaite utiliser. Une application WoT est packagée dans un fichier compressé, composé de : un fichier *Manifest*, décrivant son contenu ; la hiérarchie de fonctionnalités susmentionnées ; les modules de code correspondant aux algorithmes implémentant les fonctionnalités composées <sup>17</sup> ; le modèle de contexte de l'application, contenant une description sémantique du domaine d'application et un ensemble de règles d'adaptation (Terdjimi et al., 2016) ; un ensemble de fichiers statiques constituant l'interface de l'application et permettant aux utilisateurs finaux d'exécuter et de contrôler l'application via leur navigateur

16. Les capacités sont sémantiquement liées à ces fonctionnalités terminales lors de l'initialisation de l'avatar.

17. L'exigence d'interopérabilité impose à ces applications d'être décrites de manière générique, de manière à pouvoir être déployées et exécutées sur différentes configurations. Par conséquent, nous recommandons de les décrire à l'aide de langages déclaratifs, tels qu'une machine à états finis, et lors de l'installation, de les pré-transpiler dans des langages exécutables sur l'objet, sur la passerelle et sur la plateforme cloud sur lesquels ils peuvent être déployés.



Web en interrogeant les fonctionnalités des avatars à l'aide de standards du Web (ressources RESTful, WebSockets, etc.).

## 4 Problématiques émergentes

Représenter les données d'une manière qui soit à la fois interopérable et adaptée aux contraintes de l'IoT n'est pas le seul défi auquel fait face le SWoT : l'ensemble des technologies du Web Sémantique est concerné par ce besoin d'adaptabilité. La Section 3 donne l'architecture dans laquelle les composants du SWoT évoluent, ainsi que les fonctionnalités qui y sont regroupées par la notion d'avatar. Cette section montre les défis que représentent l'intégration du SWoT dans cette architecture. En particulier, différents aspects des objets de l'IoT amènent des problématiques dont la communauté du SWoT s'empare peu à peu. C'est notamment le cas de leur nature fortement contrainte, l'aspect éclaté des réseaux qu'ils forment, la forte contextualité des applications qui y sont liées, et les problèmes de passage à l'échelle inhérents à leur déploiement.

### 4.1 Comblent l'écart entre ontologies et objets

Les ontologies sont des vocabulaires riches, et celles que nous avons identifiées sont toutes représentées en RDF<sup>18</sup>. Basé sur des URI comme identifiants de ressource, le RDF est un langage dont certaines sérialisations peuvent s'avérer assez verbeuses, ce qui implique une consommation d'espace mémoire et de bande passante, qui sont des ressources limitées pour les objets connectés. De ce fait, pour pouvoir manipuler des données sémantisées dans les couches Fog et Cloud, il est nécessaire de considérer les contraintes de la couche Device. Cette contrainte peut être intégrée aux modules de communication ou de filtrage des avatars. Nous identifions deux approches principales : faire en sorte que les objets puissent directement manipuler du RDF, ou assurer une traduction entre RDF et des formats plus adaptés aux contraintes des objets.

**Amener le RDF aux objets....** Afin de rendre possible la manipulation du RDF par les objets, il est nécessaire de proposer des syntaxes RDF plus compactes que les syntaxes les plus répandues, comme le RDF/XML ou la notation N3. (Su et al., 2018) propose un panorama des sérialisations RDF dans l'optique de les manipuler dans un environnement contraint. Par exemple, dans (Charpenay et al., 2018d), les auteurs proposent une sérialisation binaire de JSON-LD afin de permettre l'échange plus efficace des données dans le contexte d'objets contraints. De plus, ce travail est étendu dans (Charpenay et al., 2018b), afin de s'appuyer sur ce format pour stocker et manipuler les bases de connaissance en RDF directement sur les objets connectés. De plus, ces travaux prennent aussi en compte les contraintes en terme de protocoles des communications entre les couches Fog et Device : les protocoles du Web Sémantique, comme HTTP et SPARQL, reposent sur TCP, qui demande l'établissement d'une connexion, ce qui peut être coûteux en bande passante. De la même façon, (Loseto et al., 2016) étend la spécification de Linked Data Platform (LDP)<sup>19</sup>, une recommandation du

W3C qui décrit un protocole d'exposition de données liées, potentiellement en RDF. Dans la recommandation originale, LDP est associé aux verbes HTTP, que les auteurs ont remplacé par leur équivalent CoAP<sup>20</sup>, un protocole construit sur UDP, spécialement conçu pour les échanges dans la couche Device.

**... ou assurer la traduction depuis et vers le RDF.** Même si des technologies émergent pour rendre possible la manipulation de RDF sur des objets contraints, un grand nombre d'objets et de services déjà en circulation n'ont pas cette capacité. Pour pouvoir intégrer cette masse d'objets au SWoT, il est donc nécessaire d'établir une transformation entre les formats de données spécifiques à l'IoT et le RDF. Transformer des données vers du RDF est une problématique dont la communauté du Web Sémantique s'est déjà largement emparé : on parle dans ce cas d'**enrichissement**. Cependant, les objets de l'IoT ne sont pas uniquement des producteurs de données, ils en sont aussi des consommateurs, par exemple dans le cas des actionneurs, ou pour la configuration. Il est donc non seulement nécessaire d'enrichir les données issues des objets, mais aussi d'effectuer la transformation inverse, à laquelle on pourrait faire référence comme un **appauvrissement**. En effet, la transformation du RDF vers un format moins expressif s'accompagne d'une perte de contexte, qui n'est plus explicite, mais qui correspond au contexte implicite dans lequel la donnée ainsi appauvrie sera consommée par l'objet destinataire de la communication. La problématique de la traduction bidirectionnelle entre RDF et formats contraints a été abordée dans divers travaux récents. Dans (Seydoux et al., 2016b), l'approche proposée se destine principalement aux formats arborescents aux schémas explicites, de type XML ou JSON. À partir d'associations entre les balises du langage et les ressources d'un graphe RDF, le système effectue la transformation de manière semi-automatique. Il est particulièrement adapté aux langages formalisés par un standard, dans lequel le schéma est fixé. Les travaux de (Charpenay et al., 2018c) visent eux aussi à assurer la traduction entre différents modèles standards en passant par l'enrichissement. Dans ce cas, la traduction n'est considérée que dans le sens de l'enrichissement. La méthode vise à offrir de l'interopérabilité entre les standards en construisant de manière semi-automatique des règles de traduction d'un modèle enrichi à l'autre. Dans (Lefrançois, 2018), une approche plus souple est proposée en s'ancrant sur les principes de l'architecture du Web : les agents sur le Web, potentiellement contraints, s'échangent des représentations de graphes RDF sous formes de flux d'octets typés par des mediatypes, comme l'illustre la Figure 4. Ce cadre conceptuel, nommé RDF Presentation, définit les termes RDF **lowering** (abaissement, étape 1 sur la Figure 4) et **lifting** (élévation, étape 3 sur la Figure 4) au lieu d'enrichissement et appauvrissement, étant donné que les représentations RDF sont supposées encoder toutes les caractéristiques du graphe RDF. À partir d'en-têtes HTTP, d'options CoAP, ou d'autres mécanismes, la source ou le destinataire de la communication peut indiquer les règles de transformations entre le format échangé et RDF. Un exemple de langage expressif d'enri-

18. <https://www.w3.org/RDF/>

19. <https://www.w3.org/TR/ldp/>

20. <http://coap.technology/>

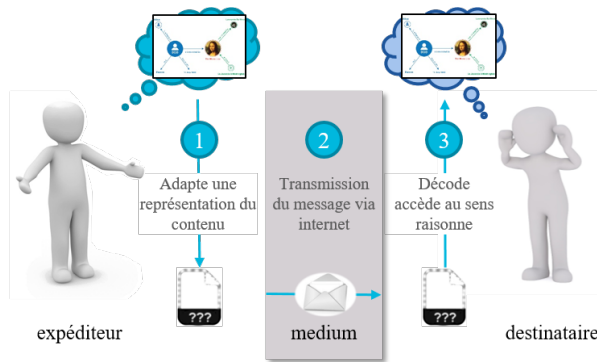


FIGURE 4 – RDF Presentation : utilisation de transformations pour échanger du RDF sur le web (Lefrançois, 2018)

chissement ou d'élévation est SPARQL-Generate<sup>21</sup> (Lefrançois et al., 2017), qui est une extension de SPARQL 1.1 qui permet de requêter à la fois du RDF et des documents dans des formats hétérogènes (XML, JSON, CBOR, CSV, HTML, GeoJSON, CBOR, ...) ou des flux de tels documents (WebSocket, MQTT).

## 4.2 Distribuer le raisonnement

Les technologies du Web Sémantique ont initialement été déployées sur des machines du Cloud pour leur permettre de traiter des données issues d'objets connectés, sans être limitées par leurs contraintes. Le développement du Fog computing offre de nouvelles ressources intégrées aux réseaux d'objets connectés, avec une couche Fog mobile et plus proches de objets. Bien que les nœuds Fog soient largement moins puissants que les serveurs du Cloud, ils peuvent cependant supporter dans une mesure raisonnable les technologies du Web Sémantique sans adaptation particulière de protocole ou de format de donnée. De ce fait, la couche Fog est un candidat privilégié dans le rôle de médiateur entre les formats et les protocoles les plus répandus du Web Sémantique, et ceux spécifiques à la couche Device de l'IoT. La traduction proposée dans la Section 4.1 est de ce fait souvent poussée vers les nœuds Fog, qui ont le double avantage de pouvoir communiquer avec les objets connectés, et de gérer un nombre limité d'objet. Il est donc possible pour les passerelles Fog d'utiliser leur connaissance du contexte et des objets auxquels elles sont connectées pour enrichir les données issues des objets, capturant dans la représentation de ces données leur contexte de collecte. Ainsi, les données peuvent être enrichies sémantiquement à partir de leur entrée dans la couche Fog.

Le découpage des avatars en modules de fonctionnalité, et la mobilité de ces modules, permet d'envisager leur distribution à travers l'architecture de référence. En particulier, la description des données à partir d'ontologies permet le déploiement de traitements tirant parti du formalisme ainsi introduit, en particulier du raisonnement. Un raisonneur peut, à partir d'une base de connaissance, identifier des inconsistances ou dériver de nouvelles connaissances. La production de nouveaux triplets peut se baser sur des règles d'inférence génériques capturées

21. <https://w3id.org/sparql-generate>

dans les langages de modélisation comme OWL, mais peut aussi être décrite par des règles d'inférence métier, décrites dans des langages spécifiques. En particulier, SHACL<sup>22</sup> est la recommandation la plus récente du W3C en matière de représentation de contrainte sur des graphes RDF, et une de ses extensions porte sur la représentation de règles<sup>23</sup>. En s'appuyant sur des requêtes SPARQL CONSTRUCT, SHACL permet d'écrire des règles d'inférence menant à la production de nouvelles connaissances. Une fois sérialisées, les règles de déduction peuvent être échangées, rendant possible la distribution du raisonnement, notamment à travers la couche Fog. Dans (Seydoux et al., 2018a,b), les applications consommant des données issues d'un réseau d'objets sont représentées sous la forme d'une collection de règles, lesquelles sont ensuite propagées de proche en proche entre les nœuds du réseau dans une approche nommée EDR. La propagation est complètement décentralisée, et elle est guidée par une définition de politique directement embarquée dans la règle, rendant l'approche complètement flexible. Un exemple de politique visant à amener les règles au plus près des objets est détaillé dans (Seydoux et al., 2018b). Les nœuds échangent entre eux des informations concernant les types de données qu'ils produisent, et les règles sont annotées pour identifier les types de données qu'elles consomment. Les nœuds transmettent la règle aux autres nœuds produisant l'ensemble des données qu'elle consomme, jusqu'à arriver à un point critique, qui est l'ancêtre commun le plus proche des différents producteurs de données nécessaires à la règle. Pour assurer un placement des règles adaptatif aux évolutions du réseau, les mises à jour de la topologie sont propagées dynamiquement par les nœuds, qui remettent en permanence en question la distribution des règles. EDR suppose une disposition hiérarchique des nœuds du réseau en un graphe acyclique, que l'on retrouve dans le patron présenté en Section 3.3. La distribution du traitement permet de réduire la quantité d'information à traiter par chaque nœud appliquant la règle, amenant ainsi à des temps de raisonnement plus court, réduisant ainsi le délai entre l'observation d'un ensemble de signes caractérisant un phénomène et sa prise en compte par une application de haut niveau. De plus, distribuer les règles permet de contextualiser leur application, en ne les considérant que dans les sous-ensemble du réseau dans lequel elles sont pertinentes.

## 4.3 Distribuer les données

Par nature, les données issues des réseaux d'objets sont fortement liées à un contexte spatio-temporel, exprimant des observations n'ayant de valeur de vérité que si elles sont considérées pour un lieu et un instant donné. Ces données sont produites par des sources très distribuées géographiquement, lesquelles sont potentiellement opérées par des entités indépendantes les unes des autres. De plus, quand bien même les données sont enrichies, soit directement par les objets par l'utilisation des formats spécifiques (*c.f.* Section 4.1) soit par l'entremise de nœuds tiers (*c.f.* Section 4.1), les ontologies utilisées par les

22. <https://www.w3.org/TR/shacl/>

23. <https://www.w3.org/TR/shacl-af/>



différentes parties ne sont pas nécessairement les mêmes, ou alignées entre elles. Pour permettre l'accès aux données, deux axes d'approche sont possibles : leur duplication, ou leur fédération.

**La duplication de données.** Dans certains cas, les données issues de sources multiples sont stockées par une entité intermédiaire qui les expose ensuite aux requêtes client. Les données sont donc potentiellement toujours disponibles dans leurs sources respectives, et elles sont dupliquées par l'intermédiaire, ce qui fait que le client interroge un interlocuteur unique pour accéder à des données issues de sources multiples. Ce modèle est instancié par le projet FIESTA-IoT, un projet H2020 visant à exposer via une plateforme unique des données liées issues d'un ensemble hétérogène de sources, telles que des villes intelligentes ou des bâtiments instrumentés. Toutes les sources ne produisant pas leurs données en RDF, n'étant pas nécessairement accessibles en continu, et ne stockant pas nécessairement d'historique, les clients viendront directement interroger la plateforme FIESTA-IoT. Celle-ci se propose donc de collecter les données dans leur format original, de les enrichir à l'aide de leur propre vocabulaire, et de les stocker dans son propre espace mémoire pour fournir un historique aux utilisateurs. FIESTA-IoT est donc hébergée sur un nœud Cloud. Dans le cas de systèmes produisant déjà leurs données en RDF, là aussi la plateforme vient les collecter et les stocker dans son propre espace, à condition que celles-ci soient annotées avec l'ontologie prévue par la plateforme. Le problème de l'interopérabilité est donc résolu par l'utilisation d'un vocabulaire unique au sein de la plateforme. Cependant, les sources de données peuvent tout à fait continuer leur activité de manière indépendante après leur intégration à la plateforme. C'est par exemple le cas du bâtiment ADREAM, au LAAS-CNRS, dont les données sont enrichies pour être publiées sur un portail de données ouvertes<sup>24</sup>, en plus d'être publiées dans la plateforme FIESTA annotées avec un vocabulaire différent.

**La fédération de données du SWoT.** Dans le cas de requêtes fédérées sur un ensemble de sources de données, contrairement au cas précédent, les données ne sont pas dupliquées. C'est le moteur de requête qui va agréger les données issues des différentes sources, afin de fournir une réponse à l'utilisateur. Ce type d'approche ne pose pas le problème de la fraîcheur des données qui peut être observé dans le cas de duplication qui ne soit pas faite en continu, mais par lots de manière ponctuelle. Dans les requêtes fédérées, les sources de données originales sont accédées au moment de la requête. Une telle approche, adaptée au SWoT, est proposée dans (Jaiswal and Lefrançois, 2017). Les auteurs proposent d'étendre la clause SERVICE des requêtes SPARQL, permettant l'interrogation fédérée, afin d'y intégrer les fonctionnalités de SPARQL-Generate. De cette façon, toute source de donnée, même non-RDF, peut être intégrée à une requête fédérée, à la seule condition d'exposer les données à travers un point d'accès HTTP. Cependant, les requêtes fédérées ont aussi leurs limitations, et le choix de l'un ou l'autre des modèles dépend des besoins applicatifs. En particulier, une requête fédérée sera

24. <https://syndream.laas.fr>

toujours aussi lente que la plus lente des sources de données interrogée, ce qui peut être un problème dans le cas de données directement exposées par des objets parfois peu fiables. De plus, interroger l'ensemble des sources de données à chaque requête peut générer un trafic réseau (et donc une consommation énergétique) plus importante que d'interroger un dépôt centrale dans lequel les données sont stockées après une unique collecte.

#### 4.4 Découvrir les objets et les services

Les réseaux IoT sont des réseaux dynamiques, dans lesquels les changements de topologie dûs à des déplacements, des connexions et des déconnexions de nœuds sont fréquents, notamment dans les couches Fog et Device. Cette accessibilité intermittente des entités amène la question de la découverte : comment, à un instant donné, identifier les nœuds offrant un service utile à une application ? Il est important de remarquer dans ce cas le rôle double des ontologies mises en jeu : non seulement elles décrivent les données métiers manipulées par les objets, mais aussi les objets eux-même. La distribution des données évoquée dans la Section 4.3 peut donc à la fois porter sur les données collectées ou sur la description des objets. Ce second cas est au cœur de la contribution de (Charpenay et al., 2018a). Le cas d'utilisation que présentent les auteurs de ce papier repose sur la description d'objets à partir de SOSA/SSN, et de Thing Description, les deux vocabulaires de référence du W3C. Chaque objet est dépositaire de sa propre description, et pour découvrir l'ensemble des objets et de leurs relations, le client envoie une requête en mode broadcast à l'ensemble du réseau. Il collecte les différentes descriptions envoyées dans les réponses à cette requête, les agrège, et exécute l'étape de découverte en attribuant une description sémantique aux nœuds physiques. La question de la découverte des objets peut aussi être comprise dans les protocoles standards. C'est par exemple le cas dans oneM2M, qui définit une procédure de découverte par laquelle un client peut récupérer l'ensemble des ressources connues par un serveur et dont la description est validée par une requête SPARQL.

#### 4.5 Respect de la vie privée

Le développement des réseaux d'objets connectés, et la présence de plus en plus pervasive d'objets dans notre environnement pose d'évidentes questions de respect de la vie privée. Au-delà des questions de sécurisation des objets eux-même, qui est déjà une problématique en soi, les multiples fuites de données récentes amènent à questionner la pertinence de la concentration de masses de données sensibles dans un entrepôt centralisé. C'est pourquoi les approches réparties, très représentées dans les problématiques émergentes (*c.f.* Section 4), sont particulièrement intéressantes. En effet, la distribution du traitement amène un changement de paradigme : les utilisateurs finaux, qui sont ceux qui produisent les données personnelles consommées par leurs fournisseurs de services, peuvent dans ce cas garder la maîtrise de leurs données. Les traitements offerts par les fournisseurs de services sont dans ce type d'approche décorrélés des serveurs contrôlés par ces même fournisseurs, et ne reposent plus sur le transfert de propriété des données par l'utilisateur. En diminuant la concentration des données, le risque de fuite

massive est réduit, ainsi que l'attrait pour un potentiel attaquant. Les ontologies comme acl<sup>25</sup>, ou les travaux comme ceux de (Daga et al., 2015), dans lequel les auteurs automatisent la vérification de la conformité des licences d'utilisation des données tout au long de leur processus d'intégration, sont des pistes par lesquelles les principes et les technologies du Web Sémantique peuvent accompagner la distribution du SWoT pour renforcer le respect de la vie privée.

## 5 Conclusion

L'émergence du SWoT représente un nouveau champ scientifique et technologique où les principes et les technologies du Web Sémantique peuvent représenter un apport majeur. Apporter une réponse aux problèmes d'interopérabilité dus à l'hétérogénéité des objets et des domaines d'application de l'IoT est le moteur du développement du SWoT. Ce besoin explique l'émergence de plusieurs ontologies et standards visant à faciliter l'interaction entre systèmes. Les déploiements d'objets connectés, de par les contraintes qui leurs sont propres, tendent à reproduire un patron architectural en trois niveaux, Cloud-Fog-Device. Dans ce patron architectural, les composants logiciels permettant le contrôle des objets, que l'on peut structurer sous forme d'avatars, sont distribués selon les besoins et les capacités des nœuds qui les supportent. Le fait que les nœuds composant les réseaux IoT aient des capacités si diverses, et en particulier les objets connectés et les fortes contraintes, est l'aspect qui demande le plus d'adaptation des technologies du Web Sémantique pour développer le SWoT. Pour être manipulables par les objets, le RDF doit être exprimé dans des syntaxes plus compactes et plus légères. Il est aussi nécessaire de prendre en compte l'incapacité de certains objets à manipuler du RDF, et d'assurer dans ce cas une traduction dans les couches Fog ou Cloud. La distribution dynamique du raisonnement et des données est aussi un aspect sur lequel des contributions de la communauté amènent un éclairage nouveau. L'adoption de vocabulaires et technologies du Web Sémantique dans des standards issus de communautés dans lesquelles ils n'étaient pas d'habitude intégrés est signe de leur démocratisation toujours plus large. Les contraintes de l'IoT, et les innovations qu'elles amènent dans le développement du SWoT, bénéficient au Web Sémantique tout entier, en faisant considérer à la communauté des défis auxquels elle doit faire face, mais aussi en lui donnant accès à une masse de données dont l'exploitation ne peut être que très riche. Les problématiques émergentes identifiées dans ce papier sont autant de domaines d'intérêt pour la communauté IC, et promettent d'intéressants développements dans le futur.

## Références

Alaya, M. B., Banouar, Y., Monteil, T., Chassot, C., and Drira, K. (2014). OM2M : Extensible ETSI-compliant M2M Service Platform with Self-configuration Capability. In *the 5th Int. Conf. on Ambient Systems Networks and Technologies*.

25. [www.w3.org/ns/auth/acl](http://www.w3.org/ns/auth/acl)

- Ashton, K. (2009). That 'internet of things' thing. *RFID journal*, 22(7).
- Ben-Alaya, M., Medjiah, S., Monteil, T., and Drira, K. (2015). Toward semantic interoperability in oneM2M architecture. *IEEE Communications Magazine*, 53(12).
- Berners-Lee, T., Hendler, J., and Lasilla, O. (2001). The Semantic Web. *Scientific American*, 284(5).
- Bonomi, F., Milito, R., Zhu, J., and Addepalli, S. (2012). Fog Computing and Its Role in the Internet of Things. In *Proc. of the MCC workshop on Mobile cloud computing*.
- Charpenay, V., Käbisch, S., and Kosch, H. (2018a). A framework for semantic discovery on the web of things. In *Emerging Topics in Semantic Technologies, ISWC 2018 Satellite Events*.
- Charpenay, V., Kabisch, S., and Kosch, H. (2018b).  $\mu$ RDF Store : Towards Extending the Semantic Web to Embedded Devices. In *ESWC*.
- Charpenay, V., Käbisch, S., and Kosch, H. (2018c). Semantic data integration on the web of things. In *IOT 2018*.
- Charpenay, V., Sebastian, K., and Kosch, H. (2018d). Towards a Binary Object Notation for RDF. In *ESWC 2018*.
- Compton, M., Henson, C., Lefort, L., Neuhaus, H., and Sheth, A. (2009). A survey of the semantic specification of sensors. In *SSN 2009*, volume 522.
- Daga, E., D'Aquin, M., Gangemi, A., and Motta, E. (2015). Propagation of Policies in Rich Data Flows. In *K-CAP 2015*.
- Daniele, L., den Hartog, F., and Roes, J. (2015). Created in close interaction with the industry : the smart appliances reference (saref) ontology. In *FOMI 2015*.
- Dastjerdi, A. V. and Buyya, R. (2016). Fog Computing : Helping the Internet of Things Realize Its Potential. *Computer*, 49(8).
- Galloway, B. and Hancke, G. P. (2013). Introduction to industrial control networks. *Communications Surveys & Tutorials, IEEE*, 15(2).
- Haller, A., Janowicz, K., Cox, S. J., Lefrançois, M., Taylor, K., Le Phuoc, D., Lieberman, J., García-Castro, R., Atkinson, R., and Stadler, C. (2018). Sosa : A lightweight ontology for sensors, observations, samples, and actuators. *Semantic Web Journal*.
- Haller, A., Janowicz, K., Cox, S. J. D., Le Phuoc, D., Taylor, K., and Lefrançois, M. (2017). Semantic Sensor Network Ontology. W3C Recommendation, W3C.
- Jaiswal, S. and Lefrançois, M. (2017). Towards federated queries for web of things devices. In *Workshop on Semantic Interoperability and Standardization in the IoT, SIS-IoT*.

- Jamont, J., Médini, L., and Mrissa, M. (2014). A Web-Based Agent-Oriented Approach to Address Heterogeneity in Cooperative Embedded Systems. In *PAAMS 2014*, volume 293 of *Advances in Intelligent Systems and Computing*.
- Jamont, J.-P. (2016). Multi-agent approach, models and tools to collective cyber-physical system engineering. Habilitation thesis, universit  Grenoble Alpes.
- Janowicz, K., Haller, A., Cox, S. J., Le Phuoc, D., and Lefran ois, M. (2018). Sosa : A lightweight ontology for sensors, observations, samples, and actuators. *Journal of Web Semantics*.
- Kaebisch, S., Kamiya, T., McCool, M., and Charpenay, V. (2019). Web of Things (WoT) Thing Description. Candidate Recommendation, W3C.
- Kovatsch, Matthias and Matsukura, R., Lagally, M., Kawaguchi, T., Toumura, K., and Kajimoto, K. (2019). Web of Things (WoT) Architecture. Candidate Recommendation, W3C.
- Lee, E. A. (2008). Cyber physical systems : Design challenges. In *ISORC 2008*. IEEE.
- Lefort, L., Henson, C., and Taylor, K. (2011). Semantic Sensor Network XG Final Report. W3C Incubator Group Report, W3C.
- Lefran ois, M. (2017). Planned ETSI SAREF Extensions based on the W3C&OGC SOSA/SSN-compatible SEAS Ontology Patterns. In *SIS-IoT 2017*.
- Lefran ois, M. (2018). Rdf presentation and correct content conveyance for legacy services and the web of things. In *Proc. of IOT 2018*.
- Lefran ois, M., Zimmermann, A., and Bakerally, N. (2017). A SPARQL Extension for Generating RDF from Heterogeneous Formats. In *ESWC*. Springer, Cham.
- Loseto, G., Ieva, S., Gramegna, F., Ruta, M., Scioscia, F., Sciascio, E. D., and I, B. (2016). Linked Data (in low-resource) Platforms : a mapping for Constrained Application Protocol. In *ISWC*.
- Mah o, Y., Le Sommer, N., Launay, P., Guidec, F., and Dragone, M. (2012). Beyond Opportunistic Networking Protocols : a Disruption-Tolerant Application Suite for Disconnected MANETs. In *ExtremeCom 2012*, Z rich, Switzerland. ACM.
- Mell, P. and Grance, T. (2011). The NIST Definition of Cloud Computing Recommendations of the National Institute of Standards and Technology. *National Institute of Standards and Technology, Information Technology Laboratory*, 145.
- Mrissa, M., M dini, L., Jamont, J., Sommer, N. L., and Laplace, J. (2015). An avatar architecture for the web of things. *IEEE Internet Computing*, 19(2).
- Mrissa, M., M dini, L., and Jamont, J.-P. (2014). Semantic Discovery and Invocation of Functionalities for the Web of Things. In *IEEE International Conference on Enabling Technologies : Infrastructure for Collaborative Enterprises*.
- Murdock et al, P. (2016). Semantic Interoperability for the Web of Things. Technical Report October.
- Poveda Villal n, M., Nguyen, Q.-D., Roussey, C., de Vaulx, C., and Chanet, J.-P. (2018). Ontological requirement specification for smart irrigation systems : A SOSA/SSN and SAREF comparison. In *SSN 2018*.
- Sagar, S., Lefran ois, M., Rebai, I., Maha, K., Garlatti, S., Feki, J., and M dini, L. (2018). Modeling smart sensors on top of SOSA/SSN and WoT TD with the semantic smart sensor network (S3N) modular ontology. In *Emerging Topics in Semantic Technologies, ISWC 2018 Satellite Events*.
- Seydoux, N., Alaya, M. B., Hernandez, N., Monteil, T., and Haemmerl , O. (2015). S mantique et Internet des objets : d'un  tat de l'art   une ontologie modulaire. In *IC 2015*.
- Seydoux, N., Drira, K., Hernandez, N., and Monteil, T. (2016a). IoT-O, a core-domain IoT ontology to represent connected devices networks. In *EKAW*.
- Seydoux, N., Drira, K., Hernandez, N., and Monteil, T. (2016b). Lowering knowledge : Making constrained devices semantically interoperable. In *ISWC*.
- Seydoux, N., Drira, K., Hernandez, N., and Monteil, T. (2018a). Towards Cooperative Semantic Computing : a Distributed Reasoning approach for Fog-enabled SWoT. In *COOPIS*.
- Seydoux, N., Khalil, D., Hernandez, N., and Monteil, T. (2018b). A Distributed Scalable Approach for Rule Processing : Computing in the Fog for the SWoT. In *Web intelligence*, Santiago, Chili.
- Sheth, A., Henson, C., and Sahoo, S. S. (2008). Semantic Sensor Web. *IEEE Internet Computing*, 12(4).
- Su, X., Li, P., Riekk , J., Liu, X., Kiljander, J., Soininen, J.-P., Prehofer, C., Flores, H., and Li, Y. (2018). Distribution of Semantic Reasoning on the Edge of Internet of Things. In *IEEE UbiComp*, number November.
- Szilagyi, I. and Wira, P. (2016). Ontologies and Semantic Web for the Internet of Things - a survey. In *IECON*. IEEE.
- Terdjimi, M., M dini, L., Mrissa, M., and Le Sommer, N. (2016). An Avatar-based Adaptation Workflow for the Web of things. In *WETICE 2016*, Paris, France.
- Wang, F., Hu, L., Zhou, J., and Zhao, K. (2015). A survey from the perspective of evolutionary process in the internet of things. *Int. J. Distributed Sensor Networks*, 2015.



