

# Conférence Nationale d'Intelligence Artificielle Année 2017

## Sections Spéciales

Sandra Bringay  
Catherine Garbay  
Matthieu Geist  
Tiago de Lima  
Catherine Roussey  
Christian de Sainte-Marie  
Bruno Zanuttini

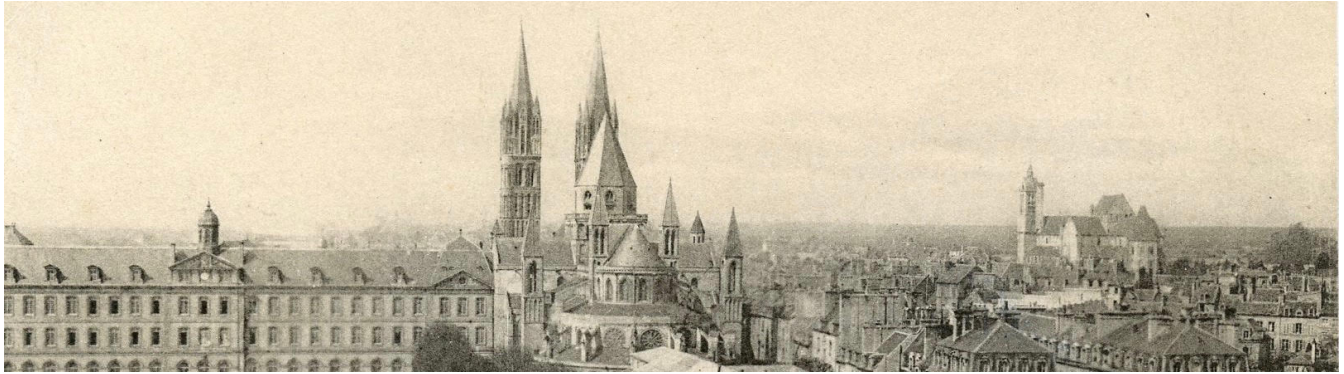
Coordination AFIA  
Carole Adam  
Yves Demazeau



**AFIA**  
Association française  
pour l'Intelligence Artificielle

---





# Conférence Nationale d'Intelligence Artificielle Année 2017

## Sections Spéciales

Sandra Bringay  
Catherine Garbay  
Matthieu Geist  
Tiago de Lima  
Catherine Roussey  
Christian de Sainte-Marie  
Bruno Zanuttini

Coordination AFIA  
Carole Adam  
Yves Demazeau



**AFIA**  
Association française  
pour l'Intelligence Artificielle

---



# **Actes CNIA 2017**

Conférence Nationale d'Intelligence Artificielle

Editeurs : Président(e)s de CP de CNIA 2017



# Table des matières

## Actes de la Conférence CNIA 2017

Flyer CNIA 2017. <b>Caen au coeur de l'IA de demain</b> .....	1
Comités 2017. <b>Comités de programmation, de programme, et d'organisation</b> .....	3
Fiona Berreby, Gauvain Bourgne et Jean-Gabriel Ganascia (IAF). <b>Cadre déclaratif modulaire pour représenter et appliquer des principes éthiques</b> .	9
Emmanuelle Claeys, Pierre Gançarski et Myriam Maumy-Bertrand (APIA). <b>Approche paramétrique de bandits pour l'A/B test</b> .....	21
Jean-Paul Delahaye et Philippe Mathieu (JFSMA). <b>Que valent les stratégies probabilistes au dilemme itéré des prisonniers?</b> .....	29
Guillaume Desquesnes, Guillaume Lozenguez, Arnaud Doniec et Eric Duviella (JFPDA). <b>Coordination distribuée et hors-ligne de planifications locales</b> .....	39
Arnaud Ferré (RJCIA). <b>Représentation de termes complexes dans un espace vectoriel relié à une ontologie pour une tâche de catégorisation</b> .....	47
Antoine Flepp, Fabrice Bourge, Julie Dugdale et Tiphaine Marie (RJCIA). <b>Extraction de connaissances à partir de l'usage des outils professionnels de communication et de collaboration</b> .....	55
Sébastien Frizzi, Rabeb Kaabi, Moez Bouchouicha, Jean-Marc Ginoux, Farhat Fnaiech et Eric Moreau (APIA). <b>Détection de la fumée et du feu par réseau de neurones convolutifs</b> .....	61
Maxime Guériau, Frédéric Armetta, Salima Hassas, Romain Billot et Nour-Eddin El Faouzi (JFSMA). <b>Contrôle par apprentissage constructiviste pour le trafic coopératif</b> .....	69
Amine Kacete, Renaud Séguier, Michel Collobert et Jérôme Royan (APIA). <b>Can Synthetic Data Handle Unconstrained Gaze Estimation?</b> .....	79
Dorian Kodelja, Romaric Besançon et Olivier Ferret (RJCIA). <b>Représentations et modèles en extraction d'événements supervisée</b> .....	87
Rémi Lacaze-Labadie, Domitile Lourdeaux et Mohamed Sallak (RJCIA). <b>Planification probabiliste : une heuristique pour garantir des solutions sûres</b> ....	95
Erwan Lecarpentier, Sebastian Rapp, Marc Melo et Emmanuel Rachelson (JFPDA). <b>Empirical evaluation of a Q-Learning Algorithm for Model-free Autonomous Soaring</b> .....	103
Domitile Lourdeaux, Azzeddine Benabbou, Lauriane Huguet et Rémi Lacaze-Labadie (APIA). <b>HUMANS : suite logicielle pour la scénarisation d'environnements virtuels pour la formation à des situations socio-techniques complexes</b> .....	111

Chun Lu, Philippe Laublet, Milan Stankovic et Filip Radulovic (IC). <b>Graphe de connaissances et folksonomie : leur performance comparative dans le calcul de l'affinité</b> .....	119
François Meunier, Christophe Marsala et Laurent Castanié (APIA). <b>3DRESC-TF : Apprentissage par transfert pour la réutilisation de connaissances en classification d'objets 3D</b> .....	125
Maxime Morge et Antoine Nongaillard (JFSMA). <b>Affectation distribuée d'individus à des activités</b> .....	133
Bilel Moulahi, Jérôme Azé et Sandra Bringay (IC). <b>Suivi et détection des idéations suicidaires dans les médias sociaux</b> .....	143
Amro Najjar, Olivier Boissier et Gauthier Picard (JFSMA). <b>Négociation «one-to-many» adaptative pour améliorer l'acceptabilité des services d'un fournisseur SaaS</b> .....	151
Joe Raad, Nathalie Pernelle et Fatiha Saïs (IC). <b>Détection de liens d'identité contextuels dans une base de connaissances</b> .....	161
Alexandre Schmitt, Florent Carlier, Valérie Renault et Pascal Leroux (RJCIA). <b>Communication multi-niveaux pour des IoT-a. Interactions autour d'un mur d'écrans connectés</b> .....	169
Anne-France Viet, Catherine Belloc et Régis Sabbadin (APIA). <b>Formalisation et résolution d'un problème en Santé Animale avec le cadre Leader-Follower MDP</b> .....	175



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

## Maroua Bouzid

Professeur à l'Université de Caen Normandie  
Présidente du comité d'organisation de la Plate-Forme IA 2017

## Yves Demazeau

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

L'AFIA et l'Université de Caen-Normandie organisent à Caen du 03 au 07 juillet 2017 la dixième 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) et Rennes (2015).

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

Pour son édition 2017, la plate-forme IA héberge les 6 conférences suivantes, ainsi que leurs ateliers associés :

<b>APIA</b>	3 <sup>ème</sup> Conférence Nationale sur les Applications Pratiques de l'IA
<b>IAF</b>	11 <sup>èmes</sup> Journées sur Intelligence Artificielle Fondamentale
<b>IC</b>	28 <sup>èmes</sup> Journées Francophones d'Ingénierie des Connaissances
<b>JFPDA</b>	12 <sup>ème</sup> Journées Francophone Planification, Décision et Apprentissage
<b>JFSMA</b>	25 <sup>ème</sup> Journée Francophones sur les Systèmes Multi-Agents
<b>RJCIA</b>	15 <sup>ème</sup> Rencontre des Jeunes Chercheurs en Intelligence Artificielle

Mais aussi les autres événements :

<b>E&amp;IA</b>	2 <sup>ème</sup> Journée Ethique et Intelligence Artificielle
<b>France@IJCAI</b>	Présentation par leurs auteurs français d'articles acceptés IJCAI 2017
<b>IA des Jeux</b>	1 <sup>ère</sup> Compétition d'Intégration Horizontale
<b>JI&amp;IA</b>	2 <sup>ème</sup> Journée Jeux Informatisés et Intelligence Artificielle

Six conférenciers ont également été invités pour des conférences plénières :

<b>Raúl García Castro</b>	Université Polytechnique de Madrid	Espagne
<b>Torben Bach Pedersen</b>	Université d'Aalborg	Danemark
<b>Michael Perrot</b>	Prix de Thèse IA 2017	France
<b>Eric Piette</b>	Prix de Thèse IA 2017	France
<b>Leon van der Torre</b>	Université du Luxembourg	Luxembourg
<b>Natalie van der Wal</b>	Université Libre d'Amsterdam	Pays-Bas

L'ensemble des informations sur ces événements est disponible à partir du site de la Plate-Forme IA 2017 (<https://pfia2017.greyc.fr/>) dont le contenu sera maintenu jusqu'à l'été 2021.