# Une Étude Empirique de la Capacité de Généralisation des Plongements de Mots Contextuels en Extraction d'Entités

Bruno Taillé<sup>1,2</sup>Vincent Guigue<sup>2</sup>Patrick Gallinari<sup>2</sup><sup>1</sup> BNP Paribas<sup>2</sup> Sorbonne Université, CNRS, Laboratoire d'Informatique de Paris 6, LIP6

{prenom.nom}@lip6.fr

## Résumé

*Les plongements de mots contextuels utilisent la capacité des modèles de langue à tirer parti de données non annotées pour apprendre des représentations de mots dépendant de leur contexte. Ils sont utiles à la généralisation, particulièrement en Reconnaissance d'Entités Nommées où détecter des mentions d'entités jamais rencontrées pendant l'entraînement est crucial. Malheureusement, les benchmarks actuels surestiment l'importance des caractéristiques lexicales par rapport aux caractéristiques contextuelles à cause d'un recoupement lexical non raisonnable entre mentions d'entraînement et d'évaluation. Dans cet article, nous proposons de mieux évaluer la capacité de généralisation des modèles en séparant les mentions par degré de nouveauté ainsi qu'avec une évaluation extra domaine. Nous montrons que les plongements contextuels sont surtout bénéfiques pour la détection des mentions non rencontrées pendant l'entraînement et mènent à une amélioration relative de +9% du score micro-F1 extra domaine contre +2% intra domaine.*

## Mots Clef

Reconnaissance d'Entités Nommées, Plongements Contextuels, Adaptation de Domaine.

## Abstract

*Contextualized embeddings use unsupervised language model pretraining to compute representations of words depending on their context. This is useful for generalization, especially in Named-Entity Recognition where it is crucial to detect mentions never seen during training. However standard benchmarks overestimate the importance of lexical over contextual features because of an unrealistic lexical overlap between train and test mentions. In this paper, we perform an empirical analysis of the generalization capabilities of state-of-the-art contextualized embeddings by separating mentions by novelty and with out-of-domain evaluation. We show that they are particularly beneficial for unseen mentions, especially out-of-domain where they result in a maximal +9% relative micro-F1 score increase against +2% in domain on CoNLL03.*

## Keywords

Named-Entity Recognition, Contextualized Embeddings, Domain Adaptation

## 1 Introduction

La Reconnaissance d'Entités Nommées (REN) consiste à détecter les mentions textuelles d'entités et à les classer selon des types prédéfinis. Cette tâche est modélisée comme de l'étiquetage de séquence dont l'architecture neuronale classique est le BiLSTM-CRF [7]. Les progrès récents de l'état de l'art proviennent principalement de l'utilisation de nouvelles formes de représentations : des plongements de mots appris à l'échelle des caractères [9] ou des plongements de mots contextuels obtenus par des modèles de langues à l'échelle des mots [14, 5] ou des caractères [1].

Cette dernière approche d'apprentissage par transfert utilise des représentations de mots apprises par des modèles de langue pour réduire la dépendance aux données annotées spécifiques à une tâche ou un domaine [6, 16]. BERT [5] établit même l'état de l'art actuel avec une simple projection linéaire des états cachés appris par un modèle de langue affiné pour la tâche. Parallèlement, l'adaptation de domaine en REN est souvent limitée au pré-entraînement d'un modèle sur des données sources dont les prédictions sont utilisées comme entrées d'un second réseau ou bien qui est réentraîné sur les données cibles [10, 17].

Dans cet article, nous montrons que les jeux de données CoNLL03 et OntoNotes découragent la généralisation aux entités non rencontrées à cause d'un recoupement lexical non réaliste entre mentions d'entraînement et d'évaluation. Cela conduit à surestimer l'importance des caractéristiques lexicales par rapport aux caractéristiques contextuelles. Aussi, nous proposons de mieux évaluer les capacités de généralisation d'une part en séparant les mentions par degré de nouveauté et d'autre part avec une évaluation extra domaine. Dans ce cadre, nous montrons que les plongements de mots contextuels sont surtout bénéfiques pour la détection des mentions non rencontrées pendant l'entraînement et mènent à une amélioration maximale relative de

	CoNLL03					ON	ON réaligné					ON réaligné / CoNLL03				
	LOC	MISC	ORG	PER	Tous	Tous	LOC	MISC	ORG	PER	Tous	LOC	MISC	ORG	PER	Tous
Exact	82%	67%	54%	14%	52%	67%	87%	93%	54%	49%	69%	70%	78%	18%	16%	42%
Partiel	4%	11%	17%	43%	20%	24%	6%	2%	32%	36%	20%	7%	10%	45%	46%	28%
Nul	14%	22%	29%	43%	28%	9%	7%	5%	14%	15%	11%	23%	12%	38%	38%	30%

TABLE 1 – Recouplement lexical des occurrences de mentions des jeux de test avec les jeux d’entraînement respectifs pour CoNLL03 et OntoNotes original et réaligné. La dernière colonne montre le recouplement entre le test de OntoNotes réaligné et l’entraînement de CoNLL03 dans l’évaluation extra domaine.

+9% du score micro-F1 extra domaine contre +2% intra domaine sur CoNLL03. Cela permet d’établir une borne inférieure simple mais efficace d’adaptation de domaine sans données cibles qui pourrait être améliorée en incorporant ces dernières.

## 2 Jeux de données

**CoNLL03** La partie anglaise de CoNLL03 [18] est le benchmark standard en REN et est composé d’articles Reuters datés de 1996 et annotés pour quatre types : Organisation (ORG), Personne (PER), Localité (LOC) et Divers (MISC).

**OntoNotes 5.0** OntoNotes 5.0 [21] est composé de documents de six domaines annotés pour la REN et la Résolution de Coréférence. Il est annoté manuellement pour onze types d’entités et sept types de valeurs qui sont généralement traités sans distinction. La partition entraînement/test classique pour la REN [19] est la même que celle de la tâche de Résolution de Coréférence de CoNLL-2012 [15].

## 3 Recouplement Lexical

Les modèles de REN neuronaux reposent sur des caractéristiques lexicales sous la forme de plongements de mots, qu’ils soient appris au niveau des caractères ou non. Bien que la syntaxe soit incorporée par leur pré-entraînement non supervisé ou l’architecture du réseau, nous prétendons que CoNLL03 et OntoNotes évaluent mal la capacité de généralisation des algorithmes à cause d’un important recouplement lexical entre les mentions présentes dans le jeu d’entraînement et les jeux de validation et de test. Nous le quantifions en séparant les occurrences de mentions dans les sets d’évaluation en trois catégories : recouplement exact, recouplement partiel et recouplement nul, de manière similaire à Augenstein et al. [2].

Une mention d’un jeu d’évaluation est un recouplement exact si elle apparaît sous l’exacte même forme sensible à la capitalisation dans le jeu d’entraînement et annotée avec le même type. Le recouplement est partiel s’il n’est pas exact mais qu’au moins un des mots non vides de la mention apparaît dans une mention de même type. Toutes les autres mentions ont un recouplement nul : leurs mots non vides ne sont jamais rencontrés pendant l’entraînement. Ainsi, la proportion de recouplements partiels et nuls reflète la capacité d’un jeu de test à évaluer la capacité de généralisation d’un algorithme aux entités non rencontrées,

ce qui est un premier pas nécessaire à l’adaptation de domaine.

Comme reporté dans la Table 1, les deux jeux de données montrent un important recouplement lexical de mentions. Dans CoNLL03, plus de la moitié des occurrences de mentions du jeu de test est présente dans le jeu d’entraînement alors que seulement 28% sont totalement nouvelles. Dans OntoNotes, le recouplement est encore pire avec 67% de recouplement exact contre 9% de nouvelles mentions. De plus, nous remarquons une influence significative du type d’entité puisque LOC et MISC présentent le recouplement le plus important alors que PER et ORG ont un vocabulaire plus varié.

Cela montre que les deux principaux jeux de données étalons en REN en anglais évaluent surtout la performance d’extraction des mentions déjà rencontrées lors de l’entraînement, bien qu’apparaissant dans des phrases différentes. De telles proportions de recouplement lexical ne sont pas réalistes dans des applications réelles où un modèle doit traiter quelques ordres de grandeurs de documents de plus en inférence qu’en entraînement pour rentabiliser le coût de l’annotation. L’amélioration spécifique des performances sur les nouvelles mentions revêt donc une importance cruciale dans un cas concret qui est sous-estimée par les benchmarks actuels.

Nous proposons donc une évaluation extra domaine en entraînant les modèles sur CoNLL03 et en les testant sur OntoNotes, plus grand et plus diversifié, ce qui correspond mieux au cas concret. Nous gardons les types de CoNLL03 et y alignons ceux d’OntoNotes : ORG et PER correspondent déjà et nous alignons LOC + GPE dans OntoNotes à LOC dans CoNLL et NORP + LANGUAGE à MISC. Cela réduit le recouplement exact à 42%, ce qui nous semble encore une surestimation du recouplement en utilisation réelle.

## 4 Représentations de Mots

**Plongements de mots classiques** Nous prenons GloVe [13] comme base de référence des plongements traditionnels. Bien que les plongements GloVe soient calculés sur un corpus important pour capturer une similarité sémantique basée sur la co-occurrence, cette représentation est purement lexicale puisque chaque mot est aligné à une unique représentation. Les plongements sont initialisés avec GloVe 840B et leurs valeurs sont affinées pendant l’entraînement.

Entraînement	Modèle	Représentation	Dim	CoNLL03				OntoNotes			
				Exact	Partiel	Nul	Tous	Exact	Partiel	Nul	Tous
CoNLL03	<b>BiLSTM-CRF</b>	BERT	4096	95.7	88.8	82.2	90.5	95.1	82.9	73.5	<b>85.0</b>
		ELMo	1024	95.9	89.2	85.8	<b>91.8</b>	94.3	79.2	72.4	83.4
		Flair	4096	95.4	88.1	83.5	90.6	94.0	76.1	62.1	79.0
		GloVe + char	350	95.3	85.5	83.1	89.9	93.9	73.9	60.4	77.9
		GloVe	300	95.1	85.3	81.1	89.3	93.7	73.0	57.4	76.9
	<b>Map-CRF</b>	BERT	4096	93.2	85.8	73.7	86.2	93.5	77.8	67.8	80.9
		ELMo	1024	93.7	87.2	80.1	<b>88.7</b>	93.6	79.1	69.5	<b>82.2</b>
		Flair	4096	94.3	85.1	78.6	88.1	93.2	74.0	59.6	77.5
		GloVe + char	350	93.1	80.7	69.8	84.4	91.8	69.3	55.6	74.8
		GloVe	300	92.2	77.0	61.7	81.5	89.6	62.8	38.5	68.1
OntoNotes	<b>BiLSTM-CRF</b>	BERT	4096					96.9	88.6	81.1	<b>93.5</b>
		ELMo	1024					97.1	88.0	79.9	93.4
		Flair	4096					96.7	85.8	75.0	92.1
		GloVe + char	350					96.3	83.3	69.9	91.0
		GloVe	300					96.2	82.9	63.8	90.4

TABLE 2 – Scores micro-F1 séparés par degré de recouplement en évaluation intra et extra domaine. Nos résultats sont obtenus en moyennant cinq entraînements.

**Plongements de mots à l’échelle des caractères** Nous reproduisons le **Char-BiLSTM** de Lample et al. [9], un BiLSTM au niveau de chaque mot qui apprend sa représentation à partir des plongements de ses caractères pour tenir compte de caractéristiques orthographiques et morphologiques. Le Char-BiLSTM est entraîné conjointement au réseau de REN et ses sorties sont concaténées aux plongements GloVe.

**Plongements de mots contextuels** Contrairement aux représentations précédentes, les plongements de mots contextuels prennent en compte le contexte d’un mot dans sa représentation. Pour se faire, un modèle de langue est préentraîné sur un corpus non annoté et on prend sa représentation interne de la prédiction d’un mot sachant son contexte. **ELMo** [14] utilise un réseau convolutif à l’échelle des caractères (Char-CNN) pour obtenir un plongement de mot indépendant du contexte et la concaténation de modèles de langue LSTM à deux couches en sens avant et inverse pour la contextualisation. **BERT** [5] adopte des plongements de sous-mots et apprend une représentation dépendant des contextes droits et gauches en entraînant l’encodeur d’un Transformer [20] pour un modèle de langue masqué et la prédiction de la phrase suivante. Nous utilisons le modèle “BERT<sub>LARGE</sub> feature-based” pour une comparaison plus juste : les poids du modèle de langue sont gelés et nous concaténons les états cachés de ses quatre dernières couches. **Flair** [1] emploie directement un modèle de langue à l’échelle du caractère. Comme pour ELMo, deux modèles de langue LSTM de sens opposés sont entraînés et leurs sorties concaténées. Flair et ELMo sont préentraînés sur le 1 Billion Word Benchmark [3] alors que BERT l’est sur la réunion de Book Corpus [22] et Wikipedia en anglais.

## 5 Expériences

### 5.1 Cadre Expérimental

Pour effectuer la REN, nous plaçons les représentations de mots dans deux modèles : un BiLSTM-CRF [7] avec une dimension cachée de 100 dans chaque direction et Map-CRF pour lequel elles sont projetées linéairement dans l’espace de sortie. Nous gardons le CRF [8] car la projection de plongements non contextuels revient à une prédiction indépendante pour chaque mot.

Nous séparons les Précision, Rappel et score F1 par degré de recouplement exact, partiel ou nul. Pour la Précision, cette séparation est effectuée a posteriori sur les prédictions du modèle. Nous utilisons le schéma d’annotations IOBES et validons sur le score-micro F1. Nous rapportons les moyennes des scores obtenus par 5 entraînements différents. Pour chaque modèle, nous choisissons le meilleur de SGD ou Adam avec un taux d’apprentissage de 0.001, des batchs de taille 64, un dropout de 0.5 et un early stopping avec patience 5.

Les scores F1 intra et extra domaine des modèles entraînés sur CoNLL03 sont rapportés dans la Table 2 ainsi que les bornes supérieures extra domaine obtenues en entraînement sur OntoNotes réaligné. Nous omettons délibérément l’évaluation extra domaine de OntoNotes vers CoNLL03 en considérant que les cas d’application concrets sont toujours limités en ressources annotées et ainsi entraîné sur le jeu de données le plus petit et le moins varié.

### 5.2 Résultats

**Performances Intra Domaine** Tout d’abord, dans toutes les configurations le score F1 est le plus haut pour les recouplements exacts, puis partiels et nuls ce qui confirme le biais dans les jeux de données avec un recouplement lexical important. Ensuite, bien que ELMo apparaît comme

	LOC				MISC				ORG				PER				Tous
	Exact	Partiel	Nul	Tous	Exact	Partiel	Nul	Tous	Exact	Partiel	Nul	Tous	Exact	Partiel	Nul	Tous	Tous
<b>CoNLL03 → CoNLL03</b>																	
BERT	96.1	68.7	76.5	92.1	93.6	53.5	46.9	79.7	95.2	82.0	79.1	88.0	99.0	98.1	93.2	96.2	90.5
ELMo	96.0	72.6	83.3	93.1	94.3	58.9	53.9	81.8	96.0	80.8	83.5	89.6	98.9	98.7	94.6	97.0	<b>91.8</b>
Flair	95.7	73.3	79.7	92.3	93.5	55.9	49.2	80.2	95.1	77.9	80.9	87.8	98.3	98.3	93.3	96.2	90.6
GloVe + char	95.6	64.1	80.5	91.8	93.3	54.0	40.8	78.9	94.9	74.4	82.0	87.5	98.7	97.2	92.0	95.2	89.9
<b>CoNLL03 → OntoNotes</b>																	
BERT	96.1	65.7	79.0	89.6	94.1	51.2	25.8	72.6	93.6	83.6	76.4	82.6	93.5	90.3	83.4	88.2	<b>85.0</b>
ELMo	94.9	63.0	77.7	88.5	94.6	56.5	37.8	78.8	92.8	80.8	74.5	80.5	91.9	84.5	75.6	82.2	83.4
Flair	95.3	59.7	67.8	86.2	94.0	52.8	28.4	74.2	89.3	77.2	59.9	72.6	92.0	82.1	70.1	78.8	79.0
GloVe + char	95.6	62.4	69.3	86.7	93.8	56.7	30.1	75.3	88.9	74.0	58.5	70.8	89.5	78.9	64.7	74.8	77.9

TABLE 3 – Scores F1 par type du BiLSTM-CRF entraîné sur CoNLL03 en évaluation intra et extra domaine.

la solution la plus stable intra domaine, il est difficile de dégager une hiérarchie claire entre plongements contextuels puisque les données de pré-entraînement ainsi que la dimension des représentations diffèrent. Pour BERT et Flair, le BiLSTM-CRF performe relativement moins bien sur CoNLL03 que sur OntoNotes, probablement par sur-apprentissage sur CoNLL03. De plus, le gain maximal de la contextualisation sur CoNLL03 est de +0.6 F1 en recoupement exact contre +3.7 en partiel et +2.7 en nul. D’autre part, Map-CRF avec ELMo ou Flair arrive presque au même niveau que BiLSTM-CRF et GloVe + char, ce qui montre que les modèles de langues capturent intrinsèquement des représentations utiles à la REN. Enfin, quelle que soit la représentation le BiLSTM réduit l’écart de performance entre mentions vues et non vues.

**Généralisation Extra Domaine** En évaluation extra domaine, les performances se dégradent et l’écart se creuse entre les mentions vues et non vues. De plus, la contextualisation est encore plus bénéfique aux mentions non vues avec +1.2 F1 en recoupement exact, +9.0 en partiel et +13.1 en nul avec le BiLSTM-CRF et BERT. Cette amélioration provient clairement du pré-entraînement du modèle de langue puisque même avec Map-CRF, les plongements contextuels atteignent au moins 77.5 F1 contre 77.9 pour BiLSTM-CRF et GloVe + char. Nous distinguons néanmoins une séparation entre plongements contextuels puisque Flair, issu d’un modèle de langue à l’échelle des caractères, généralise moins bien que ELMo ou BERT en extra domaine pour les deux modèles. Il ressort ainsi que contextualiser des mots ou sous-mots conduit à une meilleure généralisation en REN. Enfin, nous pouvons séparer les performances par genres des documents dans OntoNotes comme rapporté dans la Table 4. Pour tous les modèles, la meilleure adaptation se fait pour le type *broadcast news* qui est plus proche du domaine de CoNLL03 que *web text* ou *magazine*. Cependant, les plongements contextuels bénéficient principalement aux genres plus distants et mènent à des résultats plus homogènes.

**Influence du Type** Bien que le score micro-F1 soit souvent la seule métrique rapportée en REN, les types d’entités devraient être pris en compte. Comme montré dans la Table 3, en intra domaine MISC est le type le plus difficile à détecter alors que PER est le plus facile, certainement grâce

	bc	bn	nw	mz	tc	wb	Tous
BERT	87.2	88.4	84.7	82.4	84.5	79.5	<b>85.0</b>
ELMo	85.0	88.6	82.9	78.1	84.0	79.9	83.4
Flair	78.0	86.5	80.4	71.1	73.5	72.1	79.0
GloVe + char	80.4	86.3	77.0	70.7	79.7	69.2	77.9

TABLE 4 – Scores micro-F1 extra domaine du BiLSTM-CRF par genres. Respectivement *broadcast conversation*, *broadcast news*, *news wire*, *magazine*, *telephone conversation* et *web text*.

à un motif prénom-nom fréquent. Cependant, la contextualisation bénéficie homogènement à tous les types en intra domaine alors qu’elle bénéficie surtout à ORG et PER en extra domaine. Cela s’explique par moins de 18% de recoupement exact avec le set d’entraînement contre plus de 70% pour LOC et MISC. Ainsi, la contextualisation est plus utile pour la généralisation aux types avec le plus de variation lexicale même quand ils sont plus faciles à détecter en intra domaine.

## 6 Travaux Connexes

Augenstein et al. [2] présentent une étude quantitative de deux modèles basés sur les CRF et un réseau convolutif avec des plongements de mots classiques [4] sur sept jeux de données dont CoNLL03 et OntoNotes. Ils séparent notamment les performances sur les mentions rencontrées en entraînement (notre recoupement exact) de celles non rencontrées et montrent une chute du score F1 sur ces dernières.

Moosavi et Strube [11] soulèvent un phénomène similaire en Résolution de Coréférence sur CoNLL-2012 et montrent qu’en évaluation extra domaine l’écart de performance entre les modèles d’apprentissage profond et un système de règles disparaît. Dans [12], ils proposent d’utiliser des caractéristiques linguistiques (comme le genre, le type d’entité ou la catégorie grammaticale) pour améliorer la généralisation extra domaine. Néanmoins, de telles caractéristiques sont obtenues en utilisant des modèles à leur tour entraînés avec des caractéristiques lexicale et sur des données ou le même problème de recoupement lexical se pose, au moins pour la Reconnaissance d’Entités Nommées.



## 7 Conclusion

Les benchmarks actuels de REN sont donc biaisés en faveur des mentions déjà rencontrées, à l’exact opposé des applications concrètes. D’où la nécessité de séparer les performances par degré de recouplement des mentions pour mieux évaluer les capacités de généralisation. Dans ce cadre, les plongements contextuels bénéficient plus significativement aux mentions non rencontrées, d’autant plus en extra domaine.

Les travaux futurs peuvent chercher à réduire encore l’écart de performance entre mentions rencontrées ou non, améliorer les capacités d’adaptation de domaine zero-shot avec des données cibles additionnelles ou aborder la généralisation multilingue en utilisant des modèles de langues entraînés sur des corpus multilingues.

## Remerciements

Nous remercions Geoffrey Scoutheeten et Victor Storchan pour leurs points de vue et commentaires précieux. Ce travail est principalement financé par BNP Paribas dans le cadre de la convention CIFRE 2018/0327.

## Références

- [1] A. Akbik, D. Blythe, and R. Vollgraf. Contextual string embeddings for sequence labeling. In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 1638–1649, 2018.
- [2] I. Augenstein, L. Derczynski, and K. Bontcheva. Generalisation in named entity recognition : A quantitative analysis. *Computer Speech & Language*, 44 :61–83, 7 2017.
- [3] C. Chelba, T. Mikolov, M. Schuster, Q. Ge, T. Brants, P. Koehn, and T. Robinson. One Billion Word Benchmark for Measuring Progress in Statistical Language Modeling. *arXiv preprint arXiv :1312.3005*, 2013.
- [4] R. Collobert and J. Weston. Natural language processing (almost) from scratch. *Journal of Machine Learning Research*, 12 :2493–2537, 2011.
- [5] J. Devlin, M.-W. Chang, K. Lee, and K. Toutanova. BERT : Pre-training of Deep Bidirectional Transformers for Language Understanding. *arXiv preprint arXiv :1810.04805*, 2018.
- [6] J. Howard and S. Ruder. Universal Language Model Fine-tuning for Text Classification. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics*, pages 328–339, 2018.
- [7] Z. Huang, W. Xu, and K. Yu. Bidirectional LSTM-CRF Models for Sequence Tagging. *arXiv preprint arXiv :1508.01991*, 2015.
- [8] J. Lafferty, A. McCallum, and F. C. N. Pereira. Conditional Random Fields : Probabilistic Models for Segmenting and Labeling Sequence Data. In *Proceedings of the 18th International Conference on Machine Learning*, volume 8, pages 282–289, 2001.
- [9] G. Lample, M. Ballesteros, S. Subramanian, K. Kawakami, and C. Dyer. Neural Architectures for Named Entity Recognition. In *Proceedings of NAACL-HLT 2016*, pages 260–270, 2016.
- [10] J. Y. Lee, F. Dernoncourt, and P. Szolovits. Transfer Learning for Named-Entity Recognition with Neural Networks. In *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC-2018)*, pages 4470–4473, 2018.
- [11] N. S. Moosavi and M. Strube. Lexical Features in Coreference Resolution : To be Used With Caution. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics*, pages 14–19, 2017.
- [12] N. S. Moosavi and M. Strube. Using Linguistic Features to Improve the Generalization Capability of Neural Coreference Resolvers. In *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 193–203, 2018.
- [13] J. Pennington, R. Socher, and C. D. Manning. GloVe : Global Vectors for Word Representation. *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing*, 2014.
- [14] M. E. Peters, M. Neumann, M. Iyyer, M. Gardner, C. Clark, K. Lee, and L. Zettlemoyer. Deep contextualized word representations. In *Proceedings of the 2018 Conference of the North American Chapter of the Association for Computational Linguistics : Human Language Technologies, Volume 1 (Long Papers)*, pages 2227–2237, 2 2018.
- [15] S. Pradhan, A. Moschitti, N. Xue, O. Uryupina, and Y. Zhang. CoNLL-2012 shared task : Modeling multilingual unrestricted coreference in OntoNotes. *Joint Conference on EMNLP and CoNLL-Shared Task. Association for Computational Linguistics*, pages 1–40, 2012.
- [16] A. Radford, T. Salimans, K. Narasimhan, T. Salimans, and I. Sutskever. Improving Language Understanding by Generative Pre-Training. page 12, 2018.
- [17] J. D. Rodriguez, A. Caldwell, and A. Liu. Transfer Learning for Entity Recognition of Novel Classes. In *Proceedings of the 27th International Conference on Computational Linguistics*, pages 1974–1985, 2018.
- [18] E. F. T. K. Sang and F. De Meulder. Introduction to the CoNLL-2003 shared task. In *Proceedings of the seventh Conference on Natural Language Learning at NAACL-HLT 2003*, volume 4, pages 142–147, 2003.
- [19] E. Strubell, P. Verga, D. Belanger, and A. McCallum. Fast and Accurate Entity Recognition with Iterated Dilated Convolutions. In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 2670–2680, 2017.
- [20] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. Kaiser, and I. Polosukhin.

- Attention Is All You Need. In *Advances in Neural Information Processing Systems*, pages 5998–6008, 2017.
- [21] R. Weischedel, M. Palmer, M. Marcus, E. Hovy, S. Pradhan, L. Ramshaw, N. Xue, A. Taylor, J. Kaufman, and M. Franchini. OntoNotes Release 5.0 LDC2013T19. *Linguistic Data Consortium, Philadelphia, PA*, 2013.
- [22] Y. Zhu, R. Kiros, R. Zemel, R. Salakhutdinov, R. Urtasun, A. Torralba, and S. Fidler. Aligning Books and Movies : Towards Story-like Visual Explanations by Watching Movies and Reading Books. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 19–27, 2015.

# La génération automatique de poésie en français

Tim Van de Cruys

IRIT & CNRS

118 Route de Narbonne, 31062 Toulouse Cedex 9, France

tim.vandecruys@irit.fr

## Résumé

La génération automatique de poésie est une tâche ardue pour un système informatique. Pour qu'un poème ait du sens, il est important de prendre en compte à la fois des aspects linguistiques et littéraires. Ces dernières années, un certain nombre d'approches fructueuses sont apparues, capables de modéliser de manière adéquate divers aspects du langage naturel. En particulier, les modèles de langue basés sur les réseaux de neurones ont amélioré l'état de l'art par rapport à la modélisation prédictive de langage, tandis que les *topic models* sont capables de capturer une certaine cohérence thématique. Dans cet article, on explorera comment ces approches peuvent être adaptées et combinées afin de modéliser les aspects linguistiques et littéraires nécessaires pour la génération de poésie. Le système est exclusivement entraîné sur des textes génériques, et sa sortie est contrainte afin de conférer un caractère poétique au vers généré. Le cadre présenté est appliqué à la génération de poèmes en français, et évalué à l'aide d'une évaluation humaine.

## Mots Clefs

génération de poésie, réseaux de neurones, factorisation en matrices non-négatives

## Abstract

Automatic poetry generation is a challenging task for a computational system. For a poem to be meaningful, both linguistic and literary aspects need to be taken into account. In the last few years, a number of successful approaches have emerged that are able to adequately model various aspects of natural language. Particularly, language models based on neural networks have improved the state of the art with regard to predictive language modeling, while topic models are able to capture some form of thematic coherence. In this article, we will explore how these approaches can be adapted and combined to model the linguistic and literary aspects needed for poetry generation. The system is exclusively trained on generic text, and its output is constrained in order to confer a poetic character to the generated verse. The framework is applied to the generation of poems in French, and it is evaluated using a human evaluation.

## Keywords

poetry generation, neural networks, non-negative matrix factorization

## 1 Introduction

La génération automatique de poésie est une tâche ardue pour un système informatique. Pour qu'un poème ait du sens, il est important de prendre en compte à la fois des aspects linguistiques et littéraires. Tout d'abord, un système de génération de poésie doit modéliser de manière correcte les phénomènes de langage, tels que la syntaxe, et la cohérence sémantique et discursive. De plus, le système doit intégrer diverses contraintes (telles que la forme et la rime) liées à un genre poétique particulier. Enfin, le système doit faire preuve d'une certaine créativité littéraire, ce qui rend le poème intéressant et digne d'être lu.

Ces dernières années, dans le domaine du traitement automatique des langues, un certain nombre d'approches fructueuses sont apparues, capables de modéliser de manière adéquate divers aspects du langage naturel. En particulier, les modèles de langue basés sur les réseaux de neurones ont amélioré l'état de l'art par rapport à la modélisation prédictive de langage, tandis que les *topic models* sont capables de capturer une certaine forme de cohérence thématique. Dans cet article, on explorera comment ces approches peuvent être adaptées et combinées afin de modéliser les aspects linguistiques et littéraires nécessaires pour la génération de poésie. Plus spécifiquement, dans ce travail, on utilisera des réseaux de neurones récurrents dans une configuration encodeur-décodeur. L'encodeur construit d'abord une représentation d'une phrase entière en incorporant séquentiellement les mots de cette phrase dans un vecteur d'état caché de taille fixe. La représentation finale est ensuite donnée au décodeur, qui émet une séquence de mots selon une distribution de probabilité dérivée de l'état caché de la phrase en entrée. En apprenant au réseau à prédire la phrase suivante avec la phrase actuelle en entrée, le réseau apprend à générer du texte brut avec une certaine cohérence discursive. En transformant la distribution de probabilité fournie par le décodeur, afin d'incorporer des contraintes poétiques, le réseau peut être exploité pour la génération de vers poétiques. Il est important de noter que le système

de poésie n'est pas entraîné sur des textes poétiques; au contraire, le système est entraîné sur des textes génériques extraits du web, et ce seront alors les contraintes appliquées qui confèrent un caractère poétique aux vers générés.

Cet article est structuré comme suit. Dans la section 2, on présente un aperçu des travaux connexes sur la génération automatique de poésie. La section 3 décrit ensuite les différentes composantes du système de génération de poésie. Dans la section 4, on présentera un certain nombre d'exemples et une évaluation humaine. La section 5 conclut et examine quelques pistes pour des futurs travaux.

## 2 Travaux connexes

Il y a une longue et captivante histoire en termes de génération automatique de poésie pour le français [21, 18], que l'on qualifierait de créativité mécanique. Au-delà de la simple créativité mécanique, les premières implémentations informatiques se sont souvent appuyées sur des méthodes basées sur des règles ou sur des patrons. L'un des premiers exemples est le système ASPERA [5] pour l'espagnol, qui repose sur une base de connaissances complexe, un ensemble de règles et un raisonnement à partir de cas. D'autres approches incluent [16], qui combinent la génération basée sur des règles avec des algorithmes génétiques; le système de génération PoeTryMe de [7], qui repose sur la génération tabulaire (*chart generation*) et diverses stratégies d'optimisation; et [25], qui exploite les expressions métaphoriques en utilisant une approche basée sur les patrons.

Alors que la génération de poésie avec des modèles basés sur des règles et des patrons a une tendance inhérente à être structurellement plutôt rigide, les progrès des méthodes statistiques pour la génération de langage ont ouvert de nouvelles perspectives pour une approche plus variée et hétérogène. [9], par exemple, utilisent un modèle de langage n-gramme en combinaison avec un modèle rythmique implémenté avec des transducteurs à états finis. Et plus récemment, des réseaux de neurones récurrents ont été exploités pour la génération de la poésie. [28] utilise un encodeur-décodeur RNN pour la génération de poésie chinoise, dans lequel un premier RNN construit une représentation cachée du vers actuel dans un poème, et un deuxième RNN prédit le vers suivant mot par mot, en fonction de la représentation cachée du vers actuel. Le système est entraîné sur un corpus de poèmes chinois. [27] présente une amélioration de l'approche encodeur-décodeur en incorporant une méthode de raffinement itératif : le réseau construit un poème candidat à chaque itération, et la représentation de l'itération précédente est utilisée lors de la création de la suivante. Et [26] étendent la méthode en utilisant un mécanisme d'attention.

[6] combinent des RNNs (afin de modéliser la fluidité syntaxique) avec des calculs de similarité distributionnelle (afin de modéliser la cohérence sémantique) et des automates à états finis (pour imposer des contraintes littéraires telles que le mètre et la rime). Leur système,

HAFEZ, est capable de produire des poèmes bien formés avec un raisonnable degré de cohérence sémantique, basés sur un sujet défini par l'utilisateur. [11] se concentrent sur les vers rythmiques; ils combinent un RNN, entraîné sur une représentation phonétique de poèmes, avec une cascade de transducteurs à états finis pondérés. Et [13] présentent un modèle de réseaux de neurones pour la génération de sonnets, qui intègre l'entraînement de la rime et du rythme dans le réseau; le réseau apprend les motifs de stress iambiques à partir de données, tandis que les paires de mots qui riment sont séparées des paires de mots qui ne riment pas en utilisant une perte basée sur la marge.

Il est à noter que tous les modèles statistiques existants sont entraînés sur un corpus de poésie; à notre connaissance, notre système est le premier à ne réaliser la génération de poésie qu'avec un modèle exclusivement entraîné sur un corpus générique, ce qui signifie que le caractère poétique est conféré par le modèle lui-même. Deuxièmement, on utilise un modèle sémantique latent pour modéliser la cohérence thématique, ce qui est également nouveau.

## 3 Modèle

### 3.1 Architecture neuronale

À la base du système de poésie se trouve un modèle de langage neuronal, entraîné à prédire la phrase suivante  $S_{i+1}$  à partir de la phrase courante  $S_i$ . L'architecture neuronale est composée de deux réseaux de neurones récurrents à portes (*gated recurrent units*, ou GRUs; Cho *et al.*, 2014) fonctionnant dans une configuration encodeur-décodeur. L'encodeur prend en séquence chaque mot  $w_{1,\dots,N}^i$  de la phrase courante  $S_i$  (représenté par son plongement de mot ou *word embedding*  $\mathbf{x}$ ) de manière qu'à chaque pas de temps  $t_i$  un état caché  $\mathbf{h}_t$  est créé à la base du plongement du mot courant  $\mathbf{x}_t$  et l'état caché  $\mathbf{h}_{t-1}$  du pas de temps précédent. Pour chaque pas de temps, l'état caché  $\hat{\mathbf{h}}_t$  est calculé selon les équations suivantes :

$$\mathbf{r}_t = \sigma(\mathbf{W}_r \mathbf{x}_t + \mathbf{U}_r \hat{\mathbf{h}}_{t-1}) \quad (1)$$

$$\mathbf{z}_t = \sigma(\mathbf{W}_z \mathbf{x}_t + \mathbf{U}_z \hat{\mathbf{h}}_{t-1}) \quad (2)$$

$$\bar{\mathbf{h}}_t = \tanh(\mathbf{W} \mathbf{x}_t + \mathbf{U}(\mathbf{r}_t \odot \hat{\mathbf{h}}_{t-1})) \quad (3)$$

$$\hat{\mathbf{h}}_t = (1 - \mathbf{z}_t) \odot \hat{\mathbf{h}}_{t-1} + \mathbf{z}_t \odot \bar{\mathbf{h}}_t \quad (4)$$

où  $\mathbf{r}_t$  représente la porte de réinitialisation du GRU,  $\mathbf{z}_t$  représente la porte de mise à jour,  $\bar{\mathbf{h}}_t$  représente le nouveau état candidat, et  $\odot$  représente la multiplication élément par élément.

$\mathbf{h}_t$  peut être interprété comme une représentation de la séquence  $w_1, \dots, w_t$ , et l'état caché final  $\mathbf{h}_N$  sera donc une représentation de la phrase entière. Cet état caché final est ensuite donné comme entrée au décodeur. Le décodeur fait alors une prédiction mot par mot de la phrase suivante, conditionnée sur l'encodeur; à chaque pas de temps  $t_{i+1}$ , le décodeur crée également un état caché  $\mathbf{h}_t$  à la base du plongement  $\mathbf{x}_t$  du mot courant (prédit par le

décodateur dans le pas précédent) et l'état caché  $\mathbf{h}_{t-1}$  du pas de temps précédent (le premier état caché étant  $\mathbf{h}_N$  qui vient de l'encodeur et le premier mot étant un symbole d'initialisation). Les calculs pour chaque pas de temps  $\mathbf{h}_t$  du décodeur sont égaux à ceux utilisés dans l'encodeur (équations 1 à 4).

Afin d'exploiter pleinement la séquence complète de représentations fournie par l'encodeur, l'architecture de base est complétée par un mécanisme d'attention, notamment l'attention dite *générale* [15]. Le mécanisme d'attention permet au décodeur de consulter l'ensemble des états cachés calculés par l'encodeur; à chaque pas de temps – pour la génération de chaque mot de la phrase  $S_{i+1}$  – le décodeur détermine quels mots de la phrase  $S_i$  sont pertinents et sélectionne en conséquence une combinaison linéaire de l'ensemble des états cachés. À cette fin, on calcule d'abord un vecteur d'attention  $\mathbf{a}_t$ , qui attribue un poids à chaque état masqué  $\hat{\mathbf{h}}_i$  produit par l'encodeur (en fonction de l'état caché actuel du décodeur  $\mathbf{h}_t$ ) selon l'équation 5 :

$$\mathbf{a}_t(i) = \frac{\exp(\text{score}(\mathbf{h}_t, \hat{\mathbf{h}}_i))}{\sum_{i'} \exp(\text{score}(\mathbf{h}_t, \hat{\mathbf{h}}_{i'}))} \quad (5)$$

où

$$\text{score}(\mathbf{h}_t, \hat{\mathbf{h}}_i) = \mathbf{h}_t^T \mathbf{W}_a \hat{\mathbf{h}}_i \quad (6)$$

L'étape suivante consiste à calculer un vecteur de contexte global  $\mathbf{c}_t$ , qui est une moyenne pondérée (basée sur le vecteur d'attention  $\mathbf{a}_t$ ) de tous les états masqués de l'encodeur. Le vecteur de contexte qui en résulte est ensuite combiné avec l'état caché du décodeur d'origine afin de calculer un nouvel état caché augmenté avec l'attention,  $\tilde{\mathbf{h}}_t$  :

$$\tilde{\mathbf{h}}_t = \tanh(\mathbf{W}_c[\mathbf{c}_t; \mathbf{h}_t]) \quad (7)$$

où  $[\cdot; \cdot]$  représente la concaténation des vecteurs. Enfin, l'état caché qui en résulte  $\tilde{\mathbf{h}}_t$  est transformé en distribution de probabilité  $p(\mathbf{w}^t | \mathbf{w}^{<t}, S_i)$  sur le vocabulaire entier en utilisant une couche softmax.

$$p(\mathbf{w}^t | \mathbf{w}^{<t}, S_i) = \text{softmax}(\mathbf{W}_s \tilde{\mathbf{h}}_t) \quad (8)$$

Comme fonction objective, on optimise la somme des log-probabilités de la phrase suivante, conditionnée sur la représentation cachée de l'encodeur de la phrase actuelle.

$$J_t = \sum_{(S_i, S_{i+1}) \in \mathcal{C}} -\log p(S_i | S_{i+1}) \quad (9)$$

Au moment de l'inférence, pour la génération d'un vers, chaque mot est ensuite échantillonné de manière aléatoire en fonction de la distribution de probabilité de sortie. De manière cruciale, le décodeur est entraîné à prédire les mots de la phrase suivante en sens inverse, de sorte que le dernier mot du vers soit le premier généré. Cette opération inverse est importante pour une incorporation efficace de

la rime, comme cela sera expliqué dans la section suivante. Une représentation graphique de l'architecture, qui inclut les contraintes discutées ci-dessous, est donnée dans la figure 1.

### 3.2 Contraintes poétiques comme distributions *a priori*

Étant donné que l'architecture neuronale ci-dessus est entraînée sur des textes génériques, sa sortie ne ressemblera en rien à un poème; afin de doter la sortie générée d'un certain caractère poétique, on modifiera la distribution de probabilité de sortie du réseau de neurones en appliquant une distribution de probabilité *a priori*. On modélisera deux types de contraintes : une contrainte de rime et une contrainte thématique.

**Contrainte de rime.** Pour la modélisation de la contrainte de rime, on s'appuie sur une représentation phonétique des mots, extraite de manière automatique depuis le *Wiktionnaire* pour le français. Pour chaque mot, on détermine son rime (c'est-à-dire le groupe de voyelles final, éventuellement suivi d'un groupe de consonnes), ainsi que la groupe de consonnes précédente. Un échantillon de rimes ainsi extraites est donné dans le tableau 1.

mot	rime
reproduit	( 'dʁ', 'i' )
thérapie	( 'p', 'i' )
examen	( 'm', 'ɛ' )
canadien	( 'dj', 'ɛ' )

TABLE 1: Exemples de rimes extraits du *Wiktionnaire*

L'étape suivante consiste à créer une distribution de probabilité *a priori* pour un son de rime requis :

$$p(\mathbf{w})_{rime} = \frac{1}{Z} \mathbf{x} \text{ avec } \begin{cases} x_i = 1 & \text{if } i \in R \\ x_i = \varepsilon & \text{otherwise} \end{cases} \quad (10)$$

où  $R$  est l'ensemble des mots avec le son de rime requis,  $\varepsilon$  est une valeur très petite pour éviter les erreurs de calcul, et  $Z$  est une constante de normalisation pour assurer une distribution de probabilité. On est maintenant en mesure d'appliquer la distribution de probabilité *a priori* afin de repondérer la distribution de probabilité de sortie du réseau de neurones selon la formule 11, chaque fois que le schéma de rimes le requiert :

$$p(\mathbf{w})_{out} = \frac{1}{Z} (p(\mathbf{w}^t | \mathbf{w}^{<t}) \odot p(\mathbf{w})_{rime}) \quad (11)$$

avec  $\odot$  étant la multiplication élément par élément. Rappelons que chaque vers est généré à l'envers; la repondération par rapport à la rime est appliquée tout au début de la génération, et le mot rime est généré en premier. Cela empêche la génération d'un mot rime maladroît qui ne correspond pas au reste du vers.

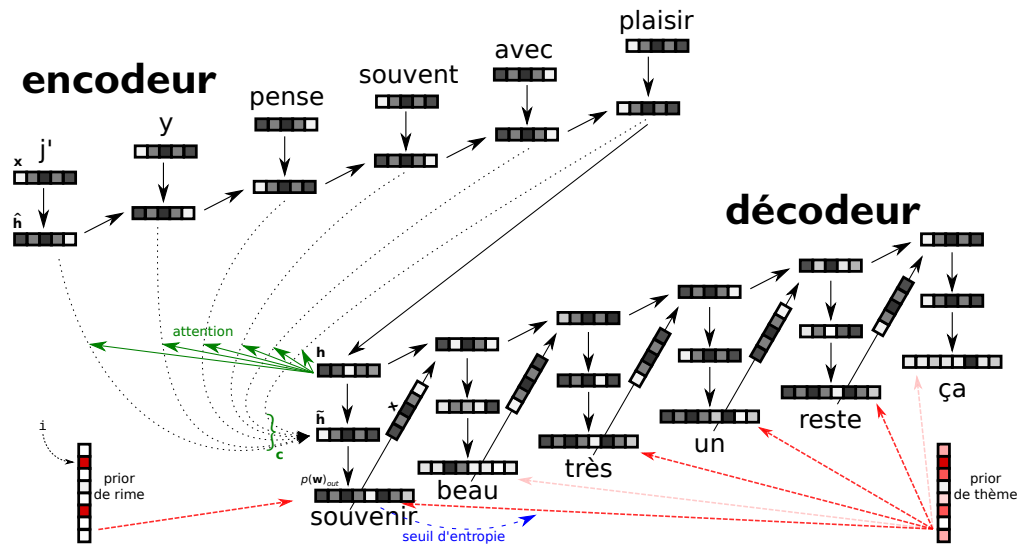


FIGURE 1: Représentation graphique du modèle de génération de poésie. L’encodeur traite le vers actuel mot par mot, et la représentation finale est donnée au décodeur, qui prédit le vers suivant mot par mot, à l’envers. Le mécanisme d’attention est représenté pour le premier pas de temps. La distribution de probabilité *a priori* pour le rime est appliquée au premier pas de temps, et celle pour le thème est facultativement appliquée à tous les pas de temps, en fonction de la valeur d’entropie de la distribution de sortie du réseau.

**Contrainte thématique.** Pour la modélisation de la contrainte thématique, on s’appuie sur un modèle de sémantique latente sous forme d’une factorisation en matrices non négatives (NMF; Lee & Seung, 2001). Des recherches antérieures ont montré que la méthode est capable de produire des dimensions thématiques bien claires et interprétables [17]. Comme entrée, on construit une matrice de fréquence **A**, qui capture les fréquences<sup>1</sup> de co-occurrence des mots du vocabulaire et leurs contextes. Cette matrice est alors factorisée en deux autres matrices non négatives, **W** et **H**.

$$A_{i \times j} \approx W_{i \times k} H_{k \times j} \quad (12)$$

où  $k$  est beaucoup plus petit que  $i, j$ , de manière que les instances et les traits sont exprimés par un nombre limité de dimensions. De manière cruciale, la factorisation en matrices non négatives impose la contrainte que les trois matrices doivent être non négatives, c’est-à-dire tous les éléments doivent être supérieurs ou égaux à zéro. En utilisant la minimisation de la divergence de Kullback-Leibler comme fonction objective, on veut trouver les matrices **W** et **H** pour lesquelles la divergence entre **A** et **WH** (la multiplication de **W** et **H**) est la plus petite. Cette factorisation est réalisée par l’application itérative de règles de mis à jour. Quelques exemples de dimensions, extraits avec la méthode, sont représentés dans le tableau 2.

La factorisation issue du modèle NMF peut être interprétée de manière probabiliste [4, 3] : la matrice **W** peut être

1. Les fréquences brutes sont pondérées en utilisant l’information mutuelle spécifique (*pointwise mutual information*; Bullinaria & Levy, 2007; Turney & Pantel, 2010).

dim 1	dim 20	dim 25	dim 90
tendresse	gare	hypocrisie	désespoir
joie	bus	mensonge	terrible
bonheur	métro	accuser	colère
sourires	tram	hypocrite	angoisse
baisers	rer	tort	violente
amour	tgw	arrogance	désarroi
joies	tramway	critiquer	frustration
merveilleux	autoroute	mensonges	souffrance
nostalgie	autobus	bêtises	humiliation
douceur	boulevard	reprocher	impuissance

TABLE 2: Exemples de dimensions thématiques issues de NMF (10 mots les plus saillants)

considérée comme  $p(w|k)$ , c’est-à-dire la probabilité d’un certain mot  $w$  du vocabulaire, étant donnée la dimension latente  $k$ . On pourrait maintenant facilement utiliser cette distribution comme une autre distribution *a priori* thématique, appliquée à chaque sortie; cependant, une telle modification à l’aveugle de la distribution de probabilité de sortie pour chaque mot de la séquence pose des problèmes par rapport à la structure syntaxique. Pour pallier à cela, on conditionne la modification de la distribution de sortie par le calcul d’une valeur d’entropie sur cette distribution : lorsque l’entropie de la distribution de sortie est faible, le réseau de neurones connaît la choix du mot correct afin de générer une phrase bien formée, donc on ne le changera pas. En revanche, lorsque l’entropie de la distribution de sortie est élevée, on modifie la distribution en utilisant la distribution thématique  $p(w|k)$  d’une dimension latente comme distribution *a priori* (analogue à la formule 11),

afin d’insérer la thématique souhaitée. Le seuil d’entropie, au-dessus duquel on utilise la distribution modifiée, est défini expérimentalement.

Notez que la contrainte de rime et la contrainte thématique peuvent facilement être combinées afin de générer un mot de rime thématique, en multipliant les trois distributions concernées, puis en procédant à une normalisation.

### 3.3 Cadre d’optimisation global

La génération d’un vers est réalisé dans un cadre d’optimisation global. On intègre le modèle de génération dans un cadre d’optimisation pour deux raisons. Premièrement, la génération d’un vers est un processus d’échantillonnage, sujet au hasard. Le cadre d’optimisation nous permet de choisir le meilleur échantillon en fonction des contraintes définies ci-dessus. Deuxièmement, l’optimisation nous permet de définir quelques critères supplémentaires qui aident dans la sélection du meilleur vers. Pour chaque vers final généré, le modèle génère un nombre considérable de vers candidats; chaque candidat est alors noté en fonction des critères suivants :

- la log-probabilité du vers généré, en fonction de l’architecture encodeur-décodeur (section 3.1);
- respect de la contrainte de rime (section 3.2); de plus, l’extraction du groupe de consonnes précédent (cf. tableau 1) permet de donner un score plus élevé aux mots rimes avec des groupes de consonnes précédents disparates, ce qui permet d’obtenir des rimes plus intéressantes;
- respect de la contrainte thématique (section 3.2); le score est modélisé comme la somme des probabilités de tous les mots pour la dimension définie;
- le nombre optimal de syllabes, modélisé comme une distribution gaussienne avec une moyenne  $\mu$  et un écart-type  $\sigma$ ;<sup>2</sup>
- la log-probabilité d’un modèle de n-grammes standard.

Le score de chaque critère est normalisé à l’intervalle  $[0, 1]$  à l’aide d’une normalisation *min-max*, et la moyenne harmonique<sup>3</sup> de tous les scores est considérée comme le score final de chaque candidat. Après la génération d’un nombre prédéfini de candidats, le candidat avec le score optimal est conservé et ajouté au poème.

## 4 Résultats et évaluation

### 4.1 Détails de mise en œuvre

L’architecture neuronale a été entraînée sur un corpus de textes web en français à caractère général, construit à base

2. On a également mené des expériences avec des contraintes basées sur le mètre et les pieds de vers, mais les premières expériences indiquaient que le système avait tendance à produire des vers très rigides. Un simple comptage des syllabes tend à donner une variation plus intéressante.

3. La moyenne harmonique est calculée par  $\frac{n}{\sum_{i=1}^n \frac{1}{x_i}}$ ; elle est choisie pour balancer les différents scores.

du corpus CommonCrawl<sup>4</sup>. Le corpus dans son intégralité contient 11 milliards de mots; cependant, on effectue un certain nombre d’étapes de filtrage afin de ne conserver que des paires de phrases propres :

- on ne garde que des phrases de 20 mots maximum;
- on ne garde que des phrases qui contiennent au moins un mot fonction (par exemple, les pronoms communs) d’une liste prédéfinie, l’idée étant de ne garder que des vraies phrases et de filtrer le bruit;
- de toutes les phrases qui restent après les deux premières étapes de filtrage, on ne garde que les phrases qui apparaissent successivement dans un document.

Après filtrage, la taille du corpus est réduite à 400 million de mots. On utilise un vocabulaire de 15 000 mots (sélectionnés par rapport à leur fréquence); au-delà, les mots sont remplacés par un token `<unk>` (dont la probabilité est fixée à zero pendant la phase de génération). L’encodeur ainsi que le décodeur sont tous les deux constitués de deux couches de GRUs avec un état caché de 2048, et la taille de plongements de mots est de 512; les plongements d’encodeur, de décodeur, et de sortie sont partagés [20]. On optimise les paramètres du modèle en utilisant une descente de gradient stochastique, partant d’un taux d’apprentissage de 0,2, qui est divisé par 4 lorsque la fonction de coût n’améliore plus sur un ensemble de validation. On utilise un *batch size* de 64, et on applique du *gradient clipping*. L’architecture neuronale a été implémentée en utilisant PyTorch [19], en s’appuyant considérablement sur le module OpenNMT [12]. Par rapport à la contrainte thématique, on utilise un seuil d’entropie de 2,70.

Le modèle n-gramme utilisé est un modèle standard d’ordre 3 lissé par Kneser-Ney, entraîné en utilisant *KenLM* [10]. Le modèle NMF est factorisé en 100 dimensions, la matrice de fréquences étant construite avec une fenêtre de phrases, et en utilisant la divergence de Kullback-Leibler comme objective. Le modèle n-gramme ainsi que le modèle NMF sont entraînés sur l’intégralité du corpus sans filtrage. Pour la contrainte de nombre de syllabes, on utilise  $\mu = 12, \sigma = 2$ .

On génère environ 2000 candidats pour chaque vers, selon un schéma de rimes fixe (ABAB CDCD). Quatre exemples représentatifs de poèmes générés par le système sont montré dans la figure 2. Notez qu’aucune sélection humaine n’a été effectuée sur les poèmes utilisés pour l’évaluation; tous les poèmes ont été générés en une seule fois, sans *cherry picking*.

### 4.2 Évaluation humaine

L’évaluation quantitative de la créativité est loin d’être simple, et cela n’est pas moins vrai pour les artefacts créatifs qui sont générés de manière automatique. Des mesures d’évaluation automatique qui calculent la similarité de la sortie du système avec des textes de

4. commoncrawl.org

Malgré mon enthousiasme, le chagrin s'allonge  
 Le bonheur est toujours superbe  
 Toi, tu es un merveilleux songe  
 Je te vois rêver de bonheur dans l'herbe

Tu trouveras le bonheur de tes rêves  
 Je t'aime comme tout le monde  
 Je t'aime mon amour, je me lève  
 Je ressens pour toi une joie profonde

~

La route vers la ville est imprenable  
 Nous décidons de prendre le bus vers le tram  
 De plus la station de métro est très agréable  
 Je suis en train de rentrer dans ma rame

La gare, plusieurs personnes m'observent  
 Je suis allée dans la rue des portes  
 Je m'aperçois que le tgv, ça énerve  
 Si je la voie, c'est que la sncf est morte

~

Rien ne prouve qu'il s'indigne  
 Dans le cas contraire, ce n'est pas grave  
 Si la vérité est fausse, c'est très mauvais signe  
 Il est vrai que les gens le savent

Et cela est faux, mais qu'importe  
 En fait, le mensonge, c'est l'effroi  
 La négation de l'homme en quelque sorte  
 Le tort n'est pas de penser cela, il est magistrat

~

Hélas, après sa mort, ce fut elle qui cède  
 Ce fut un moment d'une angoisse extrême  
 Un sentiment d'incompréhension, mais sans remède  
 Une peur en colère, et parfois même

Un véritable sentiment de panique  
 Ce qui provoque une rage étrange  
 Il s'ensuit un drame tragique  
 On sent la tragédie qui, sans excès, s'arrange

FIGURE 2: Quatre exemples représentatifs de poèmes générés par le système; les poèmes, de haut en bas, ont été générés respectivement en utilisant la dimension 1, 20, 25, et 90 (cf. tableau 2).

référence standard (telles que BLEU ou ROUGE), et qui pourraient être utilisés pour évaluer les tâches de génération standard, peuvent difficilement être qualifiées d'appropriées quand il s'agit de génération créative. C'est l'une des raisons pour lesquelles la plupart des chercheurs ont fait recours à une évaluation humaine, bien qu'il faille garder à l'esprit que l'évaluation de créativité textuelle est par nature subjective, en particulier en ce qui concerne la valeur poétique. Pour une discussion sur le sujet et un aperçu des différentes méthodes d'évaluation, voir [8].

Dans cette recherche, on adopte le cadre d'évaluation de [28], dans lequel il est demandé aux annotateurs d'évaluer les poèmes sur une échelle de cinq points, en fonction d'un certain nombre de caractéristiques, à savoir :

- *fluidité* : le poème est-il grammatical et syntaxiquement bien formé ?
- *cohérence* : le poème est-il structuré

thématiquement ?

- *signification* : le poème transmet-il un message significatif au lecteur ? Le poème a-t-il un sens ?
- *caractère poétique* : le texte affiche-t-il les caractéristiques d'un poème ?

En plus, on demande aux annotateurs de juger si le poème est écrit par un humain ou un ordinateur.

Au total, on évalue six ensembles de poèmes différents, issus de différentes instanciations de modèles. Les différents ensembles de poèmes pris en compte lors de l'évaluation sont les suivants :

1. *random* : des poèmes générés par un modèle de référence aléatoire où, pour chaque vers, on sélectionne de manière aléatoire une phrase, entre 7 et 15 mots, dans un grand corpus ; l'idée est que les phrases sélectionnées par le modèle de référence seront assez fluides (puisqu'elles proviennent d'un corpus réel), mais sans cohérence (en raison de leur sélection aléatoire) ;
2. *rnn* : des poèmes générés par l'architecture neuronale décrit en section 3.1, sans aucune contrainte supplémentaire ;
3. *rime* : des poèmes générés par l'architecture neuronale, augmenté avec la contrainte de rime ;
4. *nmf<sub>rand</sub>* : des poèmes générés par l'architecture neuronale, augmentée à la fois avec la contrainte de rime et la contrainte thématique, où l'une des dimensions NMF (induites de manière automatique) est sélectionnée de manière aléatoire ;
5. *nmf<sub>spec</sub>* : des poèmes générés par l'architecture neuronale, augmentée à la fois avec la contrainte de rime et la contrainte thématique, où l'une des dimensions NMF (induites de manière automatique) est spécifiée manuellement<sup>5</sup> ;
6. *humain* : poèmes écrits par des humains<sup>6</sup>.

22 annotateurs ont évalué 30 poèmes au total (5 pour chacun des six modèles évalués), de sorte que chaque poème a été évalué par au moins 4 annotateurs. Les résultats sont présentés dans le tableau 3.

Tout d'abord, on remarque que tous les modèles fonctionnent mieux que le modèle de référence aléatoire, même en ce qui concerne la fluidité syntaxique ( $p < 0,01$  en utilisant un test de permutation bilatéral ; notez que le modèle de référence est constituée de phrases réelles). Les bons scores obtenus pour nos modèles avec contraintes (*rime* et *nmf<sub>\*</sub>*) indiquent que l'application de contraintes ne nuit pas à la grammaticalité des vers. Deuxièmement, on constate que la contrainte de rime améliore le caractère

5. Ceci peut être considéré comme définissant manuellement le thème du poème généré. La dimension spécifiée est sélectionnée pour son caractère poétique. On prend la dimension 1 du tableau 2 comme dimension spécifiée.

6. On prend des poèmes avec le même schéma de rimes que les poèmes générés, parmi les poèmes les mieux classés sur le site short-edition.com.



modèle	fluidité	cohérence	signification	caractère poétique	écrit par humain (%)
<i>random</i>	2,95	1,86	1,68	2,18	0,00
<i>rnn</i>	3,45	2,73	2,59	2,55	0,27
<i>rime</i>	<b>3,82</b>	2,55	2,18	3,23	0,14
<i>nmf<sub>rand</sub></i>	3,64	3,32	3,09	2,86	0,27
<i>nmf<sub>spec</sub></i>	<b>3,82</b>	<b>3,82</b>	<b>3,55</b>	<b>3,95</b>	<b>0,45</b>
<i>humain</i>	4,59	4,59	4,50	4,81	0,95

TABLE 3: Résultats de l'évaluation humaine (score moyenne pour tous les annotateurs)

poétique ( $p < 0,05$  vis-à-vis *rnn*), et que la contrainte thématique améliore à la fois la cohérence ( $p < 0,05$ ) et la signification ( $p < 0,01$ ). On note également que le score pour caractère poétique est considérablement plus élevé ( $p < 0,01$ ) pour *nmf<sub>spec</sub>* (avec un thème qu'on pourrait juger poétique) que pour *nmf<sub>rand</sub>* (avec un thème aléatoire, que l'on considérerait souvent comme plus banal). Finalement, on constate que les meilleurs scores par rapport à tous les critères sont obtenus avec le modèle *nmf<sub>spec</sub>*, pour lesquels les poèmes sont jugés être écrits par un humain dans presque la moitié des cas.

## 5 Conclusion

Dans cet article, on a présenté un système pour la génération de poésie en français. On utilise des réseaux de neurones en configuration encodeur-décodeur pour générer des vers candidats, en modifiant la distribution de sortie pour incorporer des contraintes littéraires et thématiques. Dans un cadre d'optimisation, on sélectionne alors parmi un nombre de candidats le meilleur vers pour inclusion dans le poème. Les résultats d'une évaluation humaine indiquent que le système est capable de générer des poèmes crédibles, avec des bons scores en termes de fluidité et de cohérence, ainsi qu'en termes de signification et de caractère poétique. Dans notre meilleure configuration, presque la moitié des poèmes générés sont jugés être écrits par un humain. La méthode présentée est générale, ce qui signifie qu'elle peut facilement être étendue à d'autres langues.