Nous remercions les présidents de comités de programme des conférences hébergées, les membres des comités de programme, les membres du comité d'organisation, les orateurs, l'ensemble des participants, pour leurs précieux apports à la réussite de cette plateforme. Nous remercions également les nombreux partenaires qui ont voulu nous accompagner.





**AFIA**

Association française  
pour l'Intelligence Artificielle

# Comité de programmation

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

**Président :** Yves Demazeau

- Carole Adam
- Olivier Boissier
- Sandra Bringay
- Catherine Faron-Zucker
- Frédéric Maris
- Engelbert Nguifo Mephu

# Comités de programme

## 3<sup>ème</sup> Conférence Applications Pratiques de l'Intelligence Artificielle (APIA)

**Président :** Sandra Bringay, LIRMM, Montpellier et Christian de Sainte-Marie, IBM France, Gentilly

- |                         |  |
|-------------------------|--|
| • Carole Adam           | LIG, Université Grenoble Alpes         |
| • Julian Alvarez        | Play Research Lab, Université de Lille |
| • Florence Amardeilh    | Mondeca, Université Paris 11           |
| • Alain Berger          | Ardans                                 |
| • Matthieu Boussard     | Craft AI                               |
| • Bertrand Braunschweig | Inria Saclay                           |
| • Nathalie Chaignaud    | LITIS, INSA Rouen                      |
| • François Cheriaux     | EDF Recherche et Développement         |
| • Caroline Chopinaud    | MASA Group                             |
| • Jan Ciger             | Reviattech SAS                         |
| • Philippe David        | SNCF                                   |
| • Jean-Marc David       | Renault                                |
| • Etienne de Sevin      | SANPSY, Université de Bordeaux         |
| • Yves Demazeau         | LIG, CNRS Grenoble                     |
| • Sylvie Despres        | LIM&BIO, Université Paris 13           |
| • Sébastien Destercke   | HEUDIASYC, CNRS Compiègne              |
| • Jean-Yves Donnart     | THALES                                 |
| • Béatrice Fuchs        | LIRIS, Université Lyon 3               |
| • Catherine Garbay      | LIG, CNRS, Grenoble                    |
| • Abdelkader Gouaïch    | LIRMM, Université de Montpellier       |
| • Christophe Guettier   | Safran Electronique et Défense         |
| • Hudelot Céline        | Ecole Centrale Paris                   |
| • Christine Largouët    | IRISA, Agrocampus Ouest                |
| • Dominique Lenne       | HEUDIASYC, UTC Compiègne               |
| • Domitile Lourdeaux    | HEUDIASYC, UTC Compiègne               |
| • Sylvain Mahe          | EDF Recherche et Développement         |

- Jean-Charles Marty LIRIS, Université Lyon 3
- Philippe Mathieu CRISTAL, Université Lille 1
- Nada Matta UTT Troyes
- Eunika Mercier-Laurent IAE, Université Lyon 3
- Philippe Morignot INRIA Rocquencourt
- Pascal Négros Université Paris I
- Selmin Nurcan Université Paris 1
- Jean-Marc Ogier L3i, Université de La Rochelle
- Francois Pachet CSL Sony Paris
- Romaric Redon AIRBUS
- Benoit Ricard EDF Recherche et Développement
- Jean Rohmer ESILV
- Dominique Romary ODEFI
- Catherine Tessier DCSD, ONERA Toulouse
- Erwan Tranvouez LSIS, Université d'Aix-Marseille
- Brigitte Trouse INRIA Rocquencourt
- Amel Yessad LIP 6, Université Paris 6

### **11<sup>èmes</sup> Journées d'Intelligence Artificielle Fondamentale (IAF)**

**Président :** Tiago de Lima, CRIL, Lens, et Bruno Zanuttini, GREYC, Caen

- Christian Bessière LIRMM, Montpellier
- Élise Bonzon LIPADE, Paris
- Tristan Cazenave LAMSADE, Paris
- Laurence Cholvy ONERA, Toulouse
- Nadia Creignou LIF, Marseille
- Sylvie Doutre IRIT, Toulouse
- Jérôme Euzenat LIG, Grenoble
- Sébastien Konieczny CRIL, Lens
- Fred Koriche CRIL, Lens
- Jérôme Lang LAMSADE, Paris
- Jean Lieber LORIA, Nancy
- Pierre Marquis CRIL, Lens
- Nicolas Maudet LIP6, Paris
- Marie-Laure Mugnier LIRMM, Montpellier
- Amedeo Napoli LORIA, Nancy
- Odile Papini LSIS, Marseille
- Laurent Perrussel IRIT, Toulouse
- Henri Prade IRIT, Toulouse
- Pierre Siegel LIF, Marseille
- Igor Stéphane LERIA, Angers
- Serena Villata I3S, Sophia Antipolis
- Christel Vrain LIFO, Orléans

### **28<sup>èmes</sup> Journées Francophones Ingénierie des Connaissances (IC)**

**Président :** Catherine Roussey, TSCF, Irstea Clermont Ferrand

- Marie-Hélène Abel HEUDIASYC, UTC Compiègne
- Xavier Aimé Cogsonomy Nantes / LIMICS Paris
- Yamine Ait Aneur IRIT, INP Toulouse
- Florence Amardeilh Mondeca, Paris
- Fabien Amarger IRIT, Toulouse
- Nathalie Aussenac-Gilles IRIT, CNRS, Toulouse



# AFIA

Association française  
pour l'Intelligence Artificielle

- Bruno Bachimont
- Nacéra Bennacer
- Aurélien Benel
- Pierre Bourhis
- Nathalie Bricon-Souf
- Sandra Bringay
- Patrice Buche
- Davide Buscaldi
- Elena Cabrio
- Sylvie Calabretto
- Gaoussou Camara
- Pierre-Antoine Champin
- Jean-Pierre Chanet
- Jean Charlet
- Olivier Corby
- Amélie Cordier
- Mathieu d'Aquin
- Luc Damas
- Jérôme David
- Sylvie Després
- Rim Djedidi
- Jean-Pierre Evain
- Gilles Falquet
- Catherine Faron-Zucker
- Cécile Favre
- Béatrice Fuchs
- Frédéric Furst
- Fabien Gandon
- Jean-Gabriel Ganascia
- Catherine Garbay
- Serge Garlatti
- Alain Giboin
- Nathalie Guin
- Ollivier Haemmerlé
- Mounira Harzallah
- Nathalie Hernandez
- Liliana Ibanescu
- Antoine Isaac
- Clément Jonquet
- Mouna Kamel
- Gilles Kassel
- Pascale Kuntz
- Florence Le Ber
- Michel Leclere
- Maxime Lefrançois
- Alain Léger
- Dominique Lenne
- Moussa Lo
- Cédric Lopez
- Nada Matta
- Pascal Molli
- COSTECH, UTC Compiègne
- LRI, Centrale Supélec Gif-sur-Yvette
- ICD, UTT Troyes
- LIFL, CNRS Lille
- IRIT, Université Toulouse 3
- LIRMM, Université Montpellier 3
- IATE, INRA Montpellier
- LIPN, Université Paris 13
- I3S, Université de Nice Sophia Antipolis
- LIRIS, INSA de Lyon
- Université Alioune Diop de Bambey, Sénégal
- LIRIS, Université Lyon 1
- TSCF, Irstea Clermont-Ferrand
- LIMICS, AP-HP/INSERM, Paris
- I3S, INRIA Sophia Antipolis-Méditerranée
- Hoomano, Lyon
- Knowledge Media Institute, Open University, RU
- LISTIC, Université de Savoie
- LIG, INRIA Grenoble
- LIMICS, Université Paris 13
- LIMICS, Université Paris 13
- EBU, Suisse
- Université de Genève, Suisse
- I3S, Université de Nice Sophia Antipolis
- ERIC, Université Lyon 2
- LIRIS, Université Lyon 3
- MIS, Université de Picardie
- I3S, INRIA Sophia Antipolis Méditerranée
- LIP6, Université Paris 6
- LIG, CNRS, Grenoble
- IMT Atlantique, Brest
- I3S, INRIA Sophia Antipolis
- LIRIS, Université Lyon 1
- IRIT, Université Toulouse 3
- LS2N, Université de Nantes
- IRIT, Université Toulouse 3
- MIA, INRA AgroParistech, Paris
- Europeana & Vrije University Amsterdam, Pays-Bas
- LIRMM, Université de Montpellier
- IRIT, Université de Perpignan
- MIS, Université de Picardie
- LS2N, Université de Nantes
- ICUBE, ENGEES, Strasbourg
- LIRMM, Université de Montpellier
- LHC, Ecole des Mines de Saint-Etienne
- Orange Labs Telecom, Rennes
- HEUDIASYC, UTC Compiègne
- Université de Saint Louis, Sénégal
- Viséo Objet Direct, Grenoble
- ICD, UTT Troyes
- LS2N, Université de Nantes

- Alexandre Monnin I3S, INRIA Sophia Antipolis Méditerranée
- Fleur Mougin INSERM BPH, Université de Bordeaux
- Amedeo Napoli LORIA, CNRS, Nancy
- Emmanuel Nauer LORIA, Université de Lorraine
- Jérôme Nobecourt LIMICS, Université Paris 13
- Nathalie Pernelle LRI, Université Paris Sud
- Camille Pradel Synapse, Toulouse
- Yannick Prié LS2N, Université de Nantes
- Cédric Pruski LIST, Luxembourg
- Sylvie Ranwez LGI2P, Ecole des mines d'Alès
- Chantal Reynaud LRI, Université Paris Sud
- Fatiha Sais LRI, Université Paris Sud
- Pascal Salembier ICD, UTT Troyes
- Karim Sehaba LIRIS, Université Lyon 2
- Hassina Seridi- Bouchelaghem LabGED, Université de Badji Mokhtar, Algérie
- Andrea Tettamanzi I3S, Université Nice Sophia Antipolis
- Raphaël Trpncy EURECOM, Sophia Antipolis
- Serena Villata I3S, CNRS
- Amel Yessad LIP6, Université Paris 6
- Haïfa Zargayouna LIPN, Université Paris 13
- Pierre Zweigenbaum LIMSI, CNRS, Orsay

## 12<sup>èmes</sup> Journées Planification, Décision et Apprentissage (JFPDA)

**Président :** Bruno Zanuttini, GREYC, Université de Caen, et Matthieu Geist, Centrale Supélec, Metz

- Olivier Buffet INRIA / LORIA, Nancy
- Martin Cooper IRIT, Toulouse
- Laurent Jeanpierre GREYC, Caen
- Frédéric Maris IRIT, Toulouse
- Marc Métivier LIPADE, Paris Descartes
- Damien Pellier IMAG, Grenoble
- Bilal Piot INRIA / Université Lille 1
- Philippe Preux INRIA / Université Lille 3
- Vincent Thomas LORIA, Nancy
- Thierry Vidal ENIT, Tarbes
- Paul Weng SYSU-CMU, Guangzhou, Chine

## 25<sup>èmes</sup> Journées Systèmes Multi-Agents (JFSMA)

**Président :** Catherine Garbay, LIG, CNRS

- Nadia Abchiche-Mimouni IBISC, Université Evry Val d'Essonne
- Emmanuel Adam LAMIH, Université Valenciennes
- Frédéric Amblard IRIT, Université Toulouse
- Frédéric Armetta LIRIS, Université Lyon
- Flavien Balbo LHC, Université Saint-Etienne
- Jean-Paul Barthès HEUDIASYC, UTC Compiègne
- Olivier Boissier LHC, Université Saint-Etienne
- Grégory Bonnet GREYC, Université Caen Normandie
- Maroua Bouzid GREYC, Université Caen Normandie
- Philippe Caillou LRI, Université Paris Sud
- Valérie Camps IRIT, Université Toulouse
- Vincent Chevrier LORIA, Université Lorraine
- Caroline Chopinaud Craft AI



# AFIA

Association française  
pour l'Intelligence Artificielle

- Céline Clavel
- Vincent Corruble
- Pierre De Loor
- Yves Demazeau
- Alexis Drogoul
- Raphaël Duboz
- Cécile Duchene
- Julie Dugdale
- Amal El Fallah Seghrouchni
- Stéphane Galland
- Marie-Pierre Gleizes
- Abdelkader Gouaïch
- Emmanuelle Grislin
- Zahia Guessoum
- Guillaume Hutzler
- Jean-Paul Jamont
- Jean-Daniel Kant
- Domitile Lourdeaux
- René Mandiau
- Nicolas Marilleau
- Philippe Mathieu
- Laetitia Matignon
- Annabelle Mercier
- Bruno Mermet
- Frédéric Migeon
- Maxime Morge
- Gildas Morvan
- Jean-Pierre Müller
- Michel Ocello
- Denis Payet
- Sylvie Pesty
- Gauthier Picard
- Sébastien Picault
- Suzanne Pinson
- Lilia Rejeb
- Nicolas Sabouret
- Julien Saunier
- Olivier Simonin
- Erwan Tranvouez
- Mahdi Zargayouna
- LIMSI, Université Paris Sud
- LIP6, Université Paris Sorbonne
- Lab-STICC, Université Bretagne
- LIG, CNRS, Université Grenoble Alpes
- UMMISCO, IRD et Université Paris Sorbonne, Vietnam
- CIRAD & Institut Pasteur du Cambodge, Cambodge
- COGIT, IGN
- LIG, Université Grenoble Alpes
- LIP6, Université Paris Sorbonne
- IRTES-SET, Université Franche-Comté
- IRIT, Université Toulouse
- LIRMM, Université Montpellier
- LAMIH, Université Valenciennes
- LIP6 et CReSTIC, Université Reims
- IBISC, Université Evry Val d'Essonne
- LCIS, Université Grenoble Alpes)
- LIP6, Université Paris Sorbonne
- HEUDIASYC, UTC Compiègne
- LAMIH, Université Valenciennes
- UMI UMMISCO, IRD et Université Paris Sorbonne
- CRISTAL, Université Lille
- LIRIS, Université Lyon
- LCIS, Université Grenoble Alpes
- GREYC, Université Caen Normandie
- IRIT, Université Toulouse
- CRISTAL, Université Lille
- LG2IA, Université Artois
- UR Green, CIRAD Montpellier
- LCIS, Université Grenoble Alpes
- LIM, Université de la Réunion
- LIG, Université Grenoble Alpes
- LHC, Université Saint Etienne
- CRISTAL, Université Lille
- LAMSADE, Université Paris Dauphine
- SMART Lab, Institut Supérieur de Gestion, Tunisie
- LIMSI, Université Paris-Sud
- LITIS, Université Rouen Normandie
- CITI, Université Lyon
- LSIS, Université Aix Marseille
- GRETTIA, IFSTTAR

## 15<sup>ème</sup> Rencontre des Jeunes Chercheurs en Intelligence Artificielle (RJCIA)

**Président :** Matthieu Geist, CentraleSupélec, Metz

- Nicolas Béchet
- Grégory Bonnet
- Sylvain Bouveret
- Elena Cabrio
- Université de Bretagne Sud
- Université de Caen-Basse-Normandie
- LIG, Grenoble-INP
- Université de Nice Sophia Antipolis

- Jérémy Fix Centrale Supélec
- Matthieu Geist Centrale Supélec
- Abir Beatrice Karami Mines de Douai
- Jean-Marie Lagniez Centre de Recherche en Informatique de Lens
- Robin Lamarche-Perrin CNRS
- Fabien Lauer Université de Lorraine
- Marie Lefevre Université Claude Bernard Lyon 1
- Adrien Maillard ONERA
- Jean-Guy Mailly Université Paris Descartes
- Laetitia Matignon Université Claude Bernard Lyon 1
- Alexandre Niveau Université de Caen-Basse-Normandie
- Antoine Nongaillard Université de Lille 1
- Magalie Ochs Université Aix-Marseille
- Wassila Ouerdane Centrale Supélec
- Anastasia Paparrizou LIRMM
- Alexandre Pauchet INSA Rouen
- Frédéric Pennerath Centrale Supélec
- Bilal Piot Google DeepMind
- Regis Riveret DATA61- CSIRO
- Mathieu Roche CIRAD
- Francois Schwarzentruher ENS de Rennes
- Karima Sedki Université Paris 13
- Tim Van de Cruys CNRS
- Serena Villata CNRS

## Comité d'organisation

**Equipe MAD du laboratoire GREYC (UNICAEN, CNRS UMR 6072, ENSICAEN)**

**Président :** Maroua Bouzid

- Florian Bénavent
- Grégory Bonnet
- François Bourdon
- Jonathan Cohen
- Quentin Cohen-Sohal
- Guillaume Debras
- Virginie Desnos-Carreau
- Fabio Valerio Ferrari
- Laurent Jeanpierre
- Mathieu Lelerre
- Christopher Leturc
- Bruno Mermet
- Abdel-Allah Mouaddib
- Xuan Son Nguyen
- Alexandre Niveau
- Gaële Simon
- Loïs Vanhee
- Bruno Zanuttini



# Cadre Déclaratif Modulaire pour Représenter et Appliquer des Principes Éthiques

F. Berreby<sup>1</sup>G. Bourgne<sup>1</sup>J-G. Ganascia<sup>1</sup>

<sup>1</sup> CNRS & Sorbonne Universités, UPMC Université Paris 06, LIP6 UMR 7606,  
4, place Jussieu 75005 Paris, France

fiona.berreby@lip6.fr

## Résumé

*Cet article examine l'utilisation de langages de haut niveau dans la conception d'agents autonomes éthiques. Il propose un cadre logique nouveau et modulaire pour représenter et raisonner sur une variété de théories éthiques, sur la base d'une version modifiée de l'Event Calculus, implémentée en Answer Set Programming. Le processus de prise de décision éthique est conçu comme une procédure en plusieurs étapes, capturée par quatre types de modèles interdépendants qui permettent à l'agent d'évaluer son environnement, de raisonner sur sa responsabilité et de faire des choix éthiquement informés. Notre ambition est double. Tout d'abord, elle est de permettre la représentation systématique d'un nombre illimité de processus de raisonnements éthiques, à travers un cadre adaptable et extensible en vertu de sa hiérarchisation et de sa syntaxe standardisée. Deuxièmement, elle est d'éviter l'écueil de nombreux travaux d'éthique computationnelle qui intègrent directement l'information morale dans l'engin de raisonnement général sans l'explicitement -alimentant ainsi les agents avec des réponses atomiques qui ne représentent pas la dynamique sous-jacente. Nous visons à déplacer de manière globale le fardeau du raisonnement moral du programmeur vers le programme lui-même.*