Afin de permettre l'utilisation et l'expérimentation ainsi que l'inspection du modèle, le système de génération de poésie est mis à disposition sous forme de logiciel open source. La version actuelle est téléchargeable en utilisant le lien <https://github.com/timvdc/poetry>.

On conclue avec quelques pistes pour des travaux futurs. Tout d'abord, on aimerait explorer différentes architectures de réseaux de neurones. Plus précisément, on pense que des approches hiérarchiques [22] ainsi que le réseau dite *transformer* [24] conviendraient particulièrement à la génération de la poésie. Deuxièmement, on aimerait incorporer d'autres dispositifs poétiques, notamment ceux basés sur le sens. La poésie captivante repose souvent sur l'utilisation d'un langage figuré, tel que le symbolisme et la métaphore. Une incorporation spécifique de tels dispositifs

signifierait un pas important vers une génération de poésie vraiment inspirée.

## Références

- [1] J. A. Bullinaria and J. P. Levy. Extracting semantic representations from word co-occurrence statistics : A computational study. *Behavior research methods*, 39(3) :510–526, 2007.
- [2] K. Cho, B. van Merriënboer, C. Gulcehre, D. Bahdanau, F. Bougares, H. Schwenk, and Y. Bengio. Learning phrase representations using rnn encoder–decoder for statistical machine translation. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 1724–1734. Association for Computational Linguistics, 2014.
- [3] C. Ding, T. Li, and W. Peng. On the equivalence between non-negative matrix factorization and probabilistic latent semantic indexing. *Computational Statistics & Data Analysis*, 52(8) :3913–3927, 2008.
- [4] E. Gaussier and C. Goutte. Relation between pls and nmf and implications. In *Proceedings of the 28th annual international ACM SIGIR conference on Research and development in information retrieval*, pages 601–602. ACM, 2005.
- [5] P. Gervás. An expert system for the composition of formal spanish poetry. In *Applications and Innovations in Intelligent Systems VIII*, pages 19–32, London, 2001. Springer.
- [6] M. Ghazvininejad, X. Shi, Y. Choi, and K. Knight. Generating topical poetry. In *Proceedings of the 2016 Conference on Empirical Methods in Natural Language Processing*, pages 1183–1191, Austin, Texas, November 2016. Association for Computational Linguistics.
- [7] H. Gonçalo Oliveira. Poetryme : a versatile platform for poetry generation. *Computational Creativity, Concept Invention, and General Intelligence*, 1 :21, 2012.
- [8] H. Gonçalo Oliveira. A survey on intelligent poetry generation : Languages, features, techniques,

- reutilisation and evaluation. In *Proceedings of the 10th International Conference on Natural Language Generation*, pages 11–20, 2017.
- [9] E. Greene, T. Bodrumlu, and K. Knight. Automatic analysis of rhythmic poetry with applications to generation and translation. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, pages 524–533. Association for Computational Linguistics, 2010.
- [10] K. Heafield. KenLM : faster and smaller language model queries. In *Proceedings of the EMNLP 2011 Sixth Workshop on Statistical Machine Translation*, pages 187–197, Edinburgh, Scotland, United Kingdom, July 2011.
- [11] J. Hopkins and D. Kiela. Automatically generating rhythmic verse with neural networks. In *Proceedings of the 55th Annual Meeting of the Association for Computational Linguistics (Volume 1 : Long Papers)*, pages 168–178. Association for Computational Linguistics, 2017.
- [12] G. Klein, Y. Kim, Y. Deng, J. Senellart, and A. Rush. Opennmt : Open-source toolkit for neural machine translation. In *Proceedings of ACL 2017, System Demonstrations*, pages 67–72. Association for Computational Linguistics, 2017.
- [13] J. H. Lau, T. Cohn, T. Baldwin, J. Brooke, and A. Hammond. Deep-speare : A joint neural model of poetic language, meter and rhyme. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1 : Long Papers)*, pages 1948–1958. Association for Computational Linguistics, 2018.
- [14] D. D. Lee and H. S. Seung. Algorithms for non-negative matrix factorization. In *Advances in neural information processing systems*, pages 556–562, 2001.
- [15] T. Luong, H. Pham, and C. D. Manning. Effective approaches to attention-based neural machine translation. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pages 1412–1421. Association for Computational Linguistics, 2015.
- [16] R. Manurung, G. Ritchie, and H. Thompson. Using genetic algorithms to create meaningful poetic text. *Journal of Experimental & Theoretical Artificial Intelligence*, 24(1) :43–64, 2012.
- [17] B. Murphy, P. Talukdar, and T. Mitchell. Learning effective and interpretable semantic models using non-negative sparse embedding. In *Proceedings of COLING 2012*, pages 1933–1950. The COLING 2012 Organizing Committee, 2012.
- [18] OULIPO. *Atlas de Littérature Potentielle*. Gallimard, Paris, France, 1981.
- [19] A. Paszke, S. Gross, S. Chintala, G. Chanan, E. Yang, Z. DeVito, Z. Lin, A. Desmaison, L. Antiga, and A. Lerer. Automatic differentiation in pytorch. In *Advances in Neural Information Processing Systems*, 2017.
- [20] O. Press and L. Wolf. Using the output embedding to improve language models. In *Proceedings of the 15th Conference of the European Chapter of the Association for Computational Linguistics : Volume 2, Short Papers*, pages 157–163. Association for Computational Linguistics, 2017.
- [21] R. Queneau. *Cent mille milliards de poèmes*. Gallimard, Paris, France, 1961.
- [22] I. V. Serban, A. Sordoni, R. Lowe, L. Charlin, J. Pineau, A. Courville, and Y. Bengio. A hierarchical latent variable encoder-decoder model for generating dialogues. In *Thirty-First AAAI Conference on Artificial Intelligence*, 2017.
- [23] P. D. Turney and P. Pantel. From frequency to meaning : Vector space models of semantics. *Journal of artificial intelligence research*, 37 :141–188, 2010.
- [24] A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, Ł. Kaiser, and I. Polosukhin. Attention is all you need. In *Advances in Neural Information Processing Systems*, pages 5998–6008, 2017.
- [25] T. Veale. Less rhyme, more reason : Knowledge-based poetry generation with feeling, insight and wit. In *Proceedings of the international conference on computational creativity*, pages 152–159, 2013.
- [26] Q. Wang, T. Luo, D. Wang, and C. Xing. Chinese song iambics generation with neural attention-based model. In *Proceedings of International Joint Conference on Artificial Intelligence*, pages 2943–2949, 2016.
- [27] R. Yan. i, poet : Automatic poetry composition through recurrent neural networks with iterative polishing schema. In *Proceedings of International Joint Conference on Artificial Intelligence*, pages 2238–2244, 2016.
- [28] X. Zhang and M. Lapata. Chinese poetry generation with recurrent neural networks. In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, pages 670–680. Association for Computational Linguistics, 2014.

# Compression de vocabulaire de sens grâce aux relations sémantiques pour la désambiguïsation lexicale

Loïc Vial Benjamin Lecouteux Didier Schwab

Univ. Grenoble Alpes, CNRS, Grenoble INP, LIG, 38000 Grenoble, France

{loic.vial,benjamin.lecouteux,didier.schwab}@univ-grenoble-alpes.fr

## Résumé

*En Désambiguïsation Lexicale (DL), les systèmes supervisés dominent largement les campagnes d'évaluation. La performance et la couverture de ces systèmes sont cependant rapidement limités par la faible quantité de corpus annotés en sens disponibles. Dans cet article, nous présentons deux nouvelles méthodes qui visent à résoudre ce problème en exploitant les relations sémantiques entre les sens tels que la synonymie et l'hyperonymie, afin de compresser le vocabulaire de sens de WordNet, et ainsi réduire le nombre d'étiquettes différentes nécessaires pour désambiguïser tous les mots de la base lexicale. Nos méthodes permettent de réduire considérablement la taille des modèles de DL neuronaux, avec l'avantage d'améliorer leur couverture sans données supplémentaires, et sans impacter leur précision. En plus de nos méthodes, nous présentons un système de DL qui tire parti des récents travaux sur les représentations vectorielles de mots contextualisées, afin d'obtenir des résultats qui surpassent largement l'état de l'art sur toutes les tâches d'évaluation de la DL.*

## Mots Clef

Désambiguïsation lexicale, compression de vocabulaire, relations sémantiques

## Abstract

*In Word Sense Disambiguation (WSD), supervised approaches are predominant in evaluation campaigns. The limited quantity of such corpora however restricts the coverage and the performance of these systems. In this article, we present two new methods that tackle this problem by exploiting the semantic relationships between senses such as synonymy, hypernymy and hyponymy, in order to compress the sense vocabulary of WordNet, and thus reduce the number of different sense tags that must be observed to disambiguate all words of the lexical database. Our methods greatly reduce the size of neural WSD models, with the benefit of improving their coverage without additional training data, and without impacting their precision. In addition to our methods, we present a neural WSD system which relies on the recent advances in contextualized word embeddings in order to achieve results that significantly outperform the state of the art on all WSD evaluation tasks.*

## Keywords

Word Sense Disambiguation, Vocabulary Compression, Semantic Relationships

## 1 Introduction

La Désambiguïsation Lexicale (DL) est une tâche qui vise à clarifier un texte en assignant à chacun de ses mots l'étiquette de sens la plus appropriée depuis un inventaire de sens prédéfini. Il existe diverses approches pour la DL, telles que les approches à base de connaissances, qui s'appuient sur des bases de données lexicales ou des graphes de connaissances, et les méthodes supervisées, qui exploitent des corpus annotés en sens comme données d'apprentissage pour entraîner un classifieur multi-classe tel qu'un SVM ou un réseau neuronal. Les méthodes supervisées sont de loin les plus représentées car elles offrent généralement les meilleurs résultats dans les campagnes d'évaluation (par exemple (Navigli et al., 2007)). Les classifieurs état de l'art combinaient jusqu'à récemment des caractéristiques précises telles que les parties du discours et les lemmes des mots voisins, (Zhong et Ng, 2010), mais ils sont maintenant remplacés par des réseaux de neurones récurrents qui apprennent leur propre représentation des mots (Raganato et al., 2017; Le et al., 2018; Vial et al., 2019). Cependant, une des limitations majeures des systèmes supervisés est la quantité limitée de corpus manuellement annotés en sens. En effet, le SemCor (Miller et al., 1993), qui est le plus grand corpus manuellement annoté en sens disponible, contient 33 760 labels de sens différents, ce qui correspond à seulement environ 16% de l'inventaire de sens de WordNet<sup>1</sup> (Miller et al., 1990), la base de données lexicale de référence largement utilisée en DL. De nombreux travaux tentent de résoudre ce problème via la création de nouveaux corpus annotés en sens, générés soit automatiquement (Pasini et Navigli, 2017), semi-automatiquement (Taghipour et Ng, 2015), ou bien par *crowdsourcing* (Yuan et al., 2016), mais dans nos travaux, nous cherchons à résoudre ce problème en tirant parti des relations sémantiques présentes entre les sens de WordNet comme l'hyperonymie, l'hyponymie, l'antonymie, la méronymie, etc.

1. <https://wordnet.princeton.edu/documentation/wnstats7wn>

Notre méthode s'appuie sur les observations suivantes :

1. un sens et ses sens voisins dans le graphe des relations sémantiques de WordNet véhiculent tous une même idée ou concept, à des niveaux d'abstraction différents ;
2. dans certains cas, un mot peut être désambiguïsé en utilisant seulement les sens voisins de ses sens, et pas nécessairement ses sens propres ;
3. par conséquent, nous n'avons pas besoin de connaître tous les sens de WordNet pour désambiguïser tous les mots de WordNet.

Par exemple, considérons le mot « souris » et deux de ses sens : la souris *d'ordinateur* et la souris *l'animal*. Les notions plus générales comme « être vivant » (hyperonyme de souris/animal) et « appareil électronique » (hyperonyme de souris/ordinateur), permettent déjà de distinguer les deux sens, et toutes les notions plus spécialisées telles que « rongeur » ou « mammifère » sont, elles, superflues. En regroupant ces étiquettes de sens ensemble, on peut bénéficier de tous les autres exemples mentionnant un appareil électronique ou un être vivant dans un corpus d'entraînement, même si le mot « souris » n'est pas mentionné spécifiquement, pour désambiguïser le mot « souris ».

**Contributions :** Dans cet article, nous émettons l'hypothèse que seul un sous-ensemble des sens de WordNet peut être considéré pour pouvoir désambiguïser tous les mots de la base lexicale. Par conséquent, nous proposons deux méthodes différentes pour construire ce sous-ensemble que nous appelons méthodes de compression de vocabulaire de sens. Grâce à ces techniques, nous améliorons considérablement la couverture des systèmes de DL supervisés, en éliminant quasiment le besoin d'une stratégie de repli habituellement employée pour les mots jamais observés pendant l'entraînement. Nous présentons des résultats qui surpassent l'état de l'art de façon significative sur toutes les tâches d'évaluation de la DL, et nous fournissons à la communauté notre outil ainsi que nos meilleurs modèles pré-entraînés, sur un dépôt GitHub dédié<sup>2</sup>.

## 2 Désambiguïstation lexicale neuronale

Plusieurs avancées récentes ont été réalisées dans la création de nouvelles architectures neuronales pour les systèmes supervisés de désambiguïstation lexicale. Ces systèmes atteignent des performances état de l'art et certains peuvent intégrer des sources de connaissances externes. Dans cette section, nous donnons un aperçu de ces travaux.

### 2.1 Approches basées sur un modèle de langue

Dans ce type d'approches, initié par Yuan et al. (2016) et réimplémenté par Le et al. (2018), le composant principal est un modèle de langue neuronal capable de prédire un mot en tenant compte des mots qui l'entourent, grâce à un réseau neuronal entraîné sur une quantité massive de don-

nées non annotées (100 milliards de mots pour Yuan et al. (2016) et 1,8 milliards pour Le et al. (2018)).

Une fois le modèle de langue entraîné, il est utilisé pour produire des vecteurs de sens en moyennant les vecteurs de mots prédits par le modèle à l'endroit où ces mots sont annotés avec un sens particulier. Au moment du test, le modèle de langue est utilisé pour prédire un vecteur en fonction du contexte environnant, et le sens le plus proche du vecteur prédit est attribué à chaque mot.

Ces systèmes ont l'avantage de contourner le problème de l'absence de données annotées en sens en concentrant le pouvoir d'abstraction offert par les réseaux neuronaux récurrents sur un modèle de langue de bonne qualité et entraîné de manière non supervisée. Cependant, ces méthodes souffrent toujours du manque de corpus annotés en sens étant donné qu'ils restent indispensables pour la création des vecteurs de sens.

### 2.2 Approches basées sur un classifieur linéaire et la fonction *softmax*

Dans ces systèmes, le réseau neuronal principal classe et attribue directement un sens à chaque mot donné en entrée à l'aide d'une distribution de probabilité calculée par la fonction *softmax*. Les annotations en sens sont simplement considérées comme des balises placées sur chaque mot, à la manière d'une tâche d'étiquetage en parties du discours par exemple.

On peut distinguer deux branches distinctes de ces types de réseaux neuronaux :

1. Ceux dans lesquels il y a un réseau neuronal (ou classifieur) distinct et spécifique à chaque lemme du dictionnaire (Iacobacci et al., 2016; Kågebäck et Salomonsson, 2016). Chaque classifieur est capable de gérer un lemme particulier avec ses sens. Par exemple, l'un des classifieurs est spécialisé dans le choix entre les quatre sens possibles du nom « souris ». Ce type d'approche est particulièrement adapté aux tâches de *lexical sample*, où seul un petit nombre de mots distincts et très ambigus doivent être annotés dans plusieurs contextes. Mais ils nécessiteraient plusieurs milliers de réseaux différents<sup>3</sup> pour pouvoir aussi être utilisés dans les tâches de désambiguïstation lexicales *all words*, dans lesquelles tous les mots d'un document doivent être annotés en sens.
2. Ceux dans lesquels il y a un seul réseau neuronal, plus grand et capable de gérer tous les lemmes du lexique, qui attribuent à un mot un sens issu de l'ensemble de tous les sens de l'inventaire de sens utilisé (Raganato et al., 2017; Vial et al., 2019).

L'avantage de la première branche est que pour désambiguïser un mot, il est beaucoup plus facile de limiter notre choix à l'un de ses sens possibles que de chercher parmi tous les sens de tous les mots du lexique. Pour se donner une idée, le nombre moyen de sens des mots polysémiques

3. L'ensemble de WordNet contient par exemple 26 896 mots polysémiques (<https://wordnet.princeton.edu/documentation/winstats7wn>)

2. <https://github.com/getalp/disambiguate>

dans WordNet est d'environ 3, alors que le nombre total de sens en considérant tous les mots est 206 941.<sup>4</sup>

La seconde approche a cependant une propriété intéressante : tous les sens résident dans le même espace vectoriel et partagent donc des caractéristiques dans les couches cachées du réseau. Cela permet au modèle de prédire un sens identique pour deux mots différents (synonymes), mais aussi de prédire un sens pour un mot non présent dans le dictionnaire (néologisme, faute d'orthographe, etc.).

Enfin, dans deux articles récents, Luo et al. (2018a,b) ont proposé une amélioration de ce type d'architecture, en calculant une attention entre le contexte d'un mot cible et les définitions de ses différents sens. Ainsi, leur travail est le premier à incorporer les connaissances de WordNet dans un système de désambiguïsation neuronal.

### 3 Compression de vocabulaire de sens

Les systèmes supervisés neuronaux état de l'art tels que Yuan et al. (2016); Raganato et al. (2017); Le et al. (2018); Luo et al. (2018a,b); Vial et al. (2019) sont tous confrontés aux mêmes limitations :

1. La quantité de données annotés manuellement en sens étant très limitée, il se peut qu'un mot cible ne soit jamais observé pendant l'entraînement. Dans ce cas, le système ne peut pas être en mesure de l'annoter, et une stratégie de repli est généralement effectuée (par exemple utiliser le premier sens du mot dans WordNet).
2. Pour la même raison, un mot peut être observé, mais pas tous ses sens. Dans ce cas, le système va être capable d'annoter ce mot, mais si le sens attendu n'a jamais été observé, le résultat sera faux, quelle que soit l'architecture sous-jacente du système supervisé.
3. L'empreinte mémoire des modèles neuronaux ainsi que leur temps d'entraînement et d'exécution augmentent avec la quantité de données d'apprentissage et le nombre d'étiquettes de sens différentes prises en compte, nombre qui monte jusqu'à 206 941 si l'on considère toutes les étiquettes de sens de WordNet.

Afin de résoudre ces problèmes, nous proposons deux nouvelles méthodes permettant de regrouper ensemble des étiquettes de sens qui se réfèrent à des concepts similaires, tout en nous assurant que ces groupes de sens permettent toujours de discriminer les différents sens de tous les mots du lexique, afin de retrouver l'étiquette de sens originale pour un mot au moment de le désambiguïser. En conséquence, le vocabulaire de sens, c'est-à-dire le nombre total d'étiquettes de sens dans notre inventaire de sens diminue, le système est capable de mieux généraliser, et sa couverture augmente.

4. <https://wordnet.princeton.edu/documentation/wNSTATS7wn>

#### 3.1 Des sens aux *synsets* : une première compression de vocabulaire de sens grâce à la synonymie

Dans la base de données lexicale WordNet (Miller et al., 1990), les sens sont organisés en ensembles de synonymes appelés *synsets*. Un *synset* est concrètement un groupe d'un ou plusieurs sens qui ont la même définition et donc la même signification. Par exemple, les premiers sens des mots « *eye* », « *optic* » et « *oculus* » appartiennent tous au même *synset* dont la définition est « l'organe de la vue ».

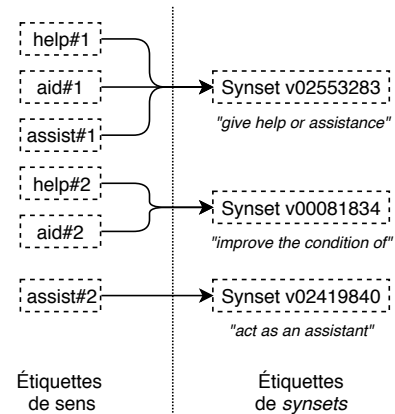


FIGURE 1 – Conversion des étiquettes de sens vers des étiquettes de *synsets*, appliqué aux deux premiers sens des mots « *help* », « *aid* » et « *assist* ». Le nombre de sens différents dans notre vocabulaire passe ainsi de six à trois.

La conversion des étiquettes de sens (« Xème sens du mot N ») aux étiquettes de *synsets* (« *synset* numéro Y »), illustré dans la figure 1, est ainsi une façon de compresser le vocabulaire qui est déjà appliquée dans plusieurs travaux (Yuan et al., 2016; Le et al., 2018; Vial et al., 2019) sans être toujours explicitement précisée. Cette méthode contribue pourtant clairement à améliorer la couverture des systèmes supervisés. En effet, si le verbe « *aid* » annoté avec son premier sens est observé dans les données d'apprentissage, le contexte autour du mot cible peut être aussi utile pour annoter ultérieurement les verbes « *assist* » ou « *help* » avec la même étiquette de *synset*.

En allant plus loin, on peut trouver d'autres informations dans WordNet qui peuvent aider à mieux généraliser. La première nouvelle méthode que nous proposons repose ainsi sur ce même principe de regroupement de sens, mais en exploitant les relations d'hyponymie et d'hyperonymie.

#### 3.2 Compression de vocabulaire de sens grâce aux relations d'hyperonymie, d'hyponymie et d'instance

Selon Polguère (2003), l'hyperonymie et l'hyponymie sont deux relations sémantiques qui correspondent à un cas particulier d'inclusion de sens : l'hyponyme d'un terme est une spécialisation de ce terme, alors que son hyperonyme est une généralisation. Par exemple, une « souris » est un

type de « rongeur » qui est à son tour un type d'« animal ». Dans WordNet, ces relations lient presque tous les noms ensemble allant de la racine générique, le nœud « entité » aux feuilles les plus spécifiques, par exemple « souris à pattes blanches ». Si l'on prend aussi en compte la relation d'instance, qui fonctionne de la même manière mais qui lie les entités nommées aux noms courants (par exemple « Einstein » est une instance de « physicien »), tous les noms de WordNet font partie de cette même hiérarchie. Ces relations sont également présentes sur plusieurs verbes : ainsi, par exemple, « additionner » est une manière de « calculer » qui est à son tour une manière de « raisonner ».

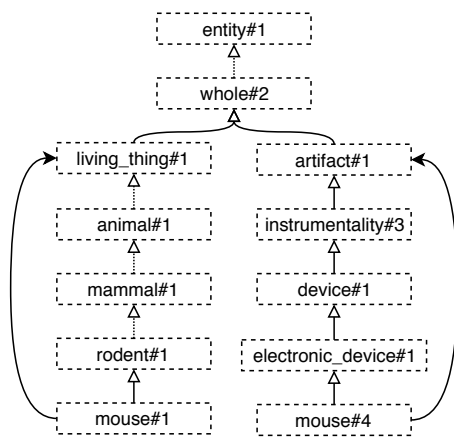


FIGURE 2 – Compression de vocabulaire utilisant la hiérarchie d'hyperonymie, appliquée au premier et quatrième sens du mot « mouse ». Les lignes en pointillés indiquent que des nœuds ont été omis par clarté.

Pour la DL, tout comme le regroupement des synonymes en *synsets* aide à mieux généraliser, nous faisons l'hypothèse que le regroupement des sens faisant partie d'une même hiérarchie d'hyperonymie va aussi aider à mieux généraliser, et que les concepts les plus spécialisés de WordNet sont souvent superflus. En effet, si l'on considère un sous-ensemble de WordNet qui ne comprend que le mot « souris », avec son premier sens (le petit rongeur), son quatrième sens (le dispositif électronique), et tous leurs hyperonymes, tel qu'illustré dans la figure 2, on voit que les concepts « artefact » et « être vivant » suffisent à différencier les deux sens, et toutes les étiquettes plus spécialisées pourraient être ramenés à ces deux concepts. Ainsi, non seulement le vocabulaire de sens, c'est à dire le nombre d'étiquettes de sens dans notre inventaire, sera réduit, mais en plus tous les autres « êtres vivants » donneront des exemples qui pourront ensuite permettre de différencier les deux sens de souris.

En considérant maintenant tout le vocabulaire de WordNet, l'objectif de notre méthode est ainsi de faire correspondre chaque sens à son ancêtre le plus haut dans sa hiérarchie d'hyperonymie, avec les contraintes suivantes : première-

ment, cet ancêtre doit permettre de discriminer tous les différents sens du mot cible ; deuxièmement, nous devons conserver les hyperonymes qui sont indispensables pour discriminer les sens des autres mots du dictionnaire. Par exemple, en prenant tout WordNet en considération, nous ne pouvons pas faire correspondre « souris#1 » à « être vivant#1 », parce qu'une étiquette plus spécifique, « animal#1 » est nécessaire pour distinguer les deux sens du mot « proie » (un sens décrit une personne et l'autre un animal). Notre méthode fonctionne donc en deux étapes :

1. Nous marquons comme « nécessaires » les enfants du premier ancêtre commun de chaque paire de sens de chaque mot de WordNet.
2. Nous faisons correspondre chaque sens à son premier ancêtre dans sa hiérarchie d'hyperonymie ayant été précédemment marqué comme « nécessaire ».

En conséquence, les sens les plus spécifiques de l'arbre qui ne sont pas indispensables pour distinguer un mot de l'inventaire lexical seront automatiquement supprimés du vocabulaire. En d'autres termes, l'ensemble de sens qui reste dans le vocabulaire est le plus petit sous-ensemble de tous les *synsets* qui sont nécessaires pour distinguer chaque sens de chaque mot de WordNet, en considérant seulement les liens d'hyperonymie, d'hyponymie et d'instance.

### 3.3 Compression de vocabulaire de sens grâce à l'ensemble des relations sémantiques de WordNet

En plus de l'hyperonymie, de l'hyponymie et de la relation d'instance, WordNet contient plusieurs autres relations entre *synsets*, telles que la méronymie (X fait partie de Y, ou X est un membre de Y) et son opposé l'holonymie, l'antonymie (X est le contraire de Y), etc.

Nous proposons ainsi une deuxième méthode de compression du vocabulaire de sens, qui prend en compte toutes les relations sémantiques offertes par WordNet, afin de former des groupes de *synsets* proches.

Par exemple, en utilisant toutes les relations sémantiques disponibles, nous pourrions former un groupe contenant « physicien », « physique » (domaine), « Einstein » (instance), « astronome » (hyponyme), mais aussi d'autres sens connexes tels que « photon », car c'est un méronyme de « rayonnement », qui est un hyponyme de « énergie », qui appartient au même domaine de « physique », etc.

Notre méthode fonctionne en construisant ces groupes de manière itérative. Soit  $S$  l'ensemble des *synsets* de WordNet et  $C$  l'ensemble des groupes de *synsets* que l'on cherche à construire, on initialise d'abord  $C$  comme des singletons contenant chacun un *synset* différent.

$$C = \{c_0, c_1, \dots, c_n\} \quad S = \{s_0, s_1, \dots, s_n\}$$

$$C = \{\{s_0\}, \{s_1\}, \dots, \{s_n\}\}$$

Ensuite, à chaque étape, on trie  $C$  par taille de groupes, et on sélectionne le plus petit groupe  $c_x$  ainsi que le plus petit groupe relié à  $c_x, c_y$ . On considère qu'un groupe  $c_a$  est relié à un groupe  $c_b$  si un *synset*  $s_a \in c_a$  est relié à un *synset*

$s_b \in c_b$  par n'importe quel lien sémantique. On fusionne  $c_x$  et  $c_y$  ensemble, si et seulement si l'opération permet toujours de discriminer les différents sens de tous les mots de la base lexicale. Si c'est le cas, on valide la fusion et on passe à l'étape suivante. Si ce n'est pas le cas, on annule la fusion et on essaye avec un autre groupe relié à  $c_x$ . S'il est impossible de fusionner un groupe avec  $c_x$ , alors on essaye avec le plus petit groupe suivant, et si aucune fusion n'est possible pour aucun des groupes, l'algorithme s'arrête.

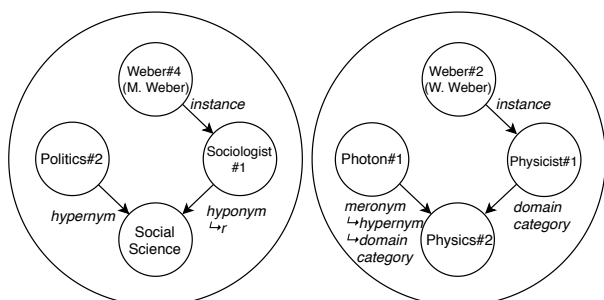


FIGURE 3 – Exemple de groupes de sens pouvant résulter de notre méthode, si on ne considère que deux sens du nom « Weber » et seulement certaines relations.

Dans la figure 3, nous montrons un ensemble possible de groupes qui pourraient résulter de notre méthode.

Cette méthode produit des groupes significativement plus grands que celle s'appuyant sur les hyperonymes. En effet, en moyenne, un groupe contient 5 *synsets* avec cette dernière, alors qu'il en contient 17 avec celle-ci. De plus, cette méthode, contrairement à la précédente, est également stochastique, parce qu'à chaque fois qu'on ordonne les groupes par taille, l'algorithme de tri place les groupes de même taille dans un ordre aléatoire. Cependant, comme nous réordonnons les groupes après chaque fusion, les groupes sont de taille assez équilibrés, et nous avons observé que la taille finale du vocabulaire (c.-à-d. le nombre de groupes) se situe toujours entre 11 000 et 13 000,

Dans la suite, on considère un ensemble  $C$  généré après que l'algorithme se soit arrêté après 105 775 étapes de fusion (générant ainsi 11 885 groupes de sens).

Méthode de compression	Taille du vocabulaire	Taux de compression	Couverture du SemCor
Référence	206 941	référence	16%
Synonymes	117 659	43%	22%
Hyperonymes	39 147	81%	32%
Toutes relations	11 885	94%	39%

TABLE 1 – Résultats de nos deux méthodes de compression de vocabulaire sur la taille du vocabulaire et la couverture du SemCor.