## Mots Clef

Éthique Computationnelle; Answer Set Programming; Event Calculus

## Abstract

*This paper investigates the use of high-level action languages for designing ethical autonomous agents. It proposes a novel and modular logic-based framework for representing and reasoning over a variety of ethical theories, based on a modified version of the Event Calculus and implemented in Answer Set Programming. The ethical decision-making process is conceived of as a multi-step procedure captured by four types of interdependent models which allow the agent to assess its environment, reason over its accountability and make ethically informed choices. The overarching ambition of the presented*

*research is twofold. First, to allow the systematic representation of an unbounded number of ethical reasoning processes, through a framework that is adaptable and extensible by virtue of its designed hierarchisation and standard syntax. Second, to avoid the pitfall of much research in current computational ethics that too readily embed moral information within computational engines, thereby feeding agents with atomic answers that fail to truly represent underlying dynamics. We aim instead to comprehensively displace the burden of moral reasoning from the programmer to the program itself.*

## Keywords

Computational Ethics; Answer Set Programming; Event Calculus

## 1 Introduction

L'étude de la morale d'un point de vue computationnel a attiré l'intérêt croissant de chercheurs en intelligence artificielle[2]. En effet, l'autonomie grandissante des agents artificiels et l'augmentation du nombre de tâches qui leur sont déléguées nous incitent à aborder leur capacité à traiter les restrictions et les préférences éthiques, que ce soit dans leur propre structure interne ou pour des interactions avec des utilisateurs humains. Des domaines aussi variés que la santé ou le transport posent des problèmes éthiques qui sont en ce sens particulièrement pressants, car ils peuvent exiger des agents des prises de décisions dont les conséquences sont immédiates ou lourdes. L'éthique computationnelle peut aussi nous aider à mieux comprendre la morale et raisonner plus clairement sur les concepts éthiques qui sont employés dans des domaines philosophiques, juridiques et technologiques. Dans ce contexte, notre objectif est de fournir une architecture modulaire qui permette la représentation systématique et adaptable des principes éthiques. Pour ce faire, nous présentons un ensemble cohérent de modèles qui, ensemble, permettent à l'agent d'évaluer son environnement, d'intégrer des règles éthiques et déterminer à partir de la mise en œuvre de ces règles soit un plan d'action, soit une évaluation du comportement d'autres agents. Ceux-ci sont

implémentés en Answer Set Programming<sup>1</sup>, sur la base d'une version modifiée de l'Event Calculus. Ainsi, notre approche est une approche logique de l'éthique, qui existe en parallèle à d'autres telles que l'éthique par étude de cas ([3][2]), ou l'éthique par conception ([5]).

Nous avons choisi l'utilisation de la *logique non monotone* car son étude a été proposée comme moyen de gérer le genre de généralisations révocables qui caractérisent souvent le raisonnement de sens commun et qui sont mal capturées par les systèmes logiques classiques [16]. Le terme couvre une famille de cadres formels conçus pour appréhender le type d'inférence où aucune conclusion n'est définitive, mais reste ouverte à la modification à la lumière d'informations complémentaires. Ce type de raisonnement par défaut est constitutif du raisonnement éthique. Des facteurs tels que la présence d'options alternatives, de conséquences indirectes ou de circonstances atténuantes peuvent renverser le jugement éthique d'une action. En conséquence, les langages non monotones sont particulièrement adaptés à la modélisation du raisonnement éthique.

L'article est structuré comme suit. Nous commençons par présenter les concepts philosophiques pertinents ainsi que les travaux connexes [Sect.2], puis nous présentons l'architecture du cadre [Sect.3]. Ensuite, nous définissons et discutons de chaque modèle [Sects.4-7], puis illustrons leur implémentation à l'aide d'un exemple jouet [Sect.8], et concluons [Sect.9].

## 2 Motivation

### 2.1 Théories Éthiques

L'étude de l'éthique est l'étude des croyances que les gens peuvent ou devraient avoir pour contrôler leur comportement. Une classification tripartite standard divise le champ entre la *méta-éthique*, qui concerne le statut ontologique des concepts éthiques, l'*éthique appliquée*, qui concerne l'application des règles morales à des environnements particuliers, et l'*éthique normative*, qui traite de la définition, de la comparaison et de l'explication de conceptions éthiques [10]. Le présent travail présente un intérêt pour l'éthique appliquée en ce sens qu'il présente un schéma de conception d'agents artificiels contraints par l'éthique qui peuvent agir dans une variété de domaines appliqués. Il présente aussi un intérêt pour l'éthique normative car son but est de modéliser les processus qui sous-tendent la prise de décision éthique normative, avec la possibilité de confronter des perspectives différentes. Il se concentre sur deux de ses principales branches : l'éthique conséquentialiste et l'éthique déontologique.

**Le Bien et Le Juste** Les éthiques conséquentialistes s'articulent autour de l'idée que les actions doivent être évaluées en fonction de leur conséquences, et ne

peuvent être justes ou injustes qu'en vertu de ce qu'elles produisent. Une action moralement juste est celle qui produit un bon, ou le meilleur, état de choses. Hors, afin de déterminer la justesse d'une action, les conséquentialistes doivent d'abord établir ce qui constitue un *bon* état de choses, c'est-à-dire déterminer ce qu'on appelle plus largement 'Le Bien' [1]. Cela leur permet ensuite d'affirmer que des actions font partie du 'Juste' dans la mesure où elles augmentent le Bien. Les théories conséquentialistes suivent donc, s'appuient et finalement dépassent, les théories du Bien. Les désaccords entre conséquentialistes sur ce qui constitue le Bien, ont engendré diverses traditions et doctrines conséquentialistes. L'*utilitarisme* voit le Bien résulter de la maximisation du bien-être collectif, l'*altruisme éthique* du bien-être des autres, l'*égoïsme éthique* de l'intérêt personnel, l'*utilitarisme des droits* du respect des droits individuels.

Les théories déontologiques (du grec *deon*, "devoir") prétendent que la valeur morale d'une action est déterminée (au moins en partie) par une caractéristique intrinsèque de l'action. Habituellement, cette caractéristique est une obligation ou une interdiction. Par exemple, une règle déontologique peut indiquer que le mensonge est contraire à l'éthique, ce qui implique que tout énoncé qui contient un mensonge est interdit. Parce qu'une action est jugée juste ou non en fonction de sa conformité avec une norme ou un devoir, son évaluation éthique est au moins en partie indépendante de ses conséquences. Le Juste est ici prioritaire sur le Bien : une action peut être injuste pour le déontologue même si elle maximise le Bien, et juste même si elle le minimise. Les tentatives de définition du Bien seront désormais appelées *théories du Bien*, et les tentatives de définition du Juste, qu'elles soient conséquentialistes ou déontologiques, seront appelées *théories du Juste*.

### 2.2 Travaux Connexes

Un certain nombre de travaux ont proposé des modèles informatiques de théories éthiques, dont la déontologie basée sur les devoirs et les règles [3] [5] [26], la déontologie de commande divine [10], le conséquentialisme [15] [19], ou l'instanciation de normes [29]. Il existe aussi des approches de vérification formelle [12][13]. Cependant, certains de ces modèles tendent à intégrer directement l'information éthiques au sein du processus de prise de décision de l'agent, sans pour autant générer un raisonnement moral. Bien qu'ils réussissent à exécuter des implémentations directes de restrictions uniques, ils ne fournissent pas une représentation explicite des relations de causalité ou des processus de pensée éthique, limitant ainsi leur applicabilité et leur portée. Par exemple, en utilisant une logique prospective, Pereira et al. [26] modélisent une règle déontologique qui prohibe le meurtre intentionnel par la règle '*falsum* ← *intentionalKilling*.' Or ils déterminent si '*intentionalKilling*' vaut pour une action en indiquant atomiquement si cette action l'implique, utilisant des

1. Pour une description de l'Answer Set Programming, voir [20].

règles de la forme '*intentionalKilling*  $\leftarrow$  *end(A,iKill(Y))*.' où A est l'action évaluée. Le problème avec cette approche est que l'évaluation éthique est *indiquée* par des énoncés spécifiques à l'action, plutôt qu'*extraite* par une forme de compréhension de l'environnement et des règles éthiques en place. Il n'y a pas de représentation de la causalité, de sorte que l'action et ses conséquences ne sont pas liées dynamiquement; leur relation est déclarée plutôt que déduite. Par conséquent, aucune notion de responsabilité éthique ne peut être élaborée sur cette base. En outre, les règles données manquant de puissance expressive, un nouveau programme est nécessaire pour modéliser chaque nouveau cas. Ces formalismes ne peuvent donc pas contraster différentes théories, ni expliciter leurs hypothèses.

Des travaux plus récents ont exploré de manière intégrale l'architecture des jugements éthiques [11] [9], représentant explicitement ces processus de raisonnement. Ce travail s'inscrit dans cette poursuite.

### 3 Schéma Structurel

#### 3.1 Modèles et Modularité

La représentation explicite du raisonnement éthique permet à un agent d'informer son processus de prise de décision ou de juger du comportement des autres. Pour y parvenir, il 'teste' les actions possibles dans des simulations spécifiques afin d'évaluer leurs conséquences ou leur mérite éthique inhérent. Le résultat de la simulation donne alors un ensemble d'actions acceptables ou inacceptables, qui dicte le comportement à venir. Le cadre présenté ici est concerné par ce processus d'évaluation, plutôt que par ce que l'application de cette évaluation.

Le processus éthique est appréhendé comme une procédure en quatre étapes définie par quatre types de modèles interdépendants : un *modèle d'action*, un *modèle de causalité*, un *modèle du Bien*, et un *modèle du Juste*. Les deux premiers modèles produisent une compréhension entièrement non-éthique du monde, les deux suivants y superposent une compréhension éthique du monde. Le modèle d'action présenté ici, et qui constitue la base du cadre, est basé sur une version modifiée de l'Event Calculus à la manière de [9]. La situation est représentée en termes de *fluents*, des propriétés du monde variant dans le temps, et d'*événements* qui modifient ces fluents. Nous définissons ces modèles ici, comme illustré dans la figure 1.

Un *modèle d'action* A permet à l'agent de représenter son environnement et les changements qui s'y déroulent. Il prend comme entrée un *ensemble d'actions effectuées*. Il est composé d'une *situation initiale* contenant les fluents vrais à T=0, une *spécification d'événements* contenant un ensemble d'événements et de relations de dépendance, et un *moteur d'événement* qui permet à la simulation d'évoluer. Il génère une *trace d'événements* de chaque si-

mulation qui désigne pour chaque moment les événements qui s'y produisent et les fluents qui y sont vrais.

Un *modèle de causalité* C piste les conséquences des actions, rendant possible un raisonnement sur la responsabilité et l'imputabilité des agents. Il prend comme entrée la *trace d'événements* produite par le modèle d'action et une *spécification d'événements* contenant un ensemble d'événements et de relations de dépendance. Il est composé d'un *moteur causal* qui permet la création d'un arbre causal représentatif de la simulation. Il génère une *trace causale* de chaque simulation qui désigne pour chaque moment les liens de cause à effets qui existent entre événements et fluents.

Un *modèle du Bien* G donne une appréciation de la valeur éthique intrinsèque de finalités ou d'événements. Il est composé d'une *spécification de modalités*, d'une *spécification éthique d'événements* composée d'un ensemble d'événements et d'un ensemble de relations de dépendance éthique, et d'une ou plusieurs *théories du Bien*. Il génère une *évaluation du Bien*, évaluant les événements comme plus ou moins en accord avec le Bien.

Un *modèle du Juste* R détermine l'action la plus juste selon des circonstances données. Il prend comme entrée la *trace causale* produite par le modèle de causalité et, dans le cas où une théorie du Juste considérée contient des principes conséquentialistes, une *évaluation du Bien* produite par le modèle du Bien. Il est composé d'une ou plusieurs *théories du Juste* et, dans le cas où une théorie du Juste considérée contient des principes déontologiques, un *ensemble de spécifications déontologiques*. Il génère une *évaluation du Juste*, évaluant les actions comme plus ou moins en accord avec le Juste -et donc admissibles ou non.

Ces quatre types de modèles sont interdépendants à degrés variables. Les modèles du Bien et du Juste reposent toujours sur un modèle d'action et un modèle de causalité. Mais alors qu'un modèle de causalité est toujours nécessaire, la formulation particulière du moteur causal peut varier, par exemple pour représenter différentes définitions des causes et des conséquences. Parce que le moteur d'événement constitue la base du cadre, il est toutefois proposé qu'il soit unique et invariable. En ce qui concerne les modèles éthiques, avoir un modèle du Bien est nécessaire pour modéliser les théories du Juste qui sont conséquentialistes, ainsi que celles qui sont déontologiques lorsqu'elles comportent des contraintes conséquentialistes. Des interdépendances peuvent également exister au sein d'un type de modèle, en particulier dans le cas des théories du Juste qui peuvent faire appel l'une à l'autre. La hiérarchie bien définie entre les différents types de modèles donne au cadre la capacité non seulement de modéliser, mais aussi de comparer un nombre potentiellement illimité de théories éthiques. De plus, la compartimentation des différents

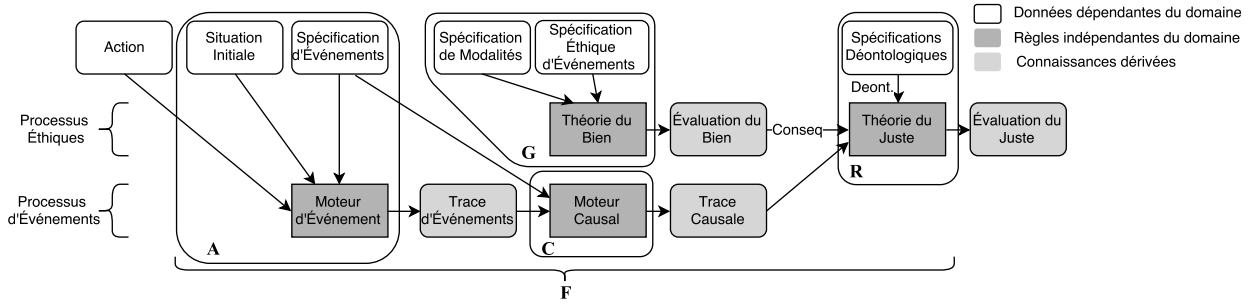


FIGURE 1 – Modèles et Modularité

types de processus permet leur analyse spécifique. Remplacer un modèle spécifique tout en maintenant les autres rend possible l'examen individualisé de ses ramifications. Sur la base de ces modèles, le *cadre d'évaluation éthique* est défini par :

$$F = \langle A_i, C_i, G_i, R_i \rangle$$

Étant donné un cadre d'évaluation éthique  $F$  et un ensemble  $\mathcal{A}$  d'actions exécutées  $\alpha$ , nous définissons l'ensemble des actions admissibles comme suit :

$$\begin{aligned} \text{Admissible}(F, \mathcal{A}) = \\ \{ \alpha \in \mathcal{A} / A_i, C_i, G_i, R_i \models \text{admissible}(\alpha) \} \end{aligned}$$

#### 4 Moteur d'évènement

**L'Event Calculus Adapté** Le moteur d'évènement présenté ici correspond à l'Event Calculus complet décrit dans [27], avec les ajouts suivants. Pour répondre aux particularités de la modélisation de scénarios éthiques complexes, nous introduisons des événements automatiques en plus des actions. Ces événements automatiques se produisent lorsque toutes leurs conditions préalables, sous la forme de fluents, tiennent, sans apport de l'agent. En outre, nous faisons une distinction entre les fluents inertiels, qui restent vrais jusqu'à ce qu'ils soient terminés par une occurrence d'évènement, et les fluents non inertiels qui sont vrais pour un moment seulement [23]. Enfin, nous introduisons un ensemble de simulations qui permettent à l'agent de simuler séparément et simultanément les effets de différentes actions sur le même scénario. Nous désignons l'ensemble des fonctions et des constantes comme suit :  $S$  est un ensemble de simulations,  $\mathcal{T}$  un ensemble de moments ;  $\mathcal{F}$  un ensemble de fluents,  $\mathcal{A}$  un ensemble d'actions,  $\mathcal{U}$  un ensemble d'évènements automatiques et  $\mathcal{E}$  un ensemble d'évènements où  $\mathcal{E} \equiv \mathcal{A} \cup \mathcal{U}$ .

**Axiomes d'Effets** Les prédicats suivants caractérisent le comportement des fluents qui contribuent à la création d'évènements.  $\text{initially}(F)$  indique que  $F$  est vrai initialement ;  $\text{effect}(E, F)$  indique que  $E$  peut causer  $F$  ;  $\text{initiates}(S, E, F, T)$  indique que  $E$  initie  $F$  à  $T$  dans  $S$  (et  $F$  n'est pas la négation d'un fluent) ;  $\text{terminates}(S, E, F, T)$  indique que  $E$  termine  $F$  à  $T$

dans  $S$  ;  $\text{clipped}(S, F, T)$  indique que  $F$  est stoppé à  $T$  dans  $S$  ;  $\text{nonInertial}(F)$  qualifie les fluents qui ne sont pas contraints par la loi d'inertie ;  $\text{holds}(S, F, T)$  indique que  $F$  est vrai à  $T$  dans  $S$ . Ces prédicats nous permettent d'axiomatiser les principes qui gouvernent les fluents : un fluent est vrai à  $T$  dans  $S$  si il a été initié par une occurrence d'évènement à  $T-1$  dans  $S$  ; un fluent vrai à  $T$  dans  $S$  reste vrai jusqu'à l'occurrence d'un évènement qui le termine, sauf s'il est non-inertiel, auquel cas il est vrai à  $T$  seulement.

```
holds(S, F, 0) :- initially(F), sim(S).
initiates(S, E, F, T) :-
    effect(E, F), occurs(S, E, T),
    not negative(S, F).
negative(S, neg(F)) :- effect(E, neg(F)), sim(S).
terminates(S, E, F, T) :-
    effect(E, neg(F)), occurs(S, E, T).
clipped(S, F, T) :- terminates(S, E, F, T).
holds(S, F, T) :-
    initiates(S, E, F, T-1), time(T).
holds(S, F, T) :-
    holds(S, F, T-1), not clipped(S, F, T-1),
    not nonInertial(F), time(T).
```

**Axiomes de Préconditions** Les prédicats suivants caractérisent le comportement d'évènements qui déterminent l'état des fluents.  $\text{prec}(F, E)$  indique que  $F$  est une précondition de  $E$  ;  $\text{incomplete}(S, E, T)$  indique que  $E$  est incomplet à  $T$  dans  $S$  ;  $\text{possible}(S, E, T)$  indique que  $E$  est possible à  $T$  dans  $S$  ;  $\text{occurs}(S, U, T)$  indique que  $U$  se produit à  $T$  dans  $S$  ;  $\text{occurs}(S, A, T)$  indique que  $A$  se produit à  $T$  dans  $S$ . Ces prédicats nous permettent d'axiomatiser les principes qui régissent l'occurrence des évènements : un évènement automatique se produit à  $T$  dans  $S$  si toutes ses préconditions sont vraies à  $T$  dans  $S$  ; une action se produit à  $T$  dans  $S$  si toutes ses préconditions sont vraies et qu'un agent effectue  $A$  à  $T$  dans  $S$ .

```
incomplete(S, E, T) :-
    prec(F, E), not holds(S, F, T),
    sim(S), time(T).
possible(S, E, T) :-
    not incomplete(S, E, T), sim(S),
    event(E), time(T).
occurs(S, U, T) :- possible(S, U, T), auto(U).
```

```
occurs(S, A, T) :-
  possible(S, A, T), performs(S, D, A, T),
  act(A).
```

## 5 Moteur Causal

**Axiomes de Causalité** En se basant sur l'architecture de l'Event Calculus, nous définissons la causalité en termes de conséquences. Cela nous permet de générer une trace fonctionnelle et dynamique des liens causaux. Nous la définissons de la manière suivante.