La table 1 montre l'effet de la compression de vocabulaire via les synonymes (sens vers *synsets*), de notre première nouvelle méthode utilisant les hyperonymes, ainsi que de notre deuxième nouvelle méthode utilisant toutes

les relations de WordNet, sur la taille du vocabulaire de sens de WordNet, et sur la couverture du SemCor. Comme nous pouvons le constater, la taille du vocabulaire diminue considérablement grâce à nos méthodes, et la couverture d'un même corpus est nettement améliorée.

## 4 Protocole expérimental

Afin d'évaluer nos méthodes de compression de vocabulaire de sens, nous les avons appliquées à un système neuronal de DL état de l'art similaire à celui de Vial et al. (2019) (voir la section 2.2). Notre réseau de neurones prend ainsi en entrée directement les mots sous forme vectorielle, à partir d'un modèle de vecteurs de mots pré-entraîné, il repose ensuite sur une ou plusieurs couches cachées, puis sur une couche de sortie, qui associe à chaque mot une distribution de probabilité sur tous les sens du vocabulaire utilisé, à l'aide de la fonction *softmax*.

### 4.1 Détails de l'architecture

En entrée de notre réseau, nous avons utilisé les vecteurs contextualisés BERT (Devlin et al., 2018). Nous avons utilisé le modèle pour l'anglais « bert-large-cased » qui est pré-entraîné sur BookCorpus et Wikipedia, et qui produit des vecteurs de dimension 1 024.

Pour les couches cachées, nous avons appliqué 6 couches d'encodeurs *Transformer* (Vaswani et al., 2017), avec les mêmes paramètres que le modèle « base » de l'article original (8 têtes d'attention, dimension cachée de 2 048, et régularisation *dropout* à 0,1). Les couches *Transformer* s'appuient sur le mécanisme d'auto-attention, et nous les avons utilisés à la place des cellules récurrentes plus classiques comme des *LSTM* ou des *GRU*, parce que plusieurs travaux récents ont montré leur plus grande efficacité dans une multitude de tâches, par exemple en traduction automatique (Vaswani et al., 2017; Ott et al., 2018) et en modélisation de la langue (Devlin et al., 2018).

De plus, étant donné que les vecteurs renvoyés par BERT encodent directement les positions des mots, il n'est pas nécessaire d'avoir une récurrence au niveau des couches cachées. Ainsi, nous n'ajoutons pas de vecteurs de positions supplémentaires en entrée de notre encodeur.

Pour tous les autres paramètres du modèle, comme la taille de *batch* ou la méthode d'optimisation, nous avons utilisé les mêmes paramètres que Vial et al. (2019).

### 4.2 Entraînement du modèle

Nous avons comparé nos méthodes sur deux ensembles de corpus d'entraînement : le SemCor (Miller et al., 1993), le plus grand corpus annoté en sens utilisé pour l'apprentissage de la plupart des systèmes supervisés de DL, et la concaténation du SemCor et du WordNet Gloss Tagged<sup>5</sup>. Ce dernier est un corpus distribué dans WordNet depuis sa version 3.0, et il contient les définitions de tous les sens de WordNet, annoté manuellement ou semi-automatiquement

5. <http://wordnetcode.princeton.edu/glosstag-files/glosstag.shtml>

Système	SE2	SE3	SE07	SE13	SE15	ALL (concat. tâches précédentes)					SE07 07
						noms	verbes	adj.	adv.	total	
Étalon du premier sens	65,6	66,0	54,5	63,8	67,1	67,7	49,8	73,1	80,5	65,5	78,9
UFSAC+1M (Vial et al., 2019)	74,6	69,4	60,7	69,8	74,2	-	-	-	-	†71,1	85,0
HCAN (Luo et al., 2018a)	72,8	70,3	-	68,5	72,8	72,7	58,2	77,4	84,1	71,1	-
LSTMLP (Yuan et al., 2016)	73,8	71,8	63,5	69,5	72,6	†73,9	-	-	-	†71,5	83,6
SemCor, référence	77,2	76,5	70,1	74,7	77,4	78,7	65,2	79,1	85,5	76,0	87,7
SemCor, hyperonymes	77,5	77,4	69,5	76,0	78,3	79,6	65,9	79,5	85,5	76,7	87,6
SemCor, toutes relations	76,6	76,9	69,0	73,8	75,4	77,2	66,0	80,1	85,0	75,4	86,7
SemCor+WNGT, référence	<b>79,7</b>	76,1	<b>74,1</b>	78,6	80,4	80,6	68,1	82,4	86,1	78,3	90,4
SemCor+WNGT, hyperonymes	<b>79,7</b>	77,8	73,4	<b>78,7</b>	<b>82,6</b>	81,4	68,7	83,7	85,5	<b>79,0</b>	90,4
SemCor+WNGT, toutes relations	79,4	<b>78,1</b>	71,4	77,8	81,4	80,7	68,6	82,8	85,5	78,5	<b>90,6</b>

TABLE 2 – Scores F1 (%) sur les tâches de DL de l’anglais des campagnes d’évaluation SensEval/SemEval. La tâche « ALL » est la concaténation de SE2, SE3, SE07 17, SE13 et SE15. La stratégie de repli est appliquée sur les mots dont aucun sens n’a été observé pendant l’entraînement. Les scores en **gras** sont à notre connaissance les meilleurs résultats obtenus sur la tâche. Les scores prefixés par un obèle (†) ne sont pas fournis par les auteurs mais sont déduits de leurs autres scores.

en sens. Nous avons utilisé les versions de ces corpus fournies avec la ressource UFSAC 2.1<sup>6</sup> (Vial et al., 2018).

Système	Sans repli	Repli sur monosémiques
SemCor, référence	93,23%	98,13%
SemCor, hyperonymes	98,75%	99,68%
SemCor, toutes relations	99,67%	99,99%
SemCor+WNGT, référence	98,26%	99,41%
SemCor+WNGT, hyperonymes	99,83%	99,96%
SemCor+WNGT, toutes relations	99,99%	100%

TABLE 3 – Couverture de nos systèmes sur la tâche ALL. « Repli sur monosémiques » est le cas où l’on considère que les mots monosémiques sont déjà annotés en sens.

Nous avons entraîné chaque modèle sur 20 passes de nos données d’entraînement. Au début de chaque passe, nous avons mélangé toutes les phrases aléatoirement, et à la fin de chaque passe, nous avons évalué notre modèle sur un jeu de développement, et nous avons conservé celui qui a obtenu le meilleur score F1 de DL. Le corpus de développement est constitué de 4 000 phrases prises aléatoirement du WNGT pour le système entraîné sur le SemCor seul, et de 4 000 phrases extraites aléatoirement de nos données d’entraînement pour les autres.

Nous avons ainsi entraînés trois systèmes :

1. un système « référence » dont le vocabulaire de sens est celui de tous les *synsets* vus pendant l’entraînement (utilisant ainsi la compression via les synonymes) ;
2. un système « hyperonymes » entraîné dans les mêmes conditions, mais avec notre première méthode de compression du vocabulaire via les hyperonymes et les hyponymes appliquée sur le corpus d’entraînement ;
3. un système « toutes relations » qui applique cette fois-ci sur le corpus d’entraînement notre deuxième méthode

6. <https://github.com/getalp/UFSAC>

de compression de vocabulaire via toutes les relations sémantiques de WordNet.

Système	Nombre de paramètres	
	SemCor	SemCor+WNGT
Référence	77,15M	120,85M
Hyperonymes	63,44M	79,85M
Toutes relations	55,16M	60,27M

TABLE 4 – Nombre de paramètres d’un modèle en fonction du corpus d’apprentissage et de notre méthode de compression de vocabulaire.

Tous les entraînements ont été effectués sur un seul GPU Titan X de Nvidia. Dans la table 4, nous montrons le nombre de paramètres des différents modèles, en fonction du corpus d’entraînement et de notre méthode de compression du vocabulaire. Comme nous pouvons le voir, ce nombre est réduit par un facteur de 1,2 à 2 grâce à nos méthodes de compression.

### 4.3 Résultats

Nous avons évalué nos modèles sur tous les corpus d’évaluation de la DL de l’anglais des campagnes d’évaluation SensEval/SemEval, c’est-à-dire les corpus d’évaluation « grain fin » de SensEval 2, SensEval 3, SemEval 2007 (tâche 17), SemEval 2013 et SemEval 2015, ainsi que le corpus « ALL » constitué de leur concaténation. Nous avons également comparé nos résultats sur la tâche « gros grain » de SemEval 2007 (tâche 7).

Pour chaque évaluation, nous avons entraîné 8 modèles indépendants, et nous donnons le score obtenu par un système « ensemble » qui moyenne leurs prédictions à l’aide d’une moyenne géométrique.

Les scores obtenus par nos systèmes en comparaison avec les meilleurs systèmes de l’état de l’art et l’étalon du premier sens sont présents dans le tableau 2, et le tableau 3 montre la couverture de nos systèmes.



Corpus d'entraînement	Vecteurs de mots pré-entraînés	Ensemble	Scores F1 sur la tâche "ALL" (%)					
			Référence		Hyperonymes		Toutes relations	
			$\bar{x}$	$\sigma$	$\bar{x}$	$\sigma$	$\bar{x}$	$\sigma$
SemCor+WNGT	BERT	Oui	78,27	-	79,00	-	78,48	-
SemCor+WNGT	BERT	Non	76,97	$\pm 0,38$	77,08	$\pm 0,17$	76,52	$\pm 0,36$
SemCor+WNGT	ELMo	Oui	75,16	-	74,65	-	70,58	-
SemCor+WNGT	ELMo	Non	74,56	$\pm 0,27$	74,36	$\pm 0,27$	68,77	$\pm 0,30$
SemCor+WNGT	GloVe	Oui	72,23	-	72,74	-	71,42	-
SemCor+WNGT	GloVe	Non	71,93	$\pm 0,35$	71,79	$\pm 0,29$	69,60	$\pm 0,32$
SemCor	BERT	Oui	76,02	-	76,73	-	75,40	-
SemCor	BERT	Non	75,06	$\pm 0,26$	75,59	$\pm 0,16$	73,91	$\pm 0,33$
SemCor	ELMo	Oui	72,55	-	73,09	-	69,43	-
SemCor	ELMo	Non	72,21	$\pm 0,13$	72,83	$\pm 0,24$	68,74	$\pm 0,29$
SemCor	GloVe	Oui	70,77	-	71,18	-	68,44	-
SemCor	GloVe	Non	70,51	$\pm 0,16$	70,77	$\pm 0,21$	67,48	$\pm 0,55$
Système « élève » Vial et al. (2019)								
SemCor+UFSAC+1M News 2016 HCAN Luo et al. (2018a)	GloVe	Oui	71,1					
SemCor+WordNet glosses	GloVe	Non	71,1					
LSTMLP Yuan et al. (2016)								
SemCor+1K (private)	private	Non	71,5					

TABLE 5 – Étude des hyperparamètres sur la tâche "ALL" (concaténation des corpus de toutes les tâches de désambiguïation lexicale à granularité fine de SensEval/SemEval). Pour les systèmes qui n'utilisent pas l'ensemble, nous montrons la moyenne des scores ( $\bar{x}$ ) de huit modèles entraînés séparément, avec l'écart type ( $\sigma$ ).

Concernant les résultats présentés dans la table 2, nous observons que nos systèmes qui utilisent nos méthodes de compression de vocabulaire, que ce soit grâce à la relation d'hyperonymie ou grâce à toutes les relations obtiennent des scores qui sont globalement équivalents ou légèrement supérieurs aux systèmes « références » qui n'utilisent pas nos méthodes.

Nos méthodes de compression améliorent cependant grandement la couverture de nos systèmes. En effet, comme nous pouvons le voir dans la table 3, sur un total de 7 253 mots à annoter pour le corpus « ALL », le système de référence entraîné sur le SemCor n'est pas capable d'annoter 491 d'entre eux, alors qu'avec la compression du vocabulaire à travers les hyperonymes, ce nombre descend à 91, et 24 avec la compression à travers toutes les relations.

Lors de l'ajout du WordNet Gloss Tagged aux données d'entraînement, seulement 12 mots ne peuvent pas être annotés avec le système « hyperonymes », et avec le système « toutes relations », plus qu'un seul mot (l'adjectif monosémique « cytotoxique ») ne peut pas être annoté parce que son sens n'a pas été vu pendant l'entraînement. Si nous prenons en compte uniquement les mots polysémiques, le système basé sur la compression à travers toutes les relations et entraîné sur le SemCor n'est pas capable d'annoter seulement un seul mot (l'adverbe « eloquently »). Si on rajoute le WNGT en plus, la couverture passe à 100%.

Par rapport aux autres travaux, nous obtenons des résultats surpassant significativement l'état de l'art dans toutes les tâches, notamment grâce à l'ajout du WordNet Gloss Tagged aux données d'entraînement, et des vecteurs BERT en entrée de notre système.

#### 4.4 Étude des hyperparamètres

Afin de mieux comprendre l'origine de nos scores, nous étudions l'impact de nos principaux paramètres sur les résultats. En plus du corpus d'entraînement et de la méthode de compression du vocabulaire, nous avons choisi deux paramètres qui nous différencient de l'état de l'art : le modèle de vecteurs de mots pré-entraînés, et la méthode d'ensemble, et nous les avons fait varier.

Pour le modèle de vecteurs de mots, nous avons expérimenté avec BERT (Devlin et al., 2018) comme pour nos résultats principaux, mais aussi avec ELMo (Peters et al., 2018) et GloVe (Pennington et al., 2014). Pour ELMo, nous avons utilisé le modèle entraîné sur Wikipedia et les données monolingues de WMT 2008-2012.<sup>7</sup> Pour GloVe, nous avons utilisé le même modèle que Luo et al. (2018a) et Vial et al. (2019) entraîné sur Wikipedia 2014 et Gigaword 5.<sup>8</sup> Comme les représentations vectorielles de GloVe n'encodent pas la position des mots (un mot a la même représentation quelque soit sa position ou son contexte), nous avons réutilisé une couche de cellules LSTM bidirectionnelles de taille 1 000 par direction pour les couches cachées (comme Vial et al. (2019)).

Pour la méthode d'ensemble, nous avons expérimenté soit en l'utilisant, comme dans nos résultats principaux, c'est-à-dire en moyennant les prédictions de 8 modèles entraînés séparément, ou bien en donnant la moyenne et l'écart type des scores des 8 modèles évalués individuellement.

Comme nous pouvons le voir dans la table 5, le corpus d'entraînement supplémentaire (WNGT) et encore plus

7. <https://allennlp.org/elmo>

8. <https://nlp.stanford.edu/projects/glove/>

l'utilisation de BERT en tant que vecteurs de mots ont tous les deux un impact majeur sur nos résultats et conduisent à des scores supérieurs à l'état de l'art. L'utilisation de BERT au lieu de ELMo ou GloVe améliore respectivement le score d'environ 3 et 5 points dans chaque expérience, et l'ajout du WNGT aux données de d'entraînement l'améliore encore d'environ 2 points. Enfin, l'utilisation d'ensembles ajoute environ 1 point au score F1 final.

Enfin, à travers les scores obtenus par les modèles individuels (sans ensemble), nous pouvons observer sur les écarts-types que la méthode de compression du vocabulaire par les hyperonymes n'a jamais d'impact significatif sur le score final. Cependant, la méthode de compression via toutes les relations semble avoir un impact négatif sur les résultats dans certains cas (en utilisant GloVe et ELMo particulièrement, et en utilisant le SemCor seul comme corpus d'entraînement).

## 5 Conclusion

Dans cet article, nous avons présenté deux nouvelles méthodes qui améliorent la couverture et la capacité de généralisation des systèmes de DL supervisés, en réduisant le nombre d'étiquettes de sens différentes dans WordNet afin de ne conserver que celles qui sont essentielles pour différencier les sens de tous les mots présents dans la base lexicale. À l'échelle de l'ensemble de la base de données lexicale, nous avons montré que ces méthodes permettaient de réduire le nombre total d'étiquettes de sens différentes dans WordNet à seulement 6% de sa taille originale, et que la couverture d'un même corpus d'entraînement est ensuite plus que doublée.

Nous avons entraîné un système de DL neuronal état de l'art et nous avons montré que nos méthodes permettaient de réduire la taille des modèles par un facteur de 1,2 à 2 et de largement augmenter leur couverture, sans dégrader leurs performances. Au final, nous obtenons une couverture de 99,99% sur l'ensemble des tâches d'évaluation (soit un seul mot manquant sur les 7 253) lorsque l'on entraîne notre système sur le SemCor uniquement, et 100% lorsque l'on ajoute le WordNet Gloss Tagged aux données d'entraînement. On élimine ainsi quasiment le besoin d'une méthode de repli pour désambiguïser n'importe quel mot du vocabulaire de WordNet.

Notre méthode combinée avec les récentes avancées en terme de vecteurs de mots pré-entraînés permet à notre système de surpasser nettement l'état de l'art dans toutes les tâches d'évaluation de la DL de l'anglais, avec une bien meilleure couverture.

## Références

J. Devlin, M.-W. Chang, K. Lee, et K. Toutanova. Bert : Pre-training of deep bidirectional transformers for language understanding, 2018.  
 I. Iacobacci, M. T. Pilehvar, et R. Navigli. Embeddings for word sense disambiguation : An evaluation study. Dans *Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1 : Long Papers)*, pages 897–907, Berlin, Germany, August 2016. Association for Computational Linguistics.

M. Kågebäck et H. Salomonsson. Word sense disambiguation using a bi-directional lstm. Dans *5th Workshop on Cognitive Aspects of the Lexicon (CogALex)*. Association for Computational Linguistics, 2016.  
 M. Le, M. Postma, J. Urbani, et P. Vossen. A deep dive into word sense disambiguation with lstm. Dans *Proceedings of the 27th International Conference on Computational Linguistics*, pages 354–365. Association for Computational Linguistics, 2018.  
 F. Luo, T. Liu, Z. He, Q. Xia, Z. Sui, et B. Chang. Leveraging gloss knowledge in neural word sense disambiguation by hierarchical co-attention. Dans *Proceedings of the 2018 Conference on Empirical Methods in Natural Language Processing*, pages 1402–1411. Association for Computational Linguistics, 2018a.  
 F. Luo, T. Liu, Q. Xia, B. Chang, et Z. Sui. Incorporating glosses into neural word sense disambiguation. Dans *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1 : Long Papers)*, pages 2473–2482. Association for Computational Linguistics, 2018b.  
 G. A. Miller, R. Beckwith, C. Fellbaum, D. Gross, et K. Miller. Wordnet : An on-line lexical database. *International Journal of Lexicography*, 3 : 235–244, 1990.  
 G. A. Miller, C. Leacock, R. Tengi, et R. T. Bunker. A semantic concordance. Dans *Proceedings of the workshop on Human Language Technology, HLT '93*, pages 303–308, Stroudsburg, PA, USA, 1993. Association for Computational Linguistics.  
 R. Navigli, K. C. Litkowski, et O. Hargraves. Semeval-2007 task 07 : Coarse-grained english all-words task. Dans *SemEval-2007*, pages 30–35, Prague, Czech Republic, June 2007.  
 M. Ott, S. Edunov, D. Grangier, et M. Auli. Scaling neural machine translation. Dans *Proceedings of the Third Conference on Machine Translation : Research Papers*, pages 1–9, Belgium, Brussels, Oct. 2018. Association for Computational Linguistics.  
 T. Pasini et R. Navigli. Train-o-matic : Large-scale supervised word sense disambiguation in multiple languages without manual training data. Dans *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 78–88. Association for Computational Linguistics, 2017.  
 J. Pennington, R. Socher, et C. D. Manning. Glove : Global vectors for word representation. Dans *Empirical Methods in Natural Language Processing (EMNLP)*, pages 1532–1543, 2014.  
 M. E. Peters, M. Neumann, M. Iyyer, M. Gardner, C. Clark, K. Lee, et L. Zettlemoyer. Deep contextualized word representations. Dans *Proc. of NAACL*, 2018.  
 A. Polguère. *Lexicologie et sémantique lexicale*. Les Presses de l'Université de Montréal, 2003.  
 A. Raganato, C. Delli Bovi, et R. Navigli. Neural sequence learning models for word sense disambiguation. Dans *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing*, pages 1167–1178. Association for Computational Linguistics, 2017.  
 K. Taghipour et H. T. Ng. One Million Sense-Tagged Instances for Word Sense Disambiguation and Induction. Dans *Proceedings of the Nineteenth Conference on Computational Natural Language Learning*, pages 338–344, Beijing, China, July 2015. Association for Computational Linguistics.  
 A. Vaswani, N. Shazeer, N. Parmar, J. Uszkoreit, L. Jones, A. N. Gomez, L. u. Kaiser, et I. Polosukhin. Attention is all you need. Dans I. Guyon, U. V. Luxburg, S. Bengio, H. Wallach, R. Fergus, S. Vishwanathan, et R. Garnett, editors, *Advances in Neural Information Processing Systems 30*, pages 5998–6008. Curran Associates, Inc., 2017.  
 L. Vial, B. Lecouteux, et D. Schwab. UFSAC : Unification of Sense Annotated Corpora and Tools. Dans *Language Resources and Evaluation Conference (LREC)*, Miyazaki, Japan, May 2018.  
 L. Vial, B. Lecouteux, et D. Schwab. Approche supervisée à base de cellules LSTM bidirectionnelles pour la désambiguïstation lexicale. *Traitement Automatique des Langues*, 2019.  
 D. Yuan, J. Richardson, R. Doherty, C. Evans, et E. Altendorf. Semi-supervised word sense disambiguation with neural models. Dans *COLING 2016*, 2016.  
 Z. Zhong et H. T. Ng. It makes sense : A wide-coverage word sense disambiguation system for free text. Dans *Proceedings of the ACL 2010 System Demonstrations, ACLDemos '10*, pages 78–83, Stroudsburg, PA, USA, 2010. Association for Computational Linguistics.

# Apprentissage en ligne par renforcement centré utilisateur pour la composition émergente d'applications ambiantes

Walid Younes<sup>1</sup>Françoise Adreit<sup>2</sup>Sylvie Trouilhet<sup>1</sup>Jean-Paul Arcangeli<sup>1</sup><sup>1</sup> Université de Toulouse, IRIT, Université Paul Sabatier<sup>2</sup> Université de Toulouse, IRIT, Université Toulouse Jean Jaurès

Prénom.Nom@irit.fr

## Résumé

*Les systèmes cyber-physiques connectés et ambiants entourent l'utilisateur humain de services plus ou moins complexes, qui doivent être le plus possible adaptés à ses préférences et à la situation courante. Nous proposons de les construire automatiquement et à la volée par composition de services plus élémentaires présents dans l'environnement, sans expression préalable des besoins de l'utilisateur. Pour cela, avec la forte variabilité dynamique de l'environnement ambiant et des besoins, ce dernier doit être sollicité mais à minima. Afin de produire les connaissances nécessaires à la composition automatique et en l'absence de données initiales, nous proposons une solution d'apprentissage en ligne par renforcement à horizon infini qui apprend incrémentalement de et pour l'utilisateur. Elle est décentralisée au sein d'un système multi-agent chargé de l'administration et de la composition des services.*

## Mots-clés

Apprentissage en ligne, apprentissage par renforcement, feedback utilisateur, services logiciels, composition de services, système multi-agent, intelligence ambiante

## Abstract

*Cyber-physical and ambient systems surround the human user with services which are more or less complex, and must be as tailored as possible to her/his preferences and the current situation. We propose to build them automatically and on the fly by composition of more elementary services present in the environment, without prior expression of the user's needs. In a context of high dynamic variability of both the ambient environment and the needs, the user must be involved at the minimum. In order to produce the knowledge necessary for automatic composition in the absence of an initial guideline, we have designed a generic solution based on online reinforcement learning. It is decentralized within a multi-agent system in charge of the administration and the composition of the services, which learns incrementally from and for the user.*

## Keywords

Online learning, reinforcement learning, user feedback, service discovery, selection and composition, multi-agent system, ambient intelligence

## 1 Introduction

Les systèmes ambiants et mobiles sont composés d'appareils fixes ou mobiles reliés par un ou plusieurs réseaux de communication. Ces appareils hébergent des composants logiciels qui fournissent des services et qui peuvent, eux-mêmes, requérir d'autres services. Ces composants sont des briques logicielles qui peuvent être assemblées en connectant des services requis à des services fournis pour composer des applications plus complexes [20]. Par exemple, l'assemblage d'un composant d'interaction non dédié présent dans un smartphone (*e.g.*, un  *curseur*, un  *bouton* ou un composant de reconnaissance vocale), d'un  *adaptateur* et d'une  *lampe connectée* peut permettre de réaliser une application permettant à un utilisateur de contrôler l'éclairage ambiant.

Les composants (matériels et logiciels) sont en général multi-propriétaires et gérés de manière indépendante : ils sont développés, installés et activés indépendamment les uns des autres. En raison de la mobilité des appareils et des utilisateurs, ils peuvent apparaître ou disparaître selon une dynamique imprévisible, conférant aux systèmes ambiants et mobiles un caractère ouvert et instable. À cela s'ajoute le nombre souvent important de composants, source de difficultés lors du passage à l'échelle. Dans ce contexte, les assemblages de composants sont difficiles à concevoir, à entretenir et à adapter.

Plongé au sein de ces systèmes, l'utilisateur humain peut utiliser les services qui sont à sa disposition. L'intelligence ambiante [24, 7, 19] vise à lui offrir un environnement personnalisé, adapté à la situation, c'est-à-dire à lui fournir les bons services au bon moment, en anticipant ses besoins qui eux-mêmes peuvent évoluer. Pour cela, l'utilisateur peut être sollicité mais dans une limite raisonnable [3].

Notre projet a pour objectif de concevoir et de réaliser un "moteur de composition" qui détecte périodiquement les

composants présents dans l’environnement ambiant et qui les assemble dynamiquement et automatiquement afin de construire des applications “composites” adaptées à l’environnement et à l’utilisateur, c’est-à-dire opérationnelles, utiles et utilisables [9]. En l’absence de besoin explicité *a priori*, le moteur apprend les préférences de l’utilisateur en fonction de la situation. Notre approche rompt avec le traditionnel mode *top-down* pour le développement d’applications : la réalisation d’un assemblage n’est pas guidée par les besoins explicites de l’utilisateur ni par des plans d’assemblage prédéfinis ; au contraire, les applications composites sont construites à la volée en mode *bottom-up* à partir des services présents sur le moment dans l’environnement ambiant. Ainsi, les applications émergent de l’environnement, en tirant profit des opportunités. Dans ce cadre, l’utilisateur ne demande pas un service ou une application : les applications émergentes lui sont fournies en mode *push*<sup>1</sup>. Pour cela, nous avons développé une solution basée sur l’apprentissage automatique. L’objectif de ce papier est de présenter les principes de cet apprentissage et de leur mise en œuvre. Il est organisé comme suit. L’architecture du système incluant le moteur de composition et l’utilisateur est présentée dans la section 2. Dans la section 3, notre problématique d’apprentissage est analysée en termes de motivations, d’objectifs et de données. Cette section se conclut par la définition du type d’apprentissage : en ligne et par renforcement. Les principes de notre solution d’apprentissage sont ensuite exposés dans la section 4. La section 5 résume l’état de l’art en matière d’apprentissage pour la composition logicielle automatique et positionne notre proposition. En conclusion, dans la section 6, le lecteur trouvera un bref résumé de la contribution, un point d’avancement sur le développement de la solution et une discussion sur les problèmes ouverts et la suite de ce travail.

## 2 Architecture du système de composition

Afin de répondre à l’exigence d’automatisation de la composition opportuniste de composants logiciels, nous avons défini une architecture logicielle du système de composition [25]. Cette section en présente les grandes lignes et la figure 1 en donne une vue simplifiée.

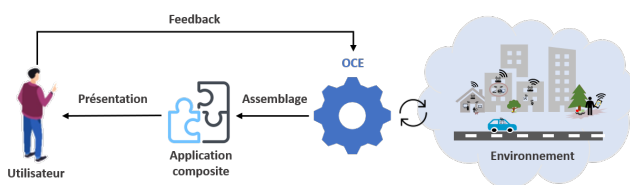


FIGURE 1 – Vue simplifiée de l’architecture

Au cœur du système, le moteur OCE (*Opportunistic Composition Engine*) a pour fonction de concevoir pour l’uti-

1. Un exemple de cas d’utilisation est développé dans [25].

lisateur des applications composites en assemblant des composants métiers et des composants d’interaction disponibles. Pour cela, il perçoit les composants et leurs services présents dans l’environnement ambiant, puis choisit d’établir des connexions entre des services requis et des services fournis faisant ainsi émerger les applications.

OCE propose à l’utilisateur les applications émergentes et c’est l’utilisateur qui décide *in fine* de leur déploiement : il peut accepter ou rejeter une application proposée ou encore la modifier [14]. Ce point est particulièrement important dans le domaine de l’interaction homme-machine pour lequel le contrôle par l’utilisateur de son environnement d’interaction est fondamental [3]. Après acceptation, l’application émergente est déployée automatiquement.

Afin de solliciter l’utilisateur *a minima*, le système de composition doit être le plus autonome possible. Notre solution se conforme aux principes de l’informatique autonome et au modèle MAPE-K (*Monitor, Analyse, Plan, Execute - Knowledge*) [12]. Cependant, alors que dans le modèle MAPE-K la phase d’exécution a pour fonction de réaliser ce qui a été planifié, l’application émergente est ici présentée à l’utilisateur sous la forme d’un assemblage de composants avant d’être éventuellement déployée. Les actions que l’utilisateur effectue en réponse sont sources de feedback et permettent à OCE d’apprendre et de faire évoluer ainsi ses connaissances (voir les sections suivantes).

Comme, fondamentalement, l’environnement ambiant est distribué et les composants matériels et logiciels sont administrés par différentes autorités, nous avons choisi de concevoir le moteur sous la forme d’un système multi-agent (SMA) dans lequel chaque service d’un composant est “administré” par un agent. De manière générale, les agents sont des entités autonomes qui coopèrent afin d’atteindre un but commun [11]. Ceux du moteur OCE ont pour rôle d’établir des connexions (ou des déconnexions) entre les services qu’ils administrent. Chaque agent a pour but de trouver un service partenaire avec qui se connecter. Pour cela, les agents échangent des messages dans le cadre du protocole d’interaction ARSA (*Advertise, Reply, Select, Agree*) [25] qui supporte la coopération entre agents en mode asynchrone dans un contexte dynamique et ouvert : ainsi, les agents continuent de coopérer même si un agent disparaît (en cas de désactivation du service administré ou de panne) ou si un message n’arrive pas à son destinataire.

## 3 Apprendre : pourquoi, quoi, à partir de quoi et comment ?

La fonction d’OCE est de concevoir des assemblages de composants logiciels qui réalisent des applications émergentes pertinentes, puis de les proposer à l’utilisateur et enfin de les déployer suivant les retours de l’utilisateur. Pour cela, le moteur doit prendre des décisions de connexion entre services dans l’objectif de satisfaire l’utilisateur.

Quatre éléments principaux contraignent ces prises de décision : la dynamique et l’imprévisibilité de l’environne-

ment ambiant, le nombre de composants et de services potentiellement important, la dynamique et la variabilité des besoins de l'utilisateur et enfin, la nécessité de solliciter *a minima* l'utilisateur.

Pour prendre des décisions pertinentes, le moteur OCE a besoin de connaissances qui, dans ce contexte, doivent être apprises automatiquement. Dans cette section, nous analysons pourquoi OCE doit apprendre, ce qu'il doit apprendre et quelles sont les données d'apprentissage.

### 3.1 Pourquoi apprendre ?

De par la dynamique et l'imprévisibilité de l'environnement ambiant, la combinatoire générée par le nombre important de composants ainsi que la dynamique et la variabilité de ses propres besoins, l'utilisateur n'est pas en mesure d'exprimer *a priori*, explicitement et d'une manière exhaustive ses besoins et ses préférences ni de les traduire en plans d'assemblages dans les différentes situations qu'il peut rencontrer. Par conséquent, OCE ne peut se baser ni sur des besoins de l'utilisateur explicités *a priori* ni sur des règles d'assemblage prédéfinies. OCE doit donc apprendre et ce, à partir de l'expérience.

### 3.2 Apprendre quoi ?

OCE doit construire les connaissances nécessaires pour proposer des applications pertinentes à l'utilisateur. Il doit apprendre ce que celui-ci préfère quand certains composants sont présents dans l'environnement ambiant. Basées sur ces préférences à un moment donné, ces connaissances seront exploitées pour prendre de futures décisions dans des situations identiques ou similaires. Par exemple, OCE peut apprendre que l'utilisateur préfère contrôler l'éclairage ambiant au moyen du curseur embarqué sur son smartphone plutôt qu'avec l'interrupteur mural connecté.

### 3.3 À partir de quoi apprendre ?

Si le moteur ne peut pas disposer initialement d'un ensemble de données d'apprentissage, en revanche, des données peuvent être observées au fil de l'utilisation du système pour servir de base aux décisions futures.

Plusieurs sources de données sont possibles. La première provient de l'observation et de l'exploitation des retours de l'utilisateur sur l'application émergente proposée (acceptation, modification ou rejet). Par exemple, en cas de modification par l'utilisateur du contrôle de l'éclairage ambiant, OCE peut modifier ses préférences de connexion entre services (le curseur du smartphone plutôt que l'interrupteur). On peut récupérer d'autres feedbacks à partir de l'éditeur utilisé par l'utilisateur pour manipuler les composants et leurs connexions [14], notamment en observant certaines actions de l'utilisateur (par exemple, pousser un composant en dehors de la fenêtre d'édition). En interne au SMA, il est également envisageable d'extraire des informations de feedback à partir des interactions des agents (par exemple, en cas d'échanges soutenus entre deux agents, on pourrait favoriser la connexion entre leurs services respectifs).

À ce stade de notre travail, nous avons choisi d'exploiter la première source de données qui traduit les préférences de l'utilisateur en fonction des composants présents dans l'environnement ambiant. En effet, nous faisons l'hypothèse que même si l'utilisateur ne peut pas expliciter *a priori* ses besoins, il est capable de réagir sur le moment à la proposition d'une application construite automatiquement. Il est alors possible de capturer cette réaction sous la forme de feedback et d'en extraire de la connaissance utile à des prises de décision futures. Ainsi, le moteur apprend à partir des interactions avec l'utilisateur sans le surcharger.

### 3.4 Comment apprendre ?

L'absence de données initiales et de solutions connues rend impossible un apprentissage supervisé ou non supervisé. De plus, la dynamique de l'environnement, avec les services qui apparaissent et disparaissent de manière imprévisible, rend très difficile voire impossible la construction d'un modèle statique de prédiction ou de classification. Pour ces raisons, nous proposons un apprentissage par adaptation progressive permettant au moteur OCE d'apprendre en continu en exploitant les informations de feedback que l'utilisateur fournit itérativement. C'est un apprentissage *en ligne, par renforcement et à horizon infini*. Ce type de solution permet à un système de s'adapter sur le long terme en interagissant avec son environnement [4].

L'objet de l'apprentissage est ici de contribuer à déterminer une action. Les agents d'OCE raisonnent en s'appuyant sur des connaissances construites et mises à jour en ligne, de manière incrémentale, au fur et à mesure de l'expérience et au gré des interactions avec l'utilisateur et des éventuelles évolutions de ses préférences. Selon le modèle de l'apprentissage en ligne [8], les agents font une "prédiction" (l'assemblage) et l'environnement (ici l'utilisateur) apporte une réponse. Cependant, le retour donné ici par l'utilisateur n'a pas le caractère d'exactitude de la réponse de l'environnement du modèle classique de l'apprentissage en ligne. Pour cette raison, nous hybridons les principes de l'apprentissage en ligne avec ceux de l'apprentissage par renforcement [8] : la réponse de l'utilisateur permet de renforcer les connaissances des agents, et les décisions à l'itération *t* s'appuient sur les connaissances cumulées lors des itérations précédentes. Enfin, considérant la dynamique, à la fois de l'environnement ambiant et de l'utilisateur, cet apprentissage doit être en plus à horizon infini [8], ce qui n'exclut pas des phases de stabilisation des connaissances.

Notons en complément que cette approche n'exclut pas non plus la possibilité d'exploiter des connaissances connues *a priori* (par exemple, des règles générales d'assemblage de composants métiers, des règles d'ergonomie d'assemblage de composants d'interaction) qui pourraient être fournies initialement et ainsi accélérer le processus d'acquisition des connaissances.

## 4 Principes de la solution

OCE est un système multi-agent dans lequel chaque agent administre un service d'un composant. Dans le cadre de l'architecture présentée en section 2, OCE opère dans un cycle, appelé *cycle moteur*. Les agents eux-mêmes fonctionnent de manière cyclique : dans un cycle classique, appelé *cycle agent*, un agent perçoit, puis décide et enfin agit. Plusieurs cycles agent s'effectuent ainsi dans un cycle moteur : les agents perçoivent les messages reçus, décident et effectuent les actions conformément au protocole ARSA. En fin de cycle moteur, après présentation d'un assemblage à l'utilisateur, OCE récupère le feedback de l'utilisateur pour apprendre. Il peut ensuite sonder une nouvelle fois l'environnement ambiant et démarrer un nouveau cycle avec les connaissances des agents mises à jour. Dans un premier temps, nous présentons ce que fait un agent dans un cycle agent, en particulier comment il exploite ses connaissances. Puis, nous expliquons comment chaque agent construit et fait évoluer ses connaissances par apprentissage, et nous discutons ce mode d'apprentissage.

### 4.1 Cycle de vie d'un agent : de la perception à l'action

Pour décider de l'action à effectuer, un agent construit une représentation de la situation courante dans la phase de perception. Il compare ensuite cette situation à des situations de référence qu'il a déjà rencontrées pour pouvoir l'évaluer et choisir l'agent à qui il répond (et donc l'action) dans la phase de décision. Enfin, il effectue l'action.

**Construction de la situation courante.** On appelle *situation courante*  $S_t^i$  pour un agent  $A^i$  la situation dans laquelle se trouve  $A^i$  dans le cycle moteur courant : elle est composée de l'ensemble des agents services perçus par  $A^i$  dans l'environnement et compatibles avec lui du point de vue de la composition<sup>2</sup>. Cette situation est construite par  $A^i$ , incrémentalement dans un cycle moteur, à partir des messages qu'il reçoit. C'est à l'étape de perception qu'elle est créée puis actualisée conformément à l'algorithme 1.

La situation courante  $S_t^i$  est formée d'un ensemble des couples de la forme  $(A^j, Type\_Message)$  où :

- $A^j$  est l'identifiant de l'agent émetteur du message,
- $Type\_Message$  est le type de message envoyé dans le cadre du protocole ARSA c'est-à-dire *Advertise*, *Reply*, *Select* ou *Agree*.

La situation courante est ensuite rapprochée des situations de référence qu'a déjà rencontrées l'agent.

**Rapprochement avec les situations de référence.** On appelle *situation de référence* pour un agent  $A^i$  une situation identifiée lors d'un précédent cycle moteur. Un agent dispose ainsi d'un ensemble de situations de référence qui constituent sa mémoire, noté  $Ref^i$ . Nous verrons plus loin comment cette connaissance est construite et maintenue par apprentissage. A l'instar d'une situation

2. Deux services sont dits compatibles si l'un est fourni  $S_F$  et l'autre est requis  $S_R$ , et si  $S_F$  rend le service  $S_R$  ( $S_R \subset S_F$ ).

---

#### Algorithme 1 Perception d'un agent $A^i$

---

- 1: Récupérer le lot de messages reçus
  - 2: **pour** chaque message reçu **faire**
  - 3:   Extraire le couple  $(A^j, Type\_Message)$
  - 4:   **si** service( $A^i$ ) et service( $A^j$ ) sont compatibles **alors**
  - 5:     Actualiser la situation courante :  $S_t^i \leftarrow S_t^i \cup \{(A^j, Type\_Message)\}$
  - 6:   **fin si**
  - 7: **fin pour**
- 

courante, une situation de référence est composée d'un ensemble d'agents services perçu par l'agent dans l'environnement à un moment donné, agents compatibles avec lui du point de vue de la composition

Une situation de référence de  $A^i$  est formée d'un ensemble de couples de la forme  $(A^j, Score_j^i)$ , où :

- $A^j$  est l'identifiant de l'agent émetteur du message,
- $Score_j^i$  est une valeur numérique qui représente pour  $A^i$  l'intérêt d'une connexion avec le service administré par  $A^j$ .

Le rapprochement a pour objectif d'identifier la situation courante parmi les situations de référence ou, à défaut, de sélectionner les situations de référence "similaires" à la situation courante. Il s'effectue sur la base des identifiants des agents présents dans les situations, en faisant abstraction des types de messages et des scores. Il est réalisé par la fonction *Calculer\_Similarité* qui construit un sous-ensemble avec les situations de référence dont le degré de similarité avec  $S_t^i$  est supérieur à un seuil  $\xi$ <sup>3</sup>. Le sous-ensemble est réduit à  $S_t^i$  si cette situation est connue (elle fait déjà partie des situations de référence); il est vide si aucune situation similaire n'a été trouvée.

Étant donné l'ensemble  $Ref^i$  des situations de référence de l'agent  $A^i$  et un seuil  $\xi$ , la fonction *Calculer\_Similarité* est définie de l'ensemble des situations courantes pour  $A^i$ ,  $Sit^i$ , vers un ensemble de couples composés d'une situation de référence et d'un degré de similarité (1) : à la situation courante rencontrée à l'instant  $t$ ,  $S_t^i$ , est associé l'ensemble des couples formés par une situation de référence  $SR_k^i$  et son degré de similarité  $d_k$  avec  $S_t^i$ , dans le cas où  $d_k$  est supérieur ou égal au seuil  $\xi$ .

$$Calculer\_Similarité : Sit^i \rightarrow \mathcal{P}(Ref^i \times \mathbb{R}) \quad (1)$$

$$S_t^i \mapsto \{(SR_k^i, d_k)\}_{k \leq |Ref^i| \text{ et } d_k \geq \xi}$$

La fonction *Calculer\_Similarité* peut implanter un algorithme de *clustering* pour regrouper les situations de référence en classes. Ceci permet d'exécuter l'opération de rapprochement plus rapidement<sup>4</sup> : il suffira à l'agent de comparer la situation courante aux représentantes dans chacune des classes.

3. Pour donner une idée, le degré de similarité entre deux situations peut être calculé sur la base la proportion d'agents en commun.

4. En particulier dans le cas des environnements ambiants où le nombre de situations de référence peut être important.

**Marquage de la situation courante.** La phase de marquage permet d'enrichir la situation courante à l'aide des situations de référence qui ont été sélectionnées dans la phase précédente, en attribuant un score aux agents de la situation courante. A l'issue de ce marquage, l'agent  $A^i$  pourra établir un classement préférentiel des agents  $A^j$  de la situation courante et en sélectionner un auquel répondre. Le marquage est réalisé par la fonction *Marquer\_Situation* qui attribue les scores à partir des situations retournées par la fonction *Calculer\_Similarité*. Si la situation courante  $S_t^i$  est connue, la fonction *Calculer\_Similarité* a renvoyé la situation de référence correspondante et les valeurs des scores sont reproduites à l'identique par la fonction *Marquer\_Situation*. Sinon, la fonction *Marquer\_Situation* calcule le score  $Score_j^i$  de chaque agent  $A^j$  de  $S_t^i$ <sup>5</sup>. Si l'un des agents de la situation courante n'apparaît pas dans les situations de référence (c'est le cas lors de l'apparition d'un service nouveau dans l'environnement ambiant), un score arbitraire lui est attribué. Le choix de la valeur de ce score permet de favoriser la prise en compte de la nouveauté<sup>6</sup>. Ce choix appartient à l'agent  $A^i$ . Il gère ce choix et le fait varier en fonction de ce qu'il a appris sur la sensibilité à la nouveauté de l'utilisateur. Dans le cas où aucune situation de référence similaire n'a été trouvée, les scores sont initialisés avec une valeur arbitraire. Comme précédemment, le choix de cette valeur initiale appartient à l'agent  $A^i$ .

**Choix de l'agent.** Dans cette phase, l'agent  $A^i$  sélectionne un agent  $A^j$  de la situation courante ( $A^i$  répondra alors au message de  $A^j$ ). Cette opération est effectuée par la fonction *Choisir\_Agent* qui prend en paramètre la situation courante marquée et renvoie l'agent qui maximise un critère d'optimisation. Pour cela, plusieurs stratégies sont possibles basées sur les scores ou le type de message (dans l'ordre *Agree*, *Select*, *Reply*, *Advertise*) ou une combinaison de ces deux critères.

L'algorithme 2 synthétise le comportement de l'agent lors de l'étape de décision.

---

#### Algorithme 2 Décision d'un agent $A^i$

---

- 1:  $Sit\_Similaires_t^i \leftarrow$   
**Calculer\_Similarité**( $S_t^i$ )
  - 2:  $Situation\_Marquée_t^i \leftarrow$   
**Marquer\_Situation**( $S_t^i, Sit\_Similaires_t^i$ )
  - 3:  $A^j \leftarrow$  **Choisir\_Agent**( $Situation\_Marquée_t^i$ )
- 

**Réalisation de l'action.** Dans cette étape, l'agent  $A^i$  donne suite au message de l'agent  $A^j$  choisi à l'étape précédente. Le type de message à envoyer suit, dans le cadre du protocole ARSA, celui de  $A^j$ . Ainsi, un agent continue à coopérer avec les autres agents jusqu'au traitement d'un message de type *Agree*. Dans ce cas, il accepte la

5. Ce calcul peut être la moyenne des scores de  $A^j$  dans les situations de référence sélectionnées, pondérée par les degrés de similarité.

6. Par exemple, en choisissant une valeur supérieure aux scores des autres agents de la situation.

connexion avec le service géré par  $A^j$  et se met en attente d'un feedback sur sa décision. L'algorithme 3 décrit le comportement de  $A^i$  dans la phase d'action.

---

#### Algorithme 3 Comportement d'un agent service

---

- 1: **selon** (*Type\_Message*)
  - 2: **cas** *Advertise* :
  - 3: envoyer un message *Reply* à  $A^j$  ;
  - 4: **cas** *Reply* :
  - 5: envoyer un message *Select* à  $A^j$  ;
  - 6: **cas** *Select* :
  - 7: envoyer un message *Agree* à  $A^j$  ;
  - 8: **cas** *Agree* :
  - 9: réaliser la connexion avec  $A^j$  ;
  - 10: se mettre en attente d'un feedback ;
  - 11: **fin selon**
- 

## 4.2 Apprentissage par feedback utilisateur

Au cours des différents cycles agent (perception-décision-action), l'agent exploite ses connaissances. L'apprentissage de ces dernières s'effectue hors cycle agent, lorsque le cycle moteur arrive à son terme et que le feedback de l'utilisateur est renvoyé à OCE. Ce feedback porte sur l'ensemble de l'assemblage proposé. Il est répercuté sur les agents qui composent l'assemblage. Chacun de ces agents  $A^i$  construit alors une nouvelle situation de référence à partir de sa situation courante marquée. Pour cela,  $A^i$  calcule une valeur de renforcement du score de chaque agent  $A^k$  de la situation courante, notée  $r_{A^k}^i$ . Le calcul des  $r_{A^k}^i$  utilise une variable  $\beta > 0$ , dont le choix appartient à  $A^i$ .

Trois cas sont possibles :

1. *L'utilisateur a accepté l'assemblage dans son intégralité.* La décision prise par tout agent  $A^i$  de l'assemblage de choisir son partenaire  $A^j$  est bonifiée :

$$\begin{aligned} r_{A^j}^i &= \beta \\ r_{A^k}^i &= 0, \forall A^k \in S_t^i \text{ t.q. } k \neq j. \end{aligned}$$

2. *L'utilisateur a rejeté l'assemblage dans son intégralité.* La décision prise par tout agent  $A^i$  de l'assemblage de choisir  $A^j$  est pénalisée :

$$\begin{aligned} r_{A^j}^i &= -\beta \\ r_{A^k}^i &= 0, \forall A^k \in S_t^i \text{ t.q. } k \neq j. \end{aligned}$$

3. *L'utilisateur a modifié l'assemblage.* Si, au moyen de l'éditeur à sa disposition (cf. section 3.3), il a déconnecté le service géré par un agent  $A^i$  du service géré par un agent  $A^j$  et reconnecté le premier au service géré par un agent  $A^h$ , on bonifie cette connexion et on pénalise celle proposée par OCE :

$$\begin{aligned} r_{A^h}^i &= (Score_j^i - Score_h^i) + \beta \\ r_{A^j}^i &= -\beta \\ r_{A^k}^i &= 0, \forall A^k \in S_t^i \text{ t.q. } k \neq h \text{ et } k \neq j. \end{aligned}$$

Les autres connexions de l'assemblage proposé sont traitées comme une acceptation (cas 1) si elles ne sont pas modifiées ou comme un rejet (cas 2) si elle sont supprimées.

Pour calculer le score  $Score_k^i$  des agents  $A^k$  figurant dans la situation courante  $S_t^i$ , l'agent  $A^i$  utilise la formule (2), inspirée de celle des algorithmes de bandits, dans laquelle  $Score_k^i$  est la valeur du score de l'agent  $A^k$  et  $\alpha \in [0, 1]$  est le facteur d'apprentissage :

$$Score_k^i = Score_k^i + \alpha(r_{A^k}^i - Score_k^i) \quad (2)$$

Dans le contexte dynamique et d'imprévisibilité de notre problème, il est difficile pour un agent de se projeter sur les prochaines situations de connexion et donc de déterminer l'importance des récompenses futures. Aussi, à ce stade de notre travail, nous faisons l'hypothèse que l'action choisie n'a pas d'effet sur les récompenses futures. Dans le cas contraire, il faudrait revenir à la formule plus complète du *Q-Learning* avec un facteur d'actualisation non nul.

Dans la formule,  $(1 - \alpha)Score_k^i$  représente la part d'information que l'agent  $A^i$  garde de son expérience passée et  $\alpha r_{A^k}^i$  celle qu'il apprend dans le cycle courant. On peut noter que le score des agents non retenus dans l'assemblage diminue systématiquement. Pour l'agent sélectionné  $A^j$ , en fonction de la valeur de  $\beta$ , le score peut aussi diminuer mais dans une moindre proportion ; la position de  $A^j$  est donc renforcée par rapport aux agents non retenus.

Une fois la situation de référence construite, l'agent  $A^i$  la stocke dans sa base de connaissances,  $Ref^i$ . Dans le cas où cette situation est déjà dans  $Ref^i$  (cas où la situation courante correspondait à une situation déjà connue),  $A^i$  se contente de mettre à jour les scores.

### 4.3 Discussion

Le but de l'apprentissage d'OCE est de maximiser la satisfaction de l'utilisateur. Notre solution ne demande pas la définition explicite des critères de satisfaction : OCE se base sur le feedback de l'utilisateur (qui accepte, refuse ou modifie un assemblage) et non sur l'évaluation de critères de qualité prédéfinis. Ceci confère à OCE un caractère générique (quel que soit l'utilisateur) et évolutif (l'utilisateur peut évoluer et ses critères de satisfaction aussi).

L'apprentissage ne porte pas sur l'algorithme de décision d'un agent (hors la part d'exploration inhérente à l'apprentissage par renforcement, un agent cherche toujours à se connecter avec le "meilleur" agent). Il porte sur la construction et l'adaptation des connaissances qui amènent l'agent à améliorer ses propositions. Ainsi, l'assemblage proposé par OCE pour une situation déjà rencontrée pourra différer de celui qu'il avait proposé auparavant parce qu'entre temps il a appris certaines préférences de l'utilisateur.

Dans le processus de composition d'un assemblage, les données remontent des agents et l'assemblage émerge des propositions locales de connexions. Ce comportement est caractéristique des systèmes multi-agents, dans lesquels la fonction du système n'est explicitement définie dans aucun des agents mais émerge de leurs interactions. De la même façon, l'apprentissage est distribué entre les agents. C'est l'agent qui apprend localement en révisant ses connaissances sur ses connexions locales, en ajoutant ou en modi-

fiant des situations de référence voire en oubliant certaines situations (par exemple celles qui sont trop anciennes). Ces situations représentent la vision locale de l'agent sur les organisations possibles au sein du système multi-agent.

L'apprentissage est donc ici un apprentissage concurrent dans lequel chaque agent est un apprenant autonome influencé par l'environnement [4]. On peut toutefois se demander si la vision purement locale est suffisante. Les agents ne pourraient-ils pas coopérer davantage et échanger, par exemple, des situations de référence pour une cohérence plus forte de leurs décisions ? Dans le même ordre d'idée, les agents qui gèrent les services d'un même composant "hôte" gagneraient probablement à se coordonner (par exemple, il peut être inutile d'annoncer un service fourni par un composant tant que les services requis par le même composant hôte ne sont pas satisfaits). Ainsi, avec des agents qui apprendraient sur d'autres agents, on s'orienterait vers un apprentissage multi-agent [2]. Il faut cependant noter la contribution de l'utilisateur à la cohérence de la décision globale : il évalue et contrôle les décisions d'OCE et son feedback est distribué aux agents. Transformé en connaissance, ce feedback global cadre les décisions des agents et donne une cohérence globale à l'agrégation des décisions individuelles.

## 5 Travaux connexes

Fondamentalement, les composants logiciels et les services sont des unités logicielles développées et déployées pour être réutilisées et composées. L'automatisation de la composition est une problématique largement traitée dans la littérature, en particulier en ce qui concerne les services Web. Pour F. Morh [16], le problème de la composition automatique des services se divise en deux grandes classes selon que la "structure" de la composition est connue au préalable ou non. Dans le premier cas, il s'agit de trouver à l'exécution les différents services qui vont permettre de réaliser un modèle donné (plan de composition, *workflow* de services...) en l'adaptant au mieux à la situation. Par exemple, MUSIC [18] permet une adaptation contextuelle de la composition en se basant sur un modèle : les plans sont sélectionnés au moment de l'exécution afin de maximiser un critère de qualité. Dans le deuxième cas, on crée de nouveaux services qui satisfont des pré-conditions et des post-conditions, ou des besoins explicités en amont. Dans tous les cas, la composition s'effectue en mode *top-down* (à l'inverse de notre approche *bottom-up*), à partir de modèles prédéfinis ou de buts explicites.

Dans [21], les auteurs présentent un état de l'art sur la composition de services en intelligence ambiante. Les solutions présentées reposent sur la formulation (sous différentes formes) d'un but à atteindre. Dans certains cas, l'utilisateur peut être impliqué et choisir une solution de composition parmi différentes compositions possibles qui lui sont présentées, comme c'est le cas dans notre architecture. Il n'est cependant pas présenté de système de composition qui repose sur l'apprentissage et, par conséquent, de contri-



bution de l'utilisateur à un processus d'apprentissage.

Dans [17], les auteurs proposent une solution à base d'apprentissage pour l'adaptation en ligne et en continu de systèmes logiciels à composants. À partir d'un but explicite et d'un ensemble de configurations connues satisfaisant ce but, il s'agit de trouver la meilleure selon des critères extrafonctionnels et non de faire émerger une fonctionnalité comme nous le proposons. L'adaptation n'est pas programmée mais apprise par renforcement à partir d'expérimentations sur les différentes configurations possibles. L'utilisateur est sollicité pour expérimenter mais pas pour donner un feedback explicite. Pour cela, les applications sont instrumentées et c'est l'environnement d'exécution qui génère les données de feedback.

Les travaux basés sur l'apprentissage pour automatiser la composition prennent le plus souvent la qualité de service (*QoS*) comme critère (par ex., [13]). Dans ce cadre, Wang *et al.* proposent une composition auto-adaptative de services Web en environnement dynamique afin de maximiser la *QoS* globale de la composition offerte à l'utilisateur [23]. Pour cela, en s'inspirant de [6, 10], ils modélisent la structure d'une composition de service sous la forme d'un processus de décision markovien contenant plusieurs *workflows*, la politique optimale pour le choix du meilleur *workflow* étant basée sur un algorithme de *Q-Learning*. La *QoS* n'est actuellement pas prise en compte dans notre solution ; ce pourrait être un élément de nature à enrichir la décision, obtenu par observation ou *via* le feedback de l'utilisateur.

Pour traiter l'évolutivité, le passage à l'échelle et l'optimisation dynamique de la composition, Wang *et al.* [22] étendent le travail exposé dans [23] et proposent un *framework* multi-agent collaboratif où les agents apprennent par renforcement en utilisant l'algorithme *Q-Learning*. Ici, le partage d'expérience entre les agents améliore l'efficacité et la vitesse d'apprentissage. Dans [15], les auteurs proposent une approche collaborative basée sur un algorithme d'apprentissage par renforcement appelé "Learning automata", pour adapter la composition de services Web et maintenir une *QoS* satisfaisante de la composition. Y. Charif et N. Sabouret utilisent eux aussi un protocole de coordination entre agents pour la chorégraphie dynamique de services : un agent dialogue à l'aide de requêtes et utilise un historique des conversations en guise de mémoire [5]. Ces approches coopératives sont intéressantes et rejoignent nos perspectives discutées en section 4.3.

## 6 Conclusion

Pour construire des applications ambiantes par assemblage automatique et dynamique de services logiciels, notre approche se démarque des solutions existantes en faisant émerger les applications de l'environnement ambiant en mode *bottom-up*, sans expression préalable des besoins ni modèle d'assemblage prédéfini. Afin de satisfaire au mieux l'utilisateur tout en limitant sa contribution dans un contexte de forte dynamique et d'imprévisibilité, cette émergence doit être contrôlée. Pour cela, notre "moteur de

composition" apprend en ligne, par renforcement et à horizon infini. C'est un système multi-agent dans lequel chaque agent gère un service. Les agents interagissent et coopèrent dans le cadre du protocole ARSA. Ils apprennent individuellement et en concurrence : chacun construit, à partir des messages qu'il reçoit, sa représentation locale de l'état du monde qui l'entoure. Il identifie ainsi la situation courante et la rapproche de situations rencontrées par le passé, ces dernières ayant été évaluées à partir de données de feedback de l'utilisateur. Le rapprochement entre situations permet à l'agent de prendre des décisions de connexion pertinentes pour le service qu'il administre. Cette méthode se rapproche du raisonnement par cas [1] basé sur la réutilisation de cas antérieurs et de solutions à un problème semblable au problème courant, apprises précédemment. En outre, le protocole ARSA et le rapprochement de situations permettent d'intégrer des services nouveaux et inconnus qui apparaissent soudainement dans l'environnement.

L'architecture fonctionnelle de notre système et le protocole ARSA ont été implémentés. Un prototype d'éditeur pour l'interaction avec l'utilisateur a été développé par ailleurs [14]. Avant de mettre en œuvre la solution d'apprentissage, il convient de finaliser la définition des différentes fonctions (calcul de similarité, marquage de situation, choix de l'agent) et les paramètres d'apprentissage. Nous pourrions alors conduire différentes expérimentations, portant sur différents cas d'utilisation identifiés, afin d'évaluer et de calibrer la solution.

Quelques questions restent ouvertes. D'une part, le temps d'apprentissage nécessaire pour que le moteur propose à l'utilisateur des applications utiles et utilisables pourrait s'avérer important. Une solution brièvement exposée en fin de section 3.4 pourrait consister à incorporer des règles métier ou des règles d'ergonomie au niveau des agents, sous la forme de situations de référence prédéfinies, afin d'accélérer l'acquisition des connaissances. Par ailleurs, notre solution d'apprentissage se base sur un feedback utilisateur capturé en phase de présentation de l'application. D'autres sources de données d'apprentissage sont possibles (cf. section 3.3). Dans tous les cas, un équilibre doit être respecté entre, d'un côté, la qualité et la quantité du feedback et de l'autre, la nature et la fréquence de la sollicitation de l'utilisateur. La précision de la valuation des situations à partir du feedback et le rapprochement de situations sont d'autres points critiques pour la qualité de la décision.

D'autre part, la formule (2), inspirée de celle du *Q-Learning*, ne prend pas en compte le facteur d'actualisation  $\gamma$  (cf. section 4.2). Cette question doit être étudiée plus précisément sur la base de cas concrets afin de déterminer en quoi le choix d'une "action" pourrait influencer sur les futures récompenses. Si tel était le cas, nous devrions réintroduire le facteur d'actualisation.

Enfin, comme nous l'avons indiqué en section 4.3, notre apprentissage est essentiellement individuel, sans échange d'information ni coordination entre les agents apprenants. L'introduction dans le système de mécanismes adéquats

devrait aussi améliorer la qualité de l'apprentissage, donc de la décision au niveau de l'ensemble des agents, et par conséquent la qualité de l'assemblage global.

## 7 Remerciements

Ce travail est en partie financé par la région Occitanie et le programme opérationnel FEDER-FSE Midi-Pyrénées et Garonne ainsi que par l'Université Paul Sabatier dans le cadre de l'opération neOCampus.