Un fluent F est une *conséquence* d'un événement E si E produit F, et les deux sont vrais. Un événement E est une *conséquence* d'un fluent F si F est une précondition à E, et les deux sont vrais.

Cette définition gère la possibilité qu'il y ait plus d'une précondition pour l'occurrence de E et que F ne soit pas considéré comme une cause de E si E ne se produit pas (par exemple car d'autres préconditions n'ont pas été remplies). Pour la modéliser, nous définissons le prédicat  $\text{cons}(S, E1, T, E2)$ , qui indique que l'événement E2 est une conséquence de l'événement E1 qui s'est produit dans S à T. Le moment référencé est le moment auquel s'est produit le *premier* événement de la chaîne causale. Une chaîne causale est composée d'une série de fluents et d'événements, mais le début et la fin d'une chaîne causale sont toujours des événements.

```
cons(S, E, T, F) :-
  occurs(S, E, T),
  effect(E, F), holds(S, F, T+1).
cons(S, F, T, E) :-
  occurs(S, E, T), prec(F, E), holds(S, F, T).
cons(S, E1, T1, E3) :-
  cons(S, E1, T1, C2), cons(S, C2, T2, E3),
  event(E1), event(E3), T2 > T1.
```

## 6 Théories du Bien

Dans cette section, nous présentons deux modes de définition du Bien, relatifs aux droits aux valeurs. Ces modes sont interchangeable et peuvent également être combinés. Nous présentons ensuite un modèle pour quantifier le bien une fois qu'il a été qualifié, ce qui lui permet à la fois d'être intégré dans les théories du Juste et donne aux événements des poids significatifs. Les droits, les valeurs ou d'autres moyens de définir le bien sont appelés *modalités*.

### 6.1 Qualifier le Bien

**Droits** L'*utilitarisme des droits* de Nozick postule que le Bien à maximiser consiste en la non violation des droits [24]. Un droit peut être défini comme une "*revendication justifiée que les individus et les groupes peuvent faire sur d'autres individus ou sur la société; avoir un droit c'est être en mesure de déterminer par ses choix, ce que les autres doivent ou ne doivent pas faire*" [7]. Cette définition capture le fait qu'un droit indique à la fois un état des

choses pour la personne concernée (l'exercice du droit) et une contrainte imposée aux autres (l'interdiction de violer le droit). Nous définissons les règles de sorte qu'un événement impliquant des personnes et qui viole un droit est *mal* par rapport à ce droit, et un événement qui implique des personnes mais ne viole pas un droit est *bien* par rapport à ce droit. Un événement peut être mal par rapport à un droit et bien par rapport à un autre. Cependant, aucun événement impliquant des personnes n'est neutre en ce qui concerne les droits : il respecte ou non chaque droit. Ce principe du tiers exclu est rendu explicite par l'utilisation de la négation par l'échec dans la règle. Notez que les droits sont définis par  $\text{right}(M)$  dans la *spécification de modalités*.

```
bad(E, X, M) :- effect(E, involves(X)),
  effect(E, neg(M)), right(M).
good(E, X, M) :- effect(E, involves(X)),
  not effect(E, neg(M)), right(M).
```

**Valeurs** Une théorie fondée sur les valeurs fournit également un moyen efficace d'évaluer le mérite initial des événements. Une valeur peut être définie comme "*une conception, explicite ou implicite, distinctive d'un individu, ou caractéristique d'un groupe, de ce qui est souhaitable et qui influence la sélection des modes, des moyens et des fins d'action disponibles.*" [18]. Une valeur est donc un type d'entité indépendante que peuvent assumer les actions et leurs conséquence. Les valeurs peuvent être générales, ou spécifiques à différents contextes, tels que le lieu de travail ou l'éducation des enfants. Nous définissons les règles de sorte qu'un événement qui démontre l'expression d'une valeur particulière est *bien* par rapport à cette valeur, et un événement qui démontre la négation d'une valeur est *mal* par rapport à cette valeur. Les autres événements sont considérés ni bons ni mauvais par rapport à celle-ci. Notez que les valeurs sont définies par  $\text{value}(M)$  dans la *spécification de modalités*.

```
good(E, X, M) :- effect(E, involves(X)),
  effect(E, displays(M)), value(M).
bad(E, X, M) :- effect(E, involves(X)),
  effect(E, neg(displays(M))), value(M).
```

### 6.2 Quantifier le Bien

Une fois déterminé le contenu du Bien, nous procédons à la quantification de ce contenu, c'est-à-dire à la 'pesée' des bonnes et mauvaises ramifications des événements. Nous définissons trois paramètres pour cela :

- Le nombre de personnes impliquées dans l'événement. Par exemple, un événement affectant cinq personnes comptera cinq fois plus qu'un événement en affectant une. Cette information est donnée dans les prédicats  $\text{good}(E, X, M)$  et  $\text{bad}(E, X, M)$  par X.
- La valeur relative des personnes impliquées dans l'événement. Par exemple, il est peut-être plus significatif de sauver des enfants plutôt que des

adultes, ou de nuire à des personnes en bonne santé plutôt qu'à des patients en déclin. Ce paramètre est pris en compte par l'attribution à chaque groupe d'une valeur numérique, exprimée dans le prédicat  $t\_Weight(E, G, N)$  où  $E$  est un événement,  $G$  son groupe cible et  $N$  le poids donné.

- L'importance de la modalité affectée par l'événement. Par exemple, être bienfaisant est peut-être plus important qu'être poli, respecter le droit à la vie est peut-être plus important que respecter le droit de propriété. Ce paramètre est pris en compte par l'attribution à chaque modalité d'une valeur numérique, exprimé dans le prédicat  $m\_Weight(M, N)$  où  $M$  est la modalité et  $N$  le poids donné.

L'attribution de poids aux modalités et aux groupes est non triviale, et la méthode proposée ici en est une parmi d'autres, qui fait acte d'introduction. Il est possible, par exemple, de l'enrichir en prenant en compte d'autres dépendances, telles que la corrélation entre certaines modalités et personnes (par exemple, l'autonomie pourrait être essentielle pour les adultes et la sécurité pour les enfants), ou l'importance de parties affectées non humaines (les animaux, l'environnement, etc.). L'étape suivante consiste à réunir tous les poids en un seul en agrégeant par un produit les différents paramètres dans les prédicats  $weightedGood(E, N, M)$  et  $weightedBad(E, N, M)$ .

Le poids global d'un événement correspond alors à la différence entre les sommes de toutes ses bonnes et mauvaises ramifications pondérées. Plus le poids d'un événement est grand, plus il participe au Bien; les événements qui ont des poids négatifs font plus de mal que de bien. Les poids sont donnés par le prédicat  $weight(E, N)$ . Ce prédicat permet l'intégration du Bien dans le Juste et participera à définir les théories du Juste. Il est à noter que les poids de cible et de modalité sont définis par  $t\_Weight(E, G, N)$  et  $m\_Weight(M, N)$  dans la *spécification de modalités*.

```
weightedGood(E, X*N1*N2, M) :- good(E, X, M),
    t_Weight(E, G, N1), m_Weight(M, N2).
weightedBad(E, X*N1*N2, M) :- bad(E, X, M),
    t_Weight(E, G, N1), m_Weight(M, N2).
weight(E, N1-N3) :-
    N1=#sum[weightedGood(E, N2, M1)=N2],
    N3=#sum[weightedBad(E, N4, M2)=N4],
    number(N1;N3), event(E).
```

## 7 Théories du Juste

### 7.1 Éthiques Conséquentialistes

L'éthique conséquentialiste existe sous de nombreuses formes, allant des principes simples d'action aux théories complexes pour maximiser le bien. Nous en décrivons et modélisons cinq.