## Références

- [1] A. Aamodt and E. Plaza. Case-based reasoning : Foundational issues, methodological variations, and system approaches. *AI Communications*, 7(1) :39–59, March 1994.
- [2] S. Albrecht and P. Stone. Autonomous Agents Modelling Other Agents : A Comprehensive Survey and Open Problems. *Artificial Intelligence*, 258 :66–95, 2018.
- [3] C. Bach and D. Scapin. Adaptation of ergonomic criteria to human-virtual environments interactions. In *Proc. of Interact'03*, pages 880–883. IOS Press, 2003.
- [4] J. Boes. *Apprentissage du contrôle de systèmes complexes par l'auto-organisation coopérative d'un système multi-agent : application à la calibration de moteurs à combustion*. PhD thesis, Université de Toulouse, UPS, 2014.
- [5] Y. Charif and N. Sabouret. Un protocole de coordination d'agents introspectifs pour la chorégraphie dynamique de services. *Revue d'Intelligence Artificielle (RSTI-RIA)*, 23(1) :47–79, 2009.
- [6] K. Chen, J. Xu, and S. Reiff-Marganiec. Markov-HTN Planning Approach to Enhance Flexibility of Automatic Web Service Composition. In *IEEE Int. Conf. on Web Services*, pages 9–16. IEEE, 2009.
- [7] D. J. Cook, J. C. Augusto, and V. R. Jakkula. Ambient intelligence : Technologies, applications, and opportunities. *Pervasive and Mobile Computing*, 5(4) :277–298, 2009.
- [8] A. Cornuéjols, L. Miclet, and V. Barra. *Apprentissage artificiel : Deep learning, concepts et algorithmes*. Eyrolles, 3ème édition, 2018.
- [9] J. Coutaz. Essai sans prétention sur l'Interaction Homme-Machine et son évolution. *1024 : Bulletin de la Société Informatique de France*, (1) :15–33, 2013.
- [10] P. Doshi, R. Goodwin, R. Akkiraju, and K. Verma. Dynamic workflow composition : Using markov decision processes. *Int. Journal of Web Services Research (IJWSR)*, 2(1) :1–17, 2005.
- [11] J. Ferber. *Multi-agent systems : An introduction to distributed artificial intelligence*. Addison Wesley, 1999.
- [12] J. O. Kephart and D. M. Chess. The vision of autonomous computing. *Computer*, 36(1) :41–50, Jan. 2003.
- [13] M. E. Khanouche, F. Attal, Y. Amirat, A. Chibani, and M. Kerkar. Clustering-based and QoS-aware services composition algorithm for ambient intelligence. *Information Sciences*, 482 :419–439, 2019.
- [14] M. Koussaifi, S. Trouilhet, J.-P. Arcangeli, and J.-M. Bruel. Ambient intelligence users in the loop : Towards a model-driven approach. In *Software Technologies : Applications and Foundations*, pages 558–572. Springer, 2018.
- [15] G. Li, D. Song, L. Liao, F. Sun, and J. Du. Learning automata-based adaptive web services composition. In *5th IEEE Int. Conf. on Software Engineering and Service Science (ICSESS)*, pages 792–795. IEEE, 2014.
- [16] F. Morh. *Automated Software and Service Composition*. SpringerBriefs in Computer Science. Springer, 2016.
- [17] R. Rodrigues Filho and B. Porter. Defining emergent software using continuous self-assembly, perception, and learning. *ACM Trans. on Autonomous and Adaptive Systems*, 12(3) :16 :1–16 :25, October 2017.
- [18] R. Rouvoy, P. Barone, Y. Ding, F. Eliassen, S. O. Hallsteinsen, J. Lorenzo, A. Mamelli, and U. Scholz. MUSIC : middleware support for self-adaptation in ubiquitous and service-oriented environments. In *Software Engineering for Self-Adaptive Systems*, volume 5525 of *LNCS*, pages 164–182. Springer, 2009.
- [19] F. Sadri. Ambient intelligence : A survey. *ACM Computing Surveys*, 43(4) :1–66, October 2011.
- [20] I. Sommerville. Component-based software engineering. In *Software Engineering*, chapter 16, pages 464–489. Pearson Education, 10th edition, 2016.
- [21] T. G. Stavropoulos, D. Vrakas, and I. Vlahavas. A survey of service composition in ambient intelligence environments. *Artificial Intelligence Review*, 40(3) :247–270, September 2011.
- [22] H. Wang, X. Wang, X. Hu, X. Zhang, and M. Gu. A multi-agent reinforcement learning approach to dynamic service composition. *Information Sciences*, 363 :96–119, 2016.
- [23] H. Wang, X. Zhou, X. Zhou, W. Liu, W. Li, and A. Bouguettaya. Adaptive service composition based on reinforcement learning. In *Proc. of the Int. Conf. on Service-Oriented Computing (ICSOC)*, pages 92–107. Springer, 2010.
- [24] M. Weiser. The computer for the 21st century. *Scientific American*, 265 :94–104, 1991.
- [25] W. Younes, S. Trouilhet, F. Adreit, and J.-P. Arcangeli. Towards an intelligent user-oriented middleware for opportunistic composition of services in ambient spaces. In *Proc. of the 5th Workshop on Middleware and Applications for the Internet of Things (M4IoT)*, pages 25–30, New York, NY, USA, 2018. ACM.

# **Activité AfIA**

1<sup>er</sup> août 2017 – 31 juillet 2019

Éditeurs : Conseil d'Administration de l'AfIA – Année 2019



# Table des matières

Flyer AfIA 2019. <b>À propos de l'AfIA</b> .....	v
Conseil d'Administration de l'AfIA. <b>CA AfIA</b> .....	vii
Collèges AfIA. <b>Collèges R&amp;R, SIC, SMAA, TLH, Compétition</b> .....	ix
Prix de Thèse IA. <b>2018 et 2019</b> .....	xix
Bulletins Trimestriels. <b>Résumé des numéros 098 à 105</b> .....	xxi
Journées propres. <b>PDIA 2017, FIJA 2018, PDIA 2018, EFIA 2019, FIJA 2019</b> .....	xxiii
Compétitions. <b>Nuits de l'Info 2017 &amp; 2018</b> .....	xxix
Journées communes. <b>Période : de août 2017 à juillet 2019</b> .....	xxxi
Plates-formes IA. <b>Affiches PFIA 2018 &amp; 2019</b> .....	xliii





**AFIA**

Association française  
pour l'Intelligence Artificielle

# 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 annuelle de grands rendez-vous. Se tient ainsi chaque été une semaine de l'IA, la « Plate-forme IA » ([PFIA 2017](#) Caen, [PFIA 2018](#) Nancy, [PFIA 2019](#) Toulouse) au sein de laquelle se tiennent les « Conférence Nationale d'Intelligence Artificielle » ([CNIA](#)), « Rencontres des Jeunes Chercheurs en IA » ([RJCIA](#)) et « Conférence sur les Applications Pratiques de l'IA » ([APIA](#)) ainsi que des conférences thématiques invitées qui évoluent d'une année à l'autre, sans récurrence obligée.

Ainsi, [PFIA](#) héberge du 1er au 5 juillet 2019 à Toulouse, outre les 22<sup>ème</sup> [CNIA](#), 17<sup>ème</sup> [RJCIA](#) et 5<sup>ème</sup> [APIA](#), les 21<sup>ème</sup> [CAp](#), 30<sup>èmes</sup> [IC](#), 14<sup>èmes</sup> [JFPDA](#), 27<sup>èmes</sup> [JFSMA](#), 13<sup>èmes</sup> [JIAF](#) et les 26<sup>èmes</sup> [TALN-RECITAL](#).

Forte du soutien de ses 470 adhérents à jour de leur cotisation en 2018, l'AFIA assure :

- Le maintien d'un [Site web](#) dédié à l'IA, reproduisant également les Brèves de l'IA,
- Une journée Recherche sur les « Perspectives et Défis en IA » ([PDIA 2018](#)),
- Une journée Enseignement intitulée « Enseignement et Formation en IA » ([EFIA 2019](#)),
- Une journée Industrielle connue sous le nom « Forum Industriel en IA » ([FIIA 2019](#)),
- La remise annuelle d'un [Prix de Thèse](#) de Doctorat en IA,
- Le soutien à plusieurs Collèges, actuellement au nombre de 6, ayant leur propre activité :
  - Collège [Représentation et Raisonnement](#) (depuis avril 2017),
  - Collège [Science de l'Ingénierie des Connaissances](#) (depuis avril 2016),
  - Collège [Systèmes Multi-Agents et Agents Autonomes](#) (depuis octobre 2016),
  - Collège [Technologies du Langage Humain](#) (depuis juillet 2019),
  - Collège [Compétitions](#) (depuis octobre 2018),
  - Collège [Industriel](#) (depuis janvier 2016),
- La parution trimestrielle des [Bulletins](#) de l'AFIA, en accès libre à tous depuis le [Site web](#),
- Un lien entre membres et sympathisants sur les réseaux sociaux [LinkedIn](#), [Facebook](#) et [Twitter](#),
- Le [Parrainage](#) scientifique, mais aussi éventuellement financier, d'événements en IA,
- La diffusion mensuelle de [Brèves](#) sur les actualités de l'IA en France,
- La réponse aux consultations officielles ou officieuses (Ministères, Missions, Organismes),
- La réponse aux questions de la presse, écrite ou orale, également sur internet,
- La divulgation d'offres de [Collaborations](#), de [Formations](#), d'[Emploi](#), de [Thèses](#) et de [Stages](#).

L'AFIA organise aussi mensuellement des Journées Communes avec d'autres Associations. Sont d'ores et déjà prévues : Cla&IA avec la [SFC](#), MACS&IA avec le [GdR MACS](#), RI&IA avec l'[ARIA](#), PS&IA avec la [SPS](#), et RO&IA avec la [ROADEF](#).

Enfin, l'AFIA encourage la participation de ses membres aux grands événements de l'IA, dont PFIA. Ainsi, les membres de l'AFIA, pour leur inscription à PFIA, bénéficient d'une réduction équivalente à deux fois le coût de leur adhésion, leur permettant d'assister à PFIA sur 5 jours pour seulement [95,00 € HT](#) !

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 2018

<b>Président</b>	Yves DEMAZEAU	GT Collèges, GT Communication, GT PFIA, EF&IA 2018
<b>V.-Président</b>	Elise BONZON	Parrainages, Collèges, Représentation
<b>Trésorier</b>	Catherine FARON-ZUCKER	GT Finances, Collèges, Enseignement, PFIA 2018
<b>Secrétaire</b>	Sandra BRINGAY	Cartographie, Enseignement, Prix Thèse, PFIA 2018
<b>Webmestre</b>	Arnaud MARTIN	GT Web & Mail, Postmaster
<b>Rédacteur</b>	Dominique LONGIN	GT Bulletins, Ouvrage 2018
<b>Membres</b>	Carole ADAM	Brèves, Compétitions, Votes
	Emmanuel ADAM	Brèves, Réseaux Sociaux, Web & Mail
	Patrick ALBERT	Communication
	Florence BANNAY	Nuit Info, Dossier n°100, Compétitions
	Alain BERGER	Dossier n°099, Collège Industriel, Communication
	Frédéric MARIS	Démos et Posters PFIA 2018, 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	Réseaux Sociaux, Web & Mail, Organisation PFIA 2018
	Philippe MORIGNOT	Bulletins, Compétitions
	Bruno PATIN	Collège Industriel, Dossier n°101
	Gauthier PICARD	GT Enseignement, Plateforme
	Olivier SIMONIN	Compétitions, PDIA 2018
	Serena VILLATA	Adhésions, Finances, Journées, FR@IJCAI 2018

## Conseil d'Administration 2019

<b>Président</b>	Yves DEMAZEAU	GT Communication, GT Plate-Forme, FIIA 2019
<b>V.-Président</b>	Elise BONZON	GT Collèges, Parrainages, Représentation
<b>Trésorier</b>	Catherine FARON-ZUCKER	GT Finances, GT Journées, EFIA 2019, PFIA 2019
<b>Secrétaire</b>	Sandra BRINGAY	Cartographie, Prix Thèse, EFIA 2019, PFIA 2019
<b>Webmestre</b>	Emmanuel ADAM	GT Web&Mail, Brèves, Réseaux Sociaux, Communication
<b>Rédacteur</b>	Dominique LONGIN	GT Bulletins, Organisation PFIA 2019, Ouvrage 2019
<b>Membres</b>	Alain BERGER	Collège Industriel, Communication, Représentation
	Gaël DIAS	Collèges, Dossier n°107
	Pierre FEILLET	FIIA 2019, Réseaux Sociaux, Dossier n°109
	Jean-Marie LAGNIEZ	Collège Industriel, PDIA 2019
	Marie LEFEVRE	Adhésions, Journées, Dossier n°108, Enseignement
	Arnaud MARTIN	Postmaster, Dossier n°103
	Frédéric MARIS	Organisation PFIA 2019, Démos et Posters PFIA 2019
	Engelbert MEPHU NGUIFO	Journées Communes, Dossier n°106
	Davy MONTICOLO	Réseaux Sociaux, Web & Mail
	Philippe MORIGNOT	Bulletins, Compétitions
	Gauthier PICARD	GT Enseignement, EFIA 2019, Tutoriels PFIA 2019
	Laurent SIMON	Dossier n°105, PDIA 2019, Réseaux Sociaux
	Olivier SIMONIN	Dossier n°104, Compétitions
	Serena VILLATA	Adhésions, Finances, Journées, FR@International 2018





**AfIA**

Association française  
pour l'Intelligence Artificielle

# 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.

## Activités d'animation générales

Le collège R&R est impliqué dans les activités suivantes :

- IAF : Journées d'Intelligence Artificielle Fondamentales. Ces journées ont lieu tous les ans, à l'initiative du comité IAF. Elles sont articulées autour de 3 ou 4 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 JFPDA et des participants à la liste de diffusion [jfpda@loria.fr](mailto:jfpda@loria.fr). Elles sont articulées autour de 2 ou 3 exposés de synthèse invités, ainsi que d'un programme constitué après appel à communication.
- JFPC : Journées Francophones de Programmation par Contraintes. Ces journées ont lieu tous les ans à l'initiative de l'AFPC (Association Française pour la Programmation par Contraintes). Elles sont

articulées autour de 2 ou 3 exposés de synthèse invités, ainsi que d'un programme constitué après appel à communication.

- des ateliers thématiques, qui ont lieu lors d'autres événements et la contribution à l'organisation de journées communes.

Le Collège consacre une partie de son budget de fonctionnement à l'attribution de bourses permettant à des étudiants d'assister à l'un des événements organisés par le Collège, ou d'autres manifestations pertinentes au regard des thématiques scientifiques.

## Comité de Pilotage

Le comité de pilotage du Collège est constitué en 2019 des personnes suivantes :

- Elise Bonzon, vice-présidente de l'AFIA  
Université Paris-Descartes, LIPADE, Paris
- Sylvie Doutre, co-présidente du comité de programme IAF  
Université Toulouse 1, IRIT, Toulouse
- Sébastien Konieczny, directeur du GDR IA  
CNRS, CRIL, Lens
- Tiago de Lima, co-président du comité de programme IAF  
Université d'Artois, CRIL, Lens
- Frédéric Maris, membre du CA et représentant pour l'AFIA  
Université Paul Sabatier, IRIT, Toulouse
- Nicolas Maudet  
Université Pierre et Marie Curie, LIP6, Paris
- François Schwarzentruher  
ENS Rennes, IRISA, Rennes
- Laurent Simon, vice-président de l'AFPC  
Université de Bordeaux, LaBRI, Bordeaux
- Elise Vareilles, membre du CA de l'AFPC  
IMT Mines, Albi
- Bruno Zanuttini, co-président du comité de programme JFPDA  
Université Caen Normandie, GREYC, Caen

La composition du comité est amenée à être modifiée bi-annuellement.

## Contacts

Coordinateur du Collège R&R : [frederic.maris@irit.fr](mailto:frederic.maris@irit.fr)

Listes de diffusion du Collège R&R : [bull-i3@irit.fr](mailto:bull-i3@irit.fr), [bull-ia@gdria.fr](mailto:bull-ia@gdria.fr), [jfpda@loria.fr](mailto:jfpda@loria.fr)

Trois membres du comité de pilotage du collège sont membres du Conseil d'Administration de l'AFIA :

- Elise Bonzon, [elise.bonzon@mi.parisdescartes.fr](mailto:elise.bonzon@mi.parisdescartes.fr)
- Frédéric Maris [frederic.maris@irit.fr](mailto:frederic.maris@irit.fr)
- Laurent Simon [lsimon@labri.fr](mailto:lsimon@labri.fr)



**AFIA**

Association française  
pour l'Intelligence Artificielle

# 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. 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. 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 sémantique, dans l'ensemble de la société engendre des mutations dans les pratiques individuelles et collectives. L'ingénierie des connaissances accompagne cette évolution, en créant les modèles, les méthodes, les algorithmes et les outils facilitant l'intégration de connaissances et de raisonnements automatiques sur ces dernières dans des environnements informatiques.

La gestion des connaissances associe ingénierie des connaissances avec recherche d'information, recherche opérationnelle, analyse de données et organisation humaine associée. La modélisation des environnements informatiques, de ses acteurs humains ou logiciels, allie ingénierie des connaissances et systèmes multi-agents. 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, raisonner, prédire, approximer, sécuriser, fiabiliser, recommander, expliquer, valider, pour pallier leur imprécision voire leur incohérence. Les approches classiques mises en œuvre sont actuellement enrichies par l'usage complémentaire d'approches hybrides (numérique et symbolique).

## Programme de Travail

Le collège Ingénierie des Connaissances a premièrement 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. Il soutient l'organisation de la conférence internationale Terminology & Ontology : Theories and applications (Toth). 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
- Djamel 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](mailto:sylvie.despres@univ-paris13.fr)

Liste de diffusion Ingénierie des Connaissances : [info-ic@inria.fr](mailto: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](mailto:faron@unice.fr)
- Sandra Bringay, [sandra.bringay@lirmm.fr](mailto:sandra.bringay@lirmm.fr)



**AFIA**  
Association française  
pour l'Intelligence Artificielle

# 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. Un partenariat avec le collège industriel de l'AFIA, notamment dans l'organisation de FIIA 2019, permettra d'afficher les réalisations industrielles de nos domaines scientifiques.

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
- Fabien Michel  
Université de Montpellier, LIRMM, Montpellier
- Frédéric Migeon  
Université Paul Sabatier, IRIT, Toulouse
- Maxime Morge  
Université de Lille 1, CRISAL, Lille
- Gauthier Picard  
Mines Saint-Etienne, LHC, Saint-Etienne
- Nicolas Sabouret  
Université Paris-Sud, LIMSI, Saclay
- Laurent Vercouter  
INSA Rouen Normandie, LITIS, Rouen
- Mahdi Zargayouna  
IFSTTAR, Paris

### **Contacts**

Coordinateur du Collège : [laurent.vercouter@insa-rouen.fr](mailto:laurent.vercouter@insa-rouen.fr)

Listes de diffusion : [sma@loria.fr](mailto:sma@loria.fr), [acai@poleia.lip6.fr](mailto:acai@poleia.lip6.fr)

Deux membres du comité de pilotage du collège sont membres du conseil d'administration de l'AFIA :

- Emmanuel Adam [emmanuel.adam@univ-valenciennes.fr](mailto:emmanuel.adam@univ-valenciennes.fr)
- Gauthier Picard [gauthier.picard@emse.fr](mailto:gauthier.picard@emse.fr)





**AFIA**

Association française  
pour l'Intelligence Artificielle

# Le Collège Technologies du Langage Humain

## Objectif du Collège

Les Technologies du Langage Humain (TLH) proposent des méthodes permettant une communication homme-machine naturelle, pouvant s'étendre à une interaction homme-homme médiée. Ainsi, les TLH permettent d'analyser, d'interpréter et de produire des actes du langage écrit, parlé ou signé, mais aussi d'interagir avec des données langagières. Ainsi, les TLH englobent traditionnellement le Traitement Automatique des Langues (TAL), la Communication Parlée (CP) et leurs applications les plus emblématiques comme la Recherche d'Information (RI) et la Traduction Automatique.

Les TLH occupent une part entière en IA sous le prisme du Test de Turing. Ainsi, elles regroupent tous les axes de recherche de l'IA dans le cadre des données langagières comme la représentation (e.g. plongements lexicaux, analyse syntaxique), le raisonnement (e.g. systèmes de question-réponse), la planification (e.g. argumentation), l'apprentissage (e.g. analyse de sentiments), ou même l'intelligence collective (e.g. détection de communautés). Créer des modèles pouvant interagir de manière élégante et naturelle en utilisant la langue nécessite une compréhension profonde de l'acoustique, de la phonétique et de la prosodie pour le langage oral d'une part (dans le cadre de la reconnaissance et de la synthèse de la parole), et d'autre part, de la morphologie, de la syntaxe, de la sémantique et de la pragmatique pour le langage écrit ou signé (dans le cadre de l'analyse et de la génération de textes). Seulement à cette condition les applications des TLH peuvent offrir la promesse d'un accès « universel » à l'information, notamment à travers les moteurs de recherche, la traduction automatique, le résumé de textes, la veille automatique ou les systèmes de question-réponse. La compréhension globale du langage permet également de caractériser les textes suivant leurs objectifs communicationnels. Ainsi, l'analyse des sentiments et des émotions, l'identification de discours haineux, la détection de plagiat, l'identification et la vérification du locuteur proposent autant de cadres applicatifs importants pour les sociétés numériques.

Parallèlement au développement de modèles des TLH, de nombreux défis connexes doivent être pris en compte dans le cadre d'une démarche globale, comme la confiance (e.g. reproductibilité, explicabilité, confidentialité), l'éthique (e.g. biais d'apprentissage, représentativité) et l'évaluation (e.g. métriques dédiées) des systèmes proposés.

Puisqu'à l'ère du numérique les données langagières prolifèrent dans des quantités toujours plus importantes et accessibles (*big data*), les champs d'applications des TLH sont nombreux. Ainsi, les archives numériques, les réseaux sociaux, les plateformes collaboratives, les entretiens clients, les textos, les courriels, les commentaires sur des plateformes de vente en ligne sont autant de matières premières pour le développement d'applications des TLH. En particulier, quelques domaines privilégiés sont la santé, l'éducation, le droit, le journalisme et le handicap, mais d'autres émergent comme la finance, l'agriculture, la sécurité, le marketing et les humanités numériques.

Le collège TLH de l'AFIA a donc pour mission de promouvoir l'animation et l'interaction scientifiques entre les communautés TAL, CP et RI, et l'ensemble des communautés en IA ayant des intérêts communs dans le but de consolider les collaborations transversales.

## Programme de Travail

Afin de mener à bien sa mission, le collège TLH s'engage à soutenir l'organisation de manifestations scientifiques (conférences, ateliers), animer des groupes de travail, éditer des dossiers techniques, organiser

des journées thématiques et diffuser et communiquer autour des recherches des communautés françaises du TAL, de la CP, de la RI et de l'IA.

Le collège TLH s'engage également à servir de canal de communication entre l'AFIA et ses collègues ainsi qu'avec l'AFCP (Association Francophone de la Communication Parlée), l'ARIA (Association Francophone de Recherche d'Information et Applications), et avec l'ATALA (Association pour le Traitement Automatique des Langues).

### **Comité de Pilotage**

Le comité de pilotage du collège est constitué de 10 chercheuses et chercheurs spécialisés dans le domaine du TAL, de la CP et de la RI.

- Florian Boudin  
Université de Nantes, LS2N, Nantes
- Davide Buscaldi  
Université Paris 13, LIPN, Villetaneuse
- Gaël Dias  
Université de Caen Normandie, GREYC, Caen
- Corinne Fredouille  
Université d'Avignon, LIA, Avignon
- José Moreno  
Université Paul Sabatier, IRIT, Toulouse
- Aurélie Névéol  
CNRS, LIMSI, Saclay
- Yannick Parmentier  
Université de Lorraine, LORIA, Nancy
- François Portet  
Grenoble INP, LIG, Grenoble
- Mathieu Roche  
CIRAD, TETIS, Montpellier
- Serena Villata  
CNRS, I3S, Sophia Antipolis

### **Contacts**

Coordinateur du collège : [mathieu.roche@cirad.fr](mailto:mathieu.roche@cirad.fr)

Listes de diffusion : [ln@cines.fr](mailto:ln@cines.fr), [parole@listes.afcp-parole.org](mailto:parole@listes.afcp-parole.org), [info-aria@lisis.org](mailto:info-aria@lisis.org)

Deux membres du comité de pilotage sont membres du conseil d'administration de l'AFIA :

- Gaël Dias [gael.dias@afia.asso.fr](mailto:gael.dias@afia.asso.fr)
- Serena Villata [villata@i3s.unice.fr](mailto:villata@i3s.unice.fr)

# Collège Compétitions

## Objectif du collège

Le collège Compétitions est impliqué dans l'ensemble des compétitions auxquelles l'AFIA est amené participer. Les activités de ce Collège prolongent les compétitions antérieures organisées par l'AFIA : « IA sur Robot » lors de RFIA'16 à Clermont-Ferrand, et la « BotContest » lors de PFIA'17 à Caen.

## Activités du collège

A ce jour, les compétitions identifiées par le Collège sont les suivantes :

- **Nuit de l'Info.** Cette compétition a lieu tous les ans en décembre. Les participants programment en une nuit sur un défi national tout en relevant divers défis, ils remettent leur programme et un document le lendemain matin à 8h. Ces résultats sont évalués par un jury national et les jurys des défis. Cette compétition a lieu tous les ans en décembre. L'AFIA propose le défi d'intégrer l'IA dans l'application développée, le jury se compose de membres du collège et d'autres membres de l'AFIA.
- **IA sur Robots.** Le but de cette compétition est de mettre en évidence l'IA au sein d'une ou plusieurs plateformes robotiques, dans un scénario figé avec des règles, et une mesure des performances. Cette compétition aura lieu en 2019 à Toulouse lors de PFIA. Elle est ouverte à tous : chercheurs, ingénieurs, étudiants et grand public. Cette nouvelle édition s'intitule "DriveToGaether" et dispose d'un scénario et d'un règlement.
- **IA et Jeux Vidéos.** Le but de cette compétition est de faire expérimenter aux conférenciers la créativité, l'originalité, la performance stratégique et/ou l'adaptativité d'un logiciel d'IA sur écran développé par les participants. Cette compétition aura lieu en 2020 lors de PFIA. La thématique de la compétition n'est cependant pas figée, et pourra être redéfinie en fonction de l'évolution du Comité de Pilotage et de l'identité du Comité d'Organisation de PFIA 2020.

## Comité de Pilotage

- Carole Adam  
Université Grenoble Alpes, LIG
- Florence Bannay  
Université Paul Sabatier, IRIT
- Dominique Busso  
Atlangames
- Anne-Gwenn Bosser  
Ecole Nationale des Ingénieurs de Brest, Lab-STICC



# AFIA

Association française  
pour l'Intelligence Artificielle

- Cindy Even  
Virtualys
- Christophe Leroux (LRI, CEA LIST, Saclay)  
Commissariat à l'Energie Atomique, LIST
- Philippe Morignot  
Aspertise
- Nicolas Pépin-Hermann  
BA Systèmes
- Syvie Saget  
CLASP/FLOV
- Olivier Simonin  
INSA Lyon, CITI

La composition de ce comité sera modifiée annuellement.

## Contacts

Coordinatrice du Collège Compétitions : Philippe Morignot, [pmorignot@yahoo.fr](mailto:pmorignot@yahoo.fr)

Site web : <http://afia-competitions.fr>

Adresse électronique du collège : [contact@afia-competitions.fr](mailto:contact@afia-competitions.fr)

Quatre membres du comité de pilotage sont membres du CA de l'AFIA :

- Florence Bannay, [bannay@irit.fr](mailto:bannay@irit.fr)
- Frédéric Maris, [maris@irit.fr](mailto:maris@irit.fr)
- Philippe Morignot, [pmorignot@yahoo.fr](mailto:pmorignot@yahoo.fr)
- Olivier Simonin, [olivier.simonin@insa-lyon.fr](mailto:olivier.simonin@insa-lyon.fr)



**AfIA**

Association française  
pour l'Intelligence Artificielle

# Prix de Thèse IA 2018

## Le Jury

- Président : Jean-Gabriel Ganascia (LIP6, UMPC)
- Membres : Rachid Alami, Salem Benferhat, Gauvain Bourgne, Nicolas Bredeche, Stéphane 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, Thomas Schiex, Julien Velcin.

## Les Lauréats

- 1<sup>er</sup> Prix (ex-aequo) : Nawal Benabbou "Procédures de décision par élicitation incrémentale de préférences en optimisation multicritère, multi-agents et dans l'incertain". 05/05/17. Dir. : Patrice Perny (LIP6, Univ. Paris 06).
- 1<sup>er</sup> Prix (ex-aequo) : Daniele Calandriello "Efficient Sequential Learning". 18/12/17.<sup>[1]</sup>Dir. : Michal Valko (CRISTAL, Univ. Lille 1) et Alessandro Lazaric (CRISTAL, Univ. Lille 1).
- Accessit : Théo Trouillon "Complex-Valued Embeddings Models for Knowledge Graphs". 29/07/17.<sup>[1]</sup>Dir. : Christopher Dance (Naver Labs, Grenoble) et Eric Gaussier (LIG, Univ. Grenoble Alpes).

# Prix de Thèse IA 2019

## Le Jury

- Président : Thomas Schiex (MIAT, INRA)
- Membres : Leila Amgoud, Salem Benferhat, Christian Bessière, Antoine Bordes, Nicolas Dobigeon, Hélène Fargier, Fabien Gandon, Yves Grandvalet, Andreas Herzig, Jérôme Lang, Pierre Marquis, Marie-Christine Rousset, Michèle Sebag.

## Les Lauréats

- 1<sup>er</sup> Prix (ex-aequo) : Anaëlle Wilczynski "Interaction entre agents modélisée par un réseau social en choix computationnel". 13/12/18. Dir. : Laurent Gourves (LAMSADE, Univ. Paris Dauphine).
- 1<sup>er</sup> Prix (ex-aequo) : Valentina Zantedeschi "Une vue unifiée de l'apprentissage local : théories et algorithmes pour l'amélioration de modèles linéaires". 18/12/18. Dir. : Marc Sebban (LHC, Univ. Jean Monnet).
- Accessit : Joe Raad "Gestion d'identité dans des graphes de connaissances". 30/11/18. Dir. : Juliette Dibie (AgroParisTech, Paris) et Nathalie Pernelle (LRI, Univ. Paris-Sud).