**Proscription d'Actions Purement Préjudiciables** Le premier principe conséquentialiste affirme que les actions ayant des effets purement préjudiciables sont inadmissibles. Cette règle intuitive est pertinente pour la plupart des scénarios éthiques et peut compléter d'autres théories du Juste qui ne se concentrent que sur des actions à effets complexes. Pour mettre en œuvre cette règle, nous définissons les prédicats  $badCons(S, A, T)$  et  $goodCons(S, A, T)$ , qui indiquent respectivement qu'une action  $A$  qui se produit à  $T$  dans  $S$  provoque au moins une mauvaise ou une bonne conséquence. Nous déclarons ensuite qu'une action est inadmissible si elle n'a que des mauvaises conséquences. Toute autre action qui n'a pas été démontrée inadmissible est admissible.

```
badCons(S, A, T) :-
    act(A), cons(S, A, T, E), bad(E, X, M).
goodCons(S, A, T) :-
    act(A), cons(S, A, T, E), good(E, X, M).
imp(pureBad, A) :-
    badCons(S, A, T), not goodCons(S, A, T).
per(pureBad, A) :- act(A), not imp(pureBad, A).
```

**Principe de la Moins Mauvaise Conséquence** Également appelé *maximum minimorum*, ce principe déclare qu'une action est inadmissible si sa pire conséquence est pire que la pire conséquence de toute autre action possible. Ce principe est particulièrement pertinent pour la prise de décision sous incertitude, où, sous 'l'hypothèse de la malchance', chaque action possible produirait sa pire conséquence. Ainsi l'agent ferait mieux de choisir l'alternative ayant le moins mauvais mauvais résultat [21]. Pour formaliser cette règle, nous déterminons d'abord une hiérarchie entre les conséquences des actions, afin d'ensuite indiquer la pire. Le prédicat  $worse(E1, E2)$  indique que la conséquence  $E1$  d'une action est pire que la conséquence  $E2$  de la même ou d'une autre action. Les prédicats  $notWorstCons(S, A, T, E)$  et  $worstCons(S, A, T, E)$  déterminent alors la limite basse d'un ordre partiel déterminé par le prédicat  $worst$ . Enfin, nous déclarons qu'une action  $A1$  est inadmissible si sa pire conséquence  $E1$  est pire que la pire conséquence  $E2$  de toute autre action  $A2$ . Toutes les autres actions sont admissibles.

```
worse(E1, E2) :-
    cons(S1, A1, T1, E1), cons(S2, A2, T2, E2),
    weight(E1, N1), weight(E2, N2), N1 < N2.
notWorstCons(S, A, T, E1) :-
    act(A), cons(S, A, T, E1), cons(S, A, T, E2),
    worse(E2, E1), not worse(E1, E2).
worstCons(S, A, T, E) :-
    act(A), event(E), cons(S, A, T, E),
    not notWorstCons(S, A, T, E).
imp(leastBad, A1) :- worstCons(S1, A1, T1, E1),
    worstCons(S2, A2, T2, E2),
    worse(E1, E2), A1 != A2.
per(leastBad, A) :-
    act(A), not imp(leastBad, A).
```

**Principe d'Analyse Coût-Avantage** Ce principe indique qu'une action est admissible si elle est globalement bénéfique, c'est-à-dire si ses bonnes conséquences l'emportent les mauvaises. Nous utilisons les prédicats  $\text{weightCons}(S, A, T, E, N)$ , qui détermine les poids  $N$  des conséquences individuelles  $E$  d'une action  $A$  qui s'est produite à  $T$  dans  $S$ , et  $\text{weightAct}(A, N2)$ , qui concatène ces poids<sup>2</sup>, pour indiquer que  $A$  est inadmissible si ce poids total  $N2$  est négatif, et admissible sinon.

```
weightCons(S, A, T, E, N) :-
  act(A), cons(S, A, T, E), weight(E, N).
weightAct(A, N2) :- act(A), number(N2),
  N2 = #sum[weightCons(S, A, T, E, N1) = N1].
imp(benCosts, A) :-
  weightAct(A, N2), number(N2), N2 < 0.
per(benCosts, A) :-
  act(A), not imp(benCosts, A).
```

### Utilitarisme de l'Acte

*"C'est le plus grand bonheur du plus grand nombre qui est la mesure du bien et du mal." J. Bentham, 1776 [8].*

L'utilitarisme de l'acte exige que l'on évalue une action directement selon le *principe d'utilité*, qui stipule que l'action moralement correcte est celle qui a les meilleures conséquences globales pour le bien-être ou l'utilité de la majorité des parties concernées [8]. Une action est donc admissible si, compte tenu de toutes les autres actions disponibles, elle a les meilleures conséquences dans l'ensemble. En utilisant le prédicat  $\text{weightAct}$  défini ci-dessus, nous déterminons un ordre de préférence entre les actions dans le domaine et déclarons qu'une action  $A1$  est inadmissible s'il existe une autre action  $A2$  dont le poids est supérieur. Toute autre action est admissible.

```
imp(actUti, A1) :-
  weightAct(A1, N1),
  weightAct(A2, N2), N1 < N2.
per(actUti, A) :- act(A), not imp(actUti, A).
```

### Utilitarisme de la Règle

*"Chaque acte, dans la vie morale, tombe sous une règle; et nous devons juger la moralité ou l'immoralité de l'acte, non par ses conséquences, mais par les conséquences de son universalisation -c'est-à-dire par les conséquences de l'adoption de la règle sous laquelle cet acte tombe." J. Hospers, 1975 [28]*

L'utilitarisme de la règle évalue une action en deux temps. La première étape consiste à évaluer les règles morales sur la base du principe d'utilité : il s'agit de déterminer si une règle ou un ensemble de règles morales engendrera les meilleures conséquences, supposant que tout ou la plupart des agents s'y plient. Dans la vie quotidienne, de telles

2. Sans spécifier de situation car une seule action est réalisée dans chaque situation.

règles peuvent inclure 'Ne pas voler', ou 'Gardez ses promesses'. La deuxième étape consiste à évaluer les actions individuelles relativement à ce qui a été justifié au cours de la première étape. Une action n'est admissible que si la règle à laquelle elle appartient respecte le principe d'utilité, outre son propre respect du principe. Par exemple, si 'Ne pas voler' est une règle adoptée, le vol sera toujours inadmissible, même dans l'instance où un vol produirait la plus grande utilité (par exemple, en alimentant un affamé). Le prédicat  $\text{ruleCount}(R, N)$  regroupe tous les poids  $N$  des actions qui appartiennent à une règle  $R$ ; le prédicat  $\text{weightRule}(R, N)$  les concatène. Nous déclarons alors qu'une action  $A$  est inadmissible si elle est une instance d'une règle  $R$  globalement nuisible, c'est-à-dire dont les mauvaises conséquences l'emportent sur les bonnes, considérant toutes ses instanciation. Toute autre action est admissible. Il est à noter que les instances de règles et les règles sont définies par  $\text{rule}(R)$  et  $\text{instance}(A, R)$  dans la *spécification de modalités*.

```
ruleCount(R, N) :-
  rule(R), instance(A, R), weightAct(A, N).
weightRule(R, N) :-
  rule(R), number(N),
  N = #sum[ruleCount(R, N1) = N1].
imp(ruleUti, A) :-
  act(A), instance(A, R),
  weightRule(R, N), N < 0.
per(ruleUti, A) :- act(A), not imp(ruleUti, A).
```

## 7.2 Éthiques Déontologiques

Dans cette section, nous présentons trois doctrines déontologiques. Deux d'entre elles sont purement déontologiques, les codes de conduite et l'éthique kantienne. La doctrine du double effet comporte des contraintes conséquentialistes. Il est à noter que notre traduction logique de ces théories constitue une possibilité parmi d'autres.

### Codes de Conduite

*"J'apporterai mon aide à mes confrères ainsi qu'à leurs familles dans l'adversité. Que les hommes et mes confrères m'accordent leur estime si je suis fidèle à mes promesses; que je sois déshonoré et méprisé si j'y manque." Serment d'Hippocrate, [25]*

Un code de conduite est un ensemble de règles qui décrit les obligations, les interdictions ou les responsabilités d'un individu, d'un groupe ou d'une organisation. Il spécifie les principes qui guident la prise de décision ou les procédures de ceux qui sont contraints par le code. Les codes de conduite varient dans leur portée et leur nature, allant des codes déontologiques professionnels aux commandements religieux. Le comportement et la moralité sont typiquement déterminés par un corps global, tel qu'une entreprise, un état ou un dieu. Nous illustrons ici ce type de contrainte en modélisant une règle commune qu'est l'interdiction de tuer. Une telle règle se trouve par exemple dans la Déclaration de Genève de l'Association Médicale Mondiale sous

la forme de la déclaration «Je garderai le respect absolu de la vie humaine» [6], ou dans le Décalogue comme commandement ‘Tu ne tueras point’(Exode 20 : 1-21). Nous modélisons une règle de ce type en déclarant qu’une action est inadmissible dans la mesure où elle provoque ou consiste en ce qui est interdit, ici, tuer.

```
imp(conduct, A) :-
  act(A), cons(S, A, T, kill(N, G)).
per(conduct, A) :- act(A), not imp(conduct, A).
```

### La Formule de la Fin en Elle-Même

*“Agis de telle sorte que tu traites l’humanité aussi bien dans ta personne que dans la personne de tout autre toujours en même temps comme une fin, et jamais simplement comme un moyen.” I. Kant, 1785 [17]*

Cette formule est un élément d’éthique kantienne qui met l’accent sur la valeur intrinsèque de la vie humaine. C’est un impératif moral qui interdit d’utiliser les personnes comme moyen pour d’autres fins, les personnes étant des fins en elles-mêmes en vertu de leur nature d’êtres rationnels [17]. La formule contraste la valeur intrinsèque, qui est persistante et souveraine, avec la valeur instrumentale, qui dépend de ce qu’elle produit. Nous présentons le prédicat `aim(A, E)` qui indique que le but de l’action A est de provoquer l’événement E et utilisons le fluent `involves(X)` pour indiquer qu’au moins une personne est impliquée dans E. Nous déclarons alors qu’une action A est inadmissible si elle provoque un événement E qui implique au moins une personne, mais où E n’est pas un but de A. Toute autre action est admissible. Les buts sont définis par `aim(A, E)` dans la *spécification déontologique*.

```
imp(kant, A) :- act(A), cons(S, A, T, E),
  effect(E, involves(X)), not aim(A, E).
per(kant, A) :- act(A), not imp(kant, A).
```

### La Doctrine du Double Effet

*‘Rien n’empêche un acte d’avoir deux effets, dont un seul est voulu, tandis que l’autre est à côté de l’intention.’ T. Aquinas, 1485 [4]*

La doctrine du double effet est un ensemble de critères éthiques utilisés pour évaluer la permissibilité éthique d’une action qui a à la fois de bonnes et de mauvaises conséquences [14]. Elle dicte qu’une personne peut licitement exécuter une action en sachant qu’elle aura des bons et mauvais effets, à condition que : 1) L’action elle-même soit bonne ou moralement neutre ; 2) Le mauvais effet ne soit pas directement voulu ; 3) Le bon effet résulte de l’acte et non du mauvais effet ; 4) Le bon effet soit plus important ou égal au mauvais effet [22]. `imp(dde1, A)` proscrit une action si elle est intrinsèquement mauvaise, correspondant à la condition 1. `imp(dde2, A)` proscrit une action si elle provoque un mauvais effet qui conduit à un bon effet. Cette règle correspond aux conditions 2 et 3, car nous estimons

que l’utilisation d’un événement comme moyen pour arriver à un autre événement équivaut à vouloir le premier événement. `imp(dde3, A)` proscrit une action si son effet global est mauvais. Cela correspond à la condition 4 qui équivaut au Principe d’Analyse Coût-Avantage. Toutes les autres actions sont permises par la doctrine.

```
imp(dde1, A) :- act(A), bad(A, X, M).
imp(dde2, A) :-
  act(A), cons(S, A, T, E1), cons(S, E1, T2, E2),
  bad(E1, X1, M1), good(E2, X2, M2).
imp(dde3, A) :- imp(benefitsCosts, A).
per(dde, A) :- act(A), not imp(dde1, A),
  not imp(dde2, A), not imp(dde3, A).
```

**Discussion** La modélisation de ces théories souligne certaines de leurs caractéristiques. D’abord, nous distinguons deux types de théories du Juste : celles qui évaluent chaque action par rapport aux autres actions possibles, et celles qui évaluent chaque action de manière indépendante. Les principes *relatifs* comparent les actions et font un choix unique. Le *principe de la moins mauvaise conséquence*, l’*utilitarisme de l’acte* et l’*utilitarisme de la règle* sont de ce genre. Cependant, l’*utilitarisme de la règle* est particulier en ce sens que la permissibilité de *toutes* les actions est déterminée selon l’impact de chacune individuellement. Ainsi, toutes ou aucune des actions considérées sous une règle sont admissibles. À l’inverse, les principes *indépendants* évaluent chaque action en elle-même sans être affectés par les options disponibles. Ainsi, ils produisent des ensembles d’actions admissibles ou inadmissibles. Toutes les autres théories du Juste présentées ici sont de ce genre. Il est important de noter que différentes théories du Juste peuvent, et parfois doivent, se compléter. Par exemple, la *doctrine du double effet* ne dit rien sur les actions dont les effets sont purement mauvais et peut être utilement complétée par un principe conséquentialiste pour y remédier.

## 8 Preuve de Concept

Dans cette section, nous illustrons la manière dont chaque contrainte éthique décrite ci-dessus gère un dilemme éthique à l’aide d’un exemple mono-agent de prise de décision. C’est un exemple jouet qui n’a pas vocation à être réaliste, mais vise à montrer la diversité des évaluations éthiques possibles. Le code source complet est téléchargeable sur un service cloud<sup>3</sup>.

**Un Dilemme Médical** Un médecin (l’agent autonome) est en possession de trois différents traitements expérimentaux pour une maladie grave et handicapante. Chaque traitement a un taux de réussite différent.

- Pour 100 patients qui essaient le traitement Alpha, 15 sont guéris, 20 perdent leur vie et 65 restent inchangés.

- Pour 100 patients qui tentent le traitement Bêta, 30 sont guéris, 25 perdent leur vie et 45 restent inchangés.

3. [https://gitlab.lip6.fr/ganascia/AAMAS\\_2017](https://gitlab.lip6.fr/ganascia/AAMAS_2017)



- Pour 100 patients qui essaient le traitement Gamma, 50 sont guéris, 30 perdent leur vie et 20 restent inchangés. Cependant, sur les 50 patients guéris, 30 ne sont entièrement guéris que parce qu'ils ont également reçu une transplantation d'organe provenant de chacun des 30 morts. Sans la transplantation, ils auraient perdu la vie.

Le gain en termes de vies sauvées (c'est-à-dire les patients guéris moins les patients tués) pour chaque traitement est donc : **Alpha -5**; **Beta 5**; **Gamma 20**. Afin de choisir quel traitement est admissible, nous considérons que le médecin simule séparément le fait de donner chaque traitement à un groupe de 100 personnes. Il administre le traitement à T=0 dans trois simulations respectives. Le cas est représenté de la manière suivante.

```
act (give (Z)) :- treatment (Z) .
prec (exist (Z), give (Z)) :- act (give (Z)) .
performs (sim1, doctor, give (alpha), 0) .
performs (sim2, doctor, give (beta), 0) .
performs (sim3, doctor, give (gamma), 0) .
initially (exist (Z)) :- treatment (Z) .
auto (cure (X, G); kill (X, G); null (X, G)) :-
    people (X), group (G) .
testgroup (group (alpha; beta; gamma)) .
treatment (alpha; beta; gamma) .
effect (give (alpha),
    posImpact (15, group (alpha))) .
effect (give (alpha),
    negImpact (20, group (alpha))) .
effect (give (alpha),
    noImpact (65, group (alpha))) .
effect (give (beta),
    posImpact (30, group (beta))) .
effect (give (beta),
    negImpact (25, group (beta))) .
effect (give (beta),
    noImpact (45, group (beta))) .
effect (give (gamma),
    posImpact (20, group (gamma))) .
effect (give (gamma),
    negImpact (30, group (gamma))) .
effect (give (gamma),
    noImpact (20, group (gamma))) .
effect (kill (X, group (gamma)),
    posImpact (X, group (gamma))) :- people (X) .
prec (posImpact (X, G), cure (X, G)) :-
    auto (cure (X, G)) .
prec (negImpact (X, G), kill (X, G)) :-
    auto (kill (X, G)) .
prec (noImpact (X, G), null (X, G)) :-
    auto (null (X, G)) .
effect (cure (X, G), involves (X)) :-
    auto (cure (X, G)), X > 0 .
effect (null (X, G), involves (X)) :-
    auto (null (X, G)), X > 0 .
effect (kill (X, G), involves (X)) :-
    auto (kill (X, G)), X > 0 .
nonInertial (posImpact (X, G); negImpact (X, G);
    noImpact (X, G); involves (X)) :-
    people (X), testgroup (G) .
```

TABLE 1 – Evaluations Basées sur la Valeur ‘helpfulness’

	Alpha	Beta	Gamma
<b>pureBad</b>	Admis	Admis	Admis
<b>leastBad</b>	Admis	Inadmis	Inadmis
<b>benCosts</b>	Inadmis	Admis	Admis
<b>actUti</b>	Inadmis	Inadmis	Admis
<b>ruleUti</b>	Admis	Admis	Admis
<b>conduct</b>	Inadmis	Inadmis	Inadmis
<b>kant</b>	Inadmis	Inadmis	Inadmis
<b>dde</b>	Inadmis	Admis	Inadmis

Le docteur considère en outre que venir en aide constitue le Bien, que guérir est venir en aide, tuer est l'opposé de venir en aide et laisser la maladie inchangée est n'avoir aucun impact. Il considère également que la valeur *venir en aide* (helpfulness) a un poids de 1 (ceci est ici trivial car il n'y a qu'une seule modalité) et que la vie de tous les patients est équivalente. Il pense aussi que donner chacun de ces traitements pourrait être généralisé comme la règle ‘donner des remèdes incertains’. Enfin, son objectif en donnant des traitements est de guérir.

```
m_Weight (M, 1) :- modality (M) .
t_Weight (E, G, 1) :-
    testgroup (G), effect (E, involves (X)) .
modality (M) :- value (M) .
value (helpfulness) .
effect (cure (X, G), displays (helpfulness)) :-
    auto (cure (X, G)) .
effect (kill (X, G),
    neg (displays (helpfulness))) :-
    auto (kill (X, G)) .
rule (uncertainCures) .
instance (give (alpha; beta; gamma),
    uncertainCures) .
aim (give (Z), cure (X, group (Z))) :-
    treatment (Z), people (X) .
```

Les résultats de l'évaluation éthique sont résumés dans le tableau 1. Pour montrer comment le remplacement d'un module par un autre peut changer le processus d'évaluation, nous modélisons également un cas dans lequel le docteur fonde sa théorie du Bien non pas sur une *valeur* mais sur le respect du *droit* à la vie. Cette modalité est définie comme suit, et les résultats résumés dans le tableau 2.

```
modality (M) :- right (M) .
right (life) .
effect (kill (X, G), neg (life)) :-
    auto (kill (X, G)) .
```

## 9 Conclusion

Le cadre présenté ici adapte et s'appuie sur l'Event