**AFIA**

Association française  
pour l'Intelligence Artificielle

# 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 2018 comprenait Dominique Longin comme rédacteur en chef, aidé de Emmanuel Adam, Claire Lefèvre, et Philippe Morignot, comité reconduit en 2019.

### 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 « Gestion des Connaissances »

janvier 2018

- Ce bulletin contient un dossier industriel sur l'« État des lieux de la Gestion des Connaissances 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 le 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 100<sup>e</sup> 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 3<sup>e</sup> 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 « Equipes industrielles en I.A. »**                      **juillet 2018**

- Le dossier de ce bulletin, dirigé par Dominique Longin, présente les travaux de six équipes de recherche industrielles dans le domaine de l'IA (Ardans, Berger-Levrault, Dassault Aviation, Mondeca, Renault et Thalès) ainsi qu'un exemple de convention entre un industriel et l'AFIA.
- Ce bulletin contient également 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 ceux des JFPC 2018 (Journées Francophones de Programmation par Contraintes) et des JIAF 2018 (Journées d'Intelligence Artificielle Fondamentale) 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.

**Bulletin 102**                      **Dossier « PFIA 2018 »**                      **septembre 2018**

- Le dossier de ce bulletin est consacré à la PFIA 2018 qui s'est déroulée du 2 au 6 juillet 2018 à Nancy, et qui hébergeait cette année : APIA, CNIA, IC, JFPDA et RJCIA ; les journées : éthique et IA, France@IJCAI2018, IA pour l'éducation, IA et Santé, Robotique et IA, Traitement des langues et IA. Les participants ont également pu assister à des démonstrations pendant deux jours entiers.
- Ce bulletin contient également le compte rendu de la journée Statistiques et IA qui s'est tenue le 25 septembre 2018 à Paris, ainsi que l'ensemble des soutenances de thèses et HDR dont nous avons eu connaissance et soutenues lors du troisième trimestre 2018.

**Bulletin 103**                      **Dossier « I.A. et Réseaux Sociaux »**                      **janvier 2019**

- Ce bulletin contient un dossier thématique sur « Intelligence artificielle et réseau sociaux » sous la direction de Arnaud Martin, regroupant pas moins de huit contributions sur ce thème.
- Il contient également les comptes rendus des 26<sup>e</sup> JFSMA (Journées Francophones sur les Systèmes Multi-Agents) qui se sont déroulées du 10 au 12 octobre 2018 à Métabief (Doubs), de la 5<sup>e</sup> journée PDIA (Perspectives et Défis en IA) qui s'est déroulée le 11 octobre à Paris, et de la Nuit de l'info qui s'est tenue la nuit du 6 au 7 décembre 2018 sous la direction de Florence Bannay. Il contient également l'ensemble des soutenances de thèses et HDR dont nous avons eu connaissance et soutenues lors du dernier trimestre 2018.

**Bulletin 104**                      **Dossier « I.A. et Robotique Mobile »**                      **avril 2019**

- Le dossier de bulletin, dirigé par Olivier Simonin, porte sur le thème « IA & Robotique mobile » et ne contient pas moins de 19 contributions permettant d'avoir une vision très précise des travaux actuellement menés en France dans ce domaine.
- Ce bulletin contient également le compte rendu de la 1<sup>re</sup> journée EFIA (Enseignement et Formation en IA) qui s'est déroulée le 10 janvier 2019 à Nanterre, et celui de la 4<sup>e</sup> journée IHM-IA organisée à Telecom ParisTech le 14 mars 2019. Ce bulletin contient également l'ensemble des soutenances de thèses et HDR dont nous avons eu connaissance et soutenues lors du premier trimestre 2019.

**Bulletin 105**                      **Dossier « Equipes Académiques en I.A. »**                      **juillet 2019**

- Le dossier de ce bulletin, dirigé par Laurent Simon, présente des équipes de recherche académiques en IA. Il est actuellement en cours de finalisation et devrait paraître sous peu.
- Ce bulletin contient également un compte rendu : de FIIA 2019 (Forum Industriel de l'IA) qui s'est tenu le 11 avril 2019 à Champ-sur-Marne ; de la journée EGC&IA (Extraction et Gestion des Connaissances & IA) organisée le 10 mai 2019 à l'Université Paris-Sud ; de la conférence internationale PAAMS (Practical Applications of Agents and Multiagent Systems) qui s'est déroulée du 26 au 28 juin 2019 à Avila (Espagne). Ce bulletin contient également l'ensemble des soutenances de thèses et HDR dont nous avons eu connaissance et soutenues lors du deuxième trimestre 2019.



# 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.



# AFIA

Association française  
pour l'Intelligence Artificielle



## FIIA 2018

Ce 3<sup>ème</sup> 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 Bayouhd (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.



# PDIA 2018

L'Association Française pour l'Intelligence Artificielle (AFIA) organise sa cinquième journée PERSPECTIVES ET DEFIS DE l'IA sur le thème : VEHICULE AUTONOME ET INTELLIGENCE ARTIFICIELLE. On parle beaucoup de Véhicule Autonome, sans toujours faire la connexion avec l'IA. On parle encore plus d'IA, mais sans forcément penser à toutes les possibilités pour les Véhicules Autonomes. Les défis sont nombreux, parmi lesquels la sûreté et la validation des algorithmes embarqués, la perception et la localisation, l'interaction avec l'humain, la prise de décision temps réel, la navigation en trafic dense. La journée vise à en dresser un panorama, au travers d'exposés et de tables rondes.

Cette journée se tient en collaboration avec le MTES Direction Générale des Infrastructures, des Transports et de la Mer (DGITM) et le partenariat du MINEF Direction Générale des Entreprises (DGE).

## Date et Lieu

- Date : 11 octobre 2018
- Lieu : Université Paris-Descartes, Amphithéâtre Fourier, 45 rue des Saints Pères, 75006 Paris

## Programme

- 08h30 Accueil
- 08h45 Ouverture de F. Dardel (*Président de l'Université Paris Descartes*), Y. Demazeau (*Président de l'AFIA*) et O. Simonin (*Membre du CA de l'AFIA*).
- 09h00 « L'IA au service du transport autonome » par P. Aknin (*IRT SystemX*)
- 09h30 « SMA pour la simulation et la modélisation du trafic » par Ph. Mathieu (*Université Lille*)
- 10h00 Pause
- 10h30 « Le défi de la validation en sécurité de l'IA pour la perception et la décision en conduite autonome de train » par Ph. David (*SNCF*)
- 11h00 « Validation et vérification du logiciel pour les systèmes autonomes » par F. Ingrand (*CNRS*)
- 11h30 Table ronde « Validation / Vérification dans les différents secteurs » avec Ph. David (*SNCF*), F. Ingrand (*LAAS*), et A. Piperno (*UTAC-CERAM*).
- 12h30 Buffet
- 14h00 « Validation des systèmes de conduite autonomes » par X. Delache (*Sous-directeur des études et de la prospectives, DGITM*)
- 14h30 « Les aspects éthiques du déploiement du véhicule autonome » par E. Dogan (*VEDECOM*)
- 15h00 « IA intégrée / à intégrer au sein des systèmes robotiques militaires pour l'autonomie » par B. Steux, J.-F. Massol et B. Ricaud (*Nexter Systems*)
- 15h30 « Navigation sociale des véhicules autonomes » par Anne Spalanzani (*UGA*)
- 16h00 Pause
- 16h30 Table ronde « Plates-formes d'intégration et relations avec l'humain » avec E. Dogan (*VEDECOM*), J. Lemaire (*DGA*) et O. Simonin (*INSA Lyon*)
- 17h30 Clôture

## Organisation

Cet événement est organisé par Olivier Simonin et Yves Demazeau 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 : <https://afia.asso.fr/inscription-pdia-2018/>. Le déjeuner est offert aux personnes qui se seront inscrites à PDIA 2018 avant le 08/10/18, et seulement à celles-ci.

# EFIA 2019

L'Association Française pour l'Intelligence Artificielle (AFIA) organise sa première journée Enseignement et Formation en IA. La journée a deux objectifs : d'une part, réaliser un tour d'horizon des besoins des entreprises et de la société civile. D'autre part, de dresser un état de l'art des formations académiques en IA qui existent en France et qui participent à la démocratisation de son utilisation.

## Date et Lieu

- Date : 10 janvier 2019
- Lieu : Université Paris Nanterre, Amphithéâtre Weber, 200 av. de la République, 92000 Nanterre

## Programme

- 09h00 Accueil
- 09h15 Ouverture de Jean-François Balaudé (*Président de l'Université Paris Nanterre*) et Gauthier Picard (*Membre du CA de l'AFIA*).
- 09h30 « Présentation de l'AFIA » par Yves Demazeau (*Président de l'AFIA*).
- 10h00 « Regard sur l'IA en France vu du CNU section 27 » par Grégory Bonnet (*CNU Section 27*)
- 10h30 Pause
- 11h00 « Développement de compétence Analytics et IA, retour d'expérience Airbus et perspectives » par Bruno Carron (*Airbus*)
- 11h30 « Retour sur la sémantique : comment caractériser et distinguer connaissance, intelligence et conscience artificielle » par Bernard Georges (*Société Générale*)
- 12h00 Buffet
- 14h00 « Évolution de l'enseignement en IA » par Jean-Gabriel Ganascia (*Sorbonne Université*)
- 14h30 « La place de l'IA dans la formation au lycée » par Christine Froidevaux (*Université Paris-Sud*)
- 15h00 « Les formations en Sciences de Données et Intelligence Artificielle à Sorbonne Université » par Bernd Amann, Vincent Guigne, Nicolas Bredeche (*Sorbonne Université*)
- 15h30 « Enseignement et formation en IA en ligne : retour d'expérience sur le MOOC "Web sémantique et Web de données" » par Catherine Faron-Zucker (*Université Nice Sophia Antipolis*)
- 16h00 « Pour une vision unifiée de l'intelligence artificielle : un nouvel ouvrage » par Odile Papini (*Université Aix-Marseille*)
- 16h30 Pause
- 17h00 Table ronde « Quels sujets tutoriels pour PFIA 2019 ? », animée par Gauthier Picard (*Ecole des Mines de Saint-Etienne*)
- 18h00 Clôture

## Organisation

Cet événement est organisé par Gauthier Picard, Sandra Bringay et Catherine Faron-Zucker pour 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 : <https://afia.asso.fr/inscription-efia/>. Le déjeuner est offert à toutes les personnes qui se seront inscrites à la journée avant le 07/01/19, et seulement à celles-ci.



# FIIA 2019

Ce quatrième FORUM INDUSTRIEL DE L'INTELLIGENCE ARTIFICIELLE est consacré aux SYSTEMES MIXTES avec pour sous-thèmes : capteurs, données et connaissances, intégration numérique et symbolique, hybridation, agents artificiels et humains, et une dominante applicative sur la ville sécurisée et durable. 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 en partenariat avec le MA Direction Générale de l'Armement.

## Date et Lieu

- Date : 11 avril 2019
- Lieu : Amphithéâtre Bienvenue, IFSTTAR, 14-20 bd Newton, 77420 Champ-sur-Marne

## Programme

- 08h45 Ouverture de Joaquin Rodriguez (IFSTTAR). Introduction par Pierre Feillet (CA AFIA)
- 09h00 THEME « Ville sécurisée et durable. L'Intelligence Artificielle à l'épreuve des territoires »
  - « IA et Territoires Urbains » par Ahmed Mabrouk (ENGIE *Saint-Denis*)
  - « La Vision par Ordinateur pour les villes intelligentes » par Levi Viana (WINTICS *Paris*)
  - Table ronde
- 10h00 Pause
- 10h30 THEME « Capteurs, données et connaissances. Intégration numérique et symbolique »
  - « Optimisation au quotidien grâce à l'IA et aux données disponibles » par Arnaud de Moissac (DC Brain *Paris*)
  - « IA et Ingénierie Augmentée (IA2), ou comment hybrider IA, connaissance métiers et modèles physiques » par Patrice Aknin (SystemX *Gif-sur-Yvette*)
  - World Sensing : quels systèmes d'IA pour aller du signal capteur à la compréhension du comportement observé » par Youssef Miloudi (Berger-Levrault *Boulogne-Billancourt*)
  - IA connexionniste et IA symbolique : une synergie prometteuse » par Juliette Mattioli (Thales *Palaiseau*) et Christophe Meyer (Thales SIX GTS *Palaiseau*)
  - « Décisions algorithmiques: des modèles prescriptifs pour encadrer l'apprentissage automatique » par Pierre Feillet (IBM France *Gentilly*)
  - Table ronde
- 12h30 Buffet
- 14h00 THEME « Hybridation des composants. Interactions entre agents artificiels et humains »
  - « Booster les performances des applications sémantiques par hybridation des composants IA » par Ghislain Atemezing (Mondeca *Paris*)
  - Exploitation des modèles *whitebox* dans les interactions humain-IA » par Matthieu Boussard (CRAFT AI *Paris*)
  - « Initiatives et interactions mixtes pour systèmes autonomes » par Christophe Guettier (Safran *Corporate*)
  - « Collaboration agent humain - agent virtuel dans l'avion de combat du futur » par Lauren Dargent et Michael Huet (Dassault Aviation *Saint Cloud*)
  - Collaboration intelligence humaine et intelligence artificielle en vue d'une société conviviale » par Christophe Denis (Sorbonne Université *Paris* & EDF *Saclay*)
  - Table Ronde
- 16h00 Pause
- 16h30 Présentations rapides de sociétés concernées par les systèmes mixtes en « Trois planches »

## Organisation

Cet événement est organisé par Pierre Feillet et Yves Demazeau pour le Collège Industriel de l'AFIA et Caroline Chopinaud pour le Collège Systèmes Multi-Agents et Agents Autonomes 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 : <https://afia.asso.fr/inscription-fjia/>. Le déjeuner est offert aux personnes qui se seront inscrites à la journée avant le 08/04/19, et seulement à celles-ci.







**AFIA**

Association française  
pour l'Intelligence Artificielle

# 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 3<sup>e</sup> 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 ».

## Le 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 bibliothèques existantes. La solution présentée a été jugée originale et adaptée au sujet national.



# AFIA

Association française  
pour l'Intelligence Artificielle

## Nuit de l'Info 2018

La Nuit de l'Info 2018 a eu lieu de 16h39 à 8h05 dans la nuit du 6 au 7 décembre sur le thème de *L'assistance numérique d'un explorateur en condition extrêmes*. Le principe de la Nuit de l'Info est très simple : Le soir d'avant à 16h39, 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 jour d'après à 8h05, les développements sont figés. Puis, pendant que les équipes 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 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 est un montant de 400€ assorti d'invitations à présenter une démonstration lors de la compétition organisée par l'AFIA au sein de la Plate-Forme IA 2019 à Toulouse.

### Les Equipes

Parmi les 12 équipes inscrites pour relever ce défi, 5 ont réussi à remettre à temps un projet : GIFCOMS, ISITCOM\_Shadow\_Brokers2, ISITCOM\_Techno\_Knights, Lord of the Ping et MLG.

### Le Jury

Le jury était composé de Florence Bannay (coordinatrice), Anne-Gwenn Bosser, Philippe Morignot et Sylvie Saget. 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

Nous avons eu du mal à départager les deux meilleures équipes : l'équipe *Lord of the Ping* a finalement été déclarée vainqueur pour le recul de ses explications et recevra un prix de 200 euros. Néanmoins vu la qualité des résultats de l'équipe *MLG*, elle recevra la même somme de 200 euros.

Le classement final est le suivant :

- 1- *Lord of the Ping* a développé un chatbot sarcastique avec détection de serpent par apprentissage. L'équipe est allée au bout de ses implémentations, a bien expliqué les problèmes et les limites des approches utilisées.
- 2- *MLG* a construit un site web permettant de saisir des données, de visualiser les courbes correspondantes et de donner les positions des robots. Le jury a apprécié les résultats présentés au travers de vidéos.
- 3- *ISITCOM\_Techno\_Knights* propose une montre connectée avec différents capteurs pour détecter l'état psychologique et pouvoir une alerte adaptée. Mais les détails techniques et implémentation étaient absents.
- 4 et 5- *ISITCOM\_Shadow\_Brokers2* et *GIFCOMS* ont proposé des projets dans lesquels le point de vue IA était peu développé.

L'AFIA remercie tous les participants pour le travail fourni et donne rendez-vous l'année prochaine pour un nouveau défi de l'AFIA !



# 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-wcfavIXVA0Y18-EhCqObG5v3HQ/viewform](https://docs.google.com/forms/d/e/1FAIpQLSfsM5Cir_LDxWov32a4eE-wcfavIXVA0Y18-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.



**AFIA**

Association française  
pour l'Intelligence Artificielle



# 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 Deringent 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](mailto: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 (*CDiscout 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.



**AFIA**

Association française  
pour l'Intelligence Artificielle

**EGC**

# 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 Benkhalif (*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](https://egc18.sciencesconf.org/resource/page/id/20). 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 incorporée 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 algorithmes » 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/>.



**AFIA**

Association française  
pour l'Intelligence Artificielle



# 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 événements 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 GNum2 de la Direction du Numérique pour l'Éducation*).
- 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/>.



**AFIA**

Association française  
pour l'Intelligence Artificielle



# Robotique & IA 2018

Cette troisième journée ROBOTIQUE et INTELLIGENCE ARTIFICIELLE fait se rencontrer les deux domaines autour des thèmes APPRENTISSAGE POUR LA ROBOTIQUE HUMANOÏDE 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 Bonnifant (*Sorbonne Université Compiègne*)
- 12h00 Déjeuner
- 13h30 « Apprentissage par Renforcement Développement » 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 REPRESENTATIONS 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://pfa2018.loria.fr/inscriptions/>.



**AFIA**

Association française  
pour l'Intelligence Artificielle



# Statistique & IA 2018

Cette première journée STATISTIQUE et INTELLIGENCE ARTIFICIELLE fait se rencontrer les deux domaines autour du thème de l'APPRENTISSAGE MACHINE.

Les progrès récents en apprentissage machine font actuellement l'objet d'investigations intenses dans diverses communautés scientifiques et les applications industrielles foisonnent. Différentes communautés s'attellent à l'étude d'une multitude d'approches, souvent complémentaires. Cette journée aura pour but de mettre en relief les liens forts entre l'Intelligence Artificielle et les Statistiques, et stimuler les discussions autour des interactions futures entre ces deux communautés. Algorithmes, challenges et impacts, et analyses mathématiques fines seront les pôles principaux des exposés invités.

## Date et Lieu

- Date 25 septembre 2018
- Lieu Institut Henri Poincaré, 11 rue Pierre et Marie Curie 75005 Paris

## Programme

- 09h00 Présentations de l'AFIA par Yves Demazeau, son Président, et de la SFdS par Gérard Biau, son Président.
- 09h15 « Unsupervised learning, from maximum entropy to deep generative networks » par Stéphane Mallat (*Collège de France Paris*)
- 10h15 « Robust machine learning via median-of-means » par Guillaume Lecue (*CNRS Palaiseau*)
- 10h45 Pause-café
- 11h00 « Paradoxes et inférence causale » par Alexandre Aussem (*Université Lyon 1*)
- 11h45 « BRICKS: Bayesian representation and inference for complex knowledge structuring » par Paul Munteanu (*Bayesia Laval*)
- 12h30 Pause
- 14h00 « A brief introduction to probabilistic argumentation » par Anthony Hunter (*University College London*)
- 14h45 « Autours des modèles graphiques qualitatifs » par Salem Benferhat (*Université d'Artois Lens*)
- 15h30 Pause-café
- 15h45 « Random matrix advances in machine learning » par R. Couillet (*Centrale-Supélec Gif-sur-Yvette*)
- 16h30 « The universality problem in dynamic machine learning » par Juan Pablo Ortega (*Université de Saint Gall*)
- 17h15 Clôture

## Organisation

Cet événement est organisé par Engelbert Mephu Nguifo ([engelbert.mephu\\_nguifo@uca.fr](mailto:engelbert.mephu_nguifo@uca.fr)) pour l'AFIA et Stéphane Chrétien ([stephane.chretien@npl.co.uk](mailto:stephane.chretien@npl.co.uk)) pour la SFdS.

## Inscriptions

L'inscription est gratuite mais obligatoire et inclut la participation aux pauses-café. Elle se fait en ligne : [https://docs.google.com/forms/d/e/1FAIpQLSf7osDlmfalZMTBXLnfwY0uAUbMJb0bgymvqTr\\_puD\\_ACrXMQ/viewform?c=0&w=1](https://docs.google.com/forms/d/e/1FAIpQLSf7osDlmfalZMTBXLnfwY0uAUbMJb0bgymvqTr_puD_ACrXMQ/viewform?c=0&w=1) L'organisation et la prise en charge du déjeuner restent à charge des participants.





**AFIA**  
Association française  
pour l'Intelligence Artificielle



# IHM&IA 2019

L'AFIHM et l'AFIA organisent leur quatrième journée commune INTERACTION HOMME-MACHINE & INTELLIGENCE ARTIFICIELLE.

Cette journée s'organise 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 14 mars 2019
- Lieu Amphi B310, Télécom ParisTech, 46 rue Barrault, 75013 Paris

## Programme

- 09h00 Accueil
- 09h30 Présentations de l'Association Française pour l'Intelligence Artificielle, par Yves Demazeau (*Président de l'AFIA*) et de l'AFIHM par Eric Lecolinet (*Président de l'AFIHM*).
- 10h00 « L'IA pour la conception et l'utilisation des systèmes interactifs » par Gilles Bailly (CNRS *Paris*)
- 10h30 « Prendre en compte le facteur humain pour construire la coopération Homme-IA » par Myriam Fréjus (EDF *Palaiseau*)
- 11h00 « Exploration interdisciplinaire de corpus d'interaction Humain-Humain et Humain-Machine pour le développement d'Agents Conversationnels Animés » par Magalie Ochs (Université *Aix-Marseille*)
- 11h30 « En quoi et comment l'IA pourrait-elle permettre d'améliorer l'entraînement et l'apprentissage des utilisateurs d'Interfaces Cerveau-Ordinateur ? » par Camille Jeunet (CNRS *Toulouse*)
- 12h00 « Apprentissage des Interfaces Multi-Touch : Reconnaissance de Gestes et Conception de Modèles d'Interaction Automatique » par Quentin Debard (INSA *Lyon*)
- 12h30 Déjeuner / Posters et Démonstrations
- 14h00 « Using Interactive Machine Learning in Movement-Based Interactive Systems » par Frédéric Bevilacqua (IRCAM *Paris*)
- 14h30 « Collaboration Humain-Machine à l'aide de motifs dialogiques » par Jean-Baptiste Louvet (INSA *Rouen*)
- 15h00 « FlexTech : Intégration de l'Humain dans les Systèmes Complexes à Autonomie Croissante » par Guy-André Boy (CentraleSupélec *Saclay*)
- 15h30 « Comment créer des Partenariats Humain-Machine ? » par Wendy Mackay (INRIA *Saclay*)
- 16h00 Pause / Posters et Démonstrations
- 17h00 Clôture

## Organisation

Cet événement est organisé par Nicolas Sabouret pour l'Association Française pour l'Intelligence Artificielle (AFIA) et Baptiste Caramiaux et Jules Françoise pour L'Association Francophone d'Interaction Homme-Machine (AFIHM).

## Inscriptions

Les inscriptions gratuites mais obligatoires, sont à réaliser sur <https://ihmia.afihm.org/inscriptions/>. L'inscription inclut la participation aux pause-café. L'organisation et la prise en charge du déjeuner restent à la discrétion des participants.





**AfIA**

Association française  
pour l'Intelligence Artificielle

**EGC**

## EGC&IA 2019

Nous assistons aujourd'hui à une production de données structurées sans précédent, publiées sous la forme de graphes de connaissances accessibles via le Web de données. Bien que incomplets, hétérogènes et pouvant contenir des erreurs, ces graphes de connaissance contiennent un gisement d'informations important qui peut être exploité pour découvrir automatiquement de nouvelles connaissances.

C'est ainsi que cette quatrième journée EXTRACTION ET GESTION DES CONNAISSANCES et INTELLIGENCE ARTIFICIELLE réunit les deux communautés autour du thème « DECOUVERTE DE CONNAISSANCES DANS LE WEB DES DONNEES ».

### Date et Lieu

- Date 10 mai 2019
- Lieu LRI, Bât. 660 Claude Shannon, rue Noetzelin, 91190 Gif-sur-Yvette

### Programme

- 09h45 Présentations de l'AFIA et du Collège Sciences de l'Ingénierie des Connaissances, par Yves Demazeau (*Président de l'AFIA*) et de EGC par Arnaud Martin (*Président de EGC*).
- 10h00 Conférence invitée « Explicabilité / Interprétabilité des algorithmes d'apprentissage profond » par Luis Galarraga (*INRIA Rennes*).
- 11h00 « Learning How to Correct a Knowledge Base » par Thomas Pellissier-Tanon (*U. Paris Saclay Paris*), Camille Bourgaux et Fabian Suchanek.
- 11h30 « Towards Interactive Causal Relation Discovery Driven by an Ontology » par Mélanie Munch (*INRA Paris*) Juliette Dibie, Pierre-Henri Wuillemin, et Cristina Manfredotti.
- 12h00 Pause-Déjeuner
- 13h30 « Enhancing the interlinking quality : A structure-based approach » par Pierre-Henri Paris (*Sorbonne Université Paris*), Fayçal Hamdi et Samira Si-Said Cherfi.
- 14h00 « Scalable Schema Discovery for RDF Data » par Redouane Bouhamoum (*U. Paris Saclay Paris*), Zoubida Kedad and Stéphane Lopes.
- 14h30 « Revealing the Conceptual Schemas of RDF Datasets » par Subhi Issa (*CNAM Paris*), Pierre-Henri Paris, Fayçal Hamdi et Samira Si-Said Cherfi
- 15h00 Pause-café
- 15h30 « Re-descriptions and Formal Concept Analysis for Mining Definitions in Linked Data » par Justine Reynaud (*U. Lorraine Nancy*) Yannick Toussaint et Amedeo Napoli.
- 16h00 « Linkex : A Tool for Link Key Discovery Based on Pattern Structure » par Nacira Abbas (*U. Lorraine Nancy*), Jérôme David et Amedeo Napoli.
- 16h30 Clôture

### Organisation

Cet événement est organisé par Nathalie Pernelle et Fathia Sais pour le Collège SIC de l'Association Française pour l'Intelligence Artificielle (AFIA) et Cyril de Runz pour l'Association Extraction et Gestion des Connaissances (EGC).

### Inscriptions

Les inscriptions gratuites mais obligatoires, se font sur <https://jtegcafia.sciencesconf.org/registration>. Elles incluent le déjeuner et la pause-café, offerts à tous les participants inscrits à la journée.



# APIA & CNIA & IC & JFPDA & RJCIA

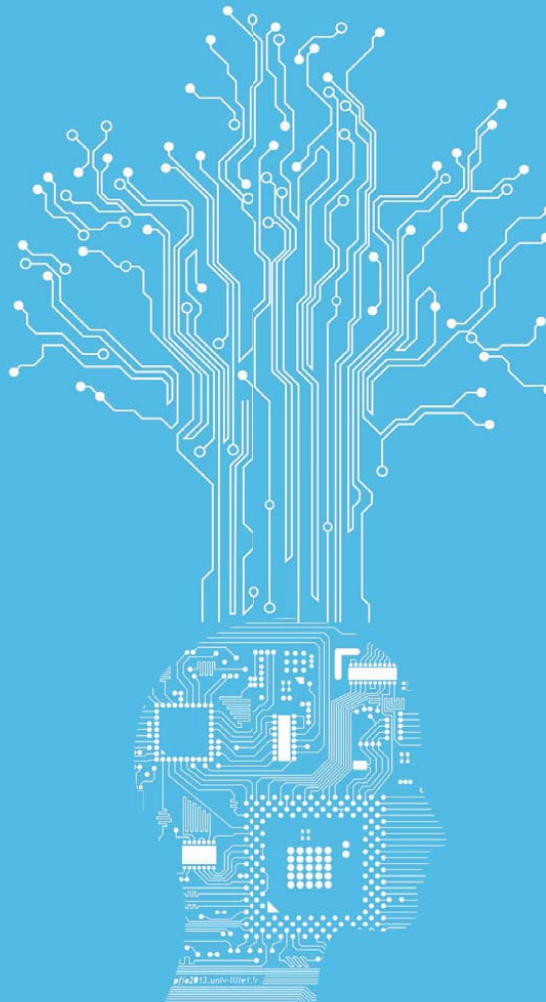
IA pour l'éducation   IA & Santé   TAL & IA   Ethique & IA  
Robotique & IA   France@IJCAI2018

# PFIA 2018

11<sup>e</sup> Plate-forme  
Intelligence Artificielle

2 au 6 juillet 2018 - Nancy

Campus Sciences - Université de Lorraine  
Vandœuvre-lès-Nancy



PFIA2018.LORIA.FR





**AFIA**  
Association française  
pour l'Intelligence Artificielle



Institut de Recherche  
en Informatique de Toulouse  
CNRS - INP - UT3 - UT1 - UT2J

La rencontre annuelle de la communauté  
**Intelligence Artificielle**  
autour de conférences et ateliers thématiques



Plate-Forme Intelligence Artificielle  
en coopération avec JSAI 2019

**TOULOUSE**  
1 - 5 Juillet 2019  
Université Toulouse 1 Capitole

**Conférences hébergées :**

- APIA : Conférence sur les Applications Pratiques de l'Intelligence Artificielle
- CAP : Conférence sur l'Apprentissage Automatique
- CNA : Conférence Nationale en Intelligence Artificielle
- IC : Journées Francophones Ingénierie des Connaissances
- JFPDA : Journées Planification, Décision et Apprentissage
- JFSMA : Journées Francophones sur les Systèmes Multi-Agents
- JIAF : Journées d'Intelligence Artificielle Fondamentale
- RJCA : Rencontre des Jeunes Chercheurs en Intelligence Artificielle
- TALN - RECITAL : Conférence sur le Traitement Automatique des Langues Naturelles

[www.irit.fr/pfia2019](http://www.irit.fr/pfia2019)

Sponsors



Soutiens institutionnels





En partenariat avec



L'AVENIR EST AUX VALEURS SÛRES

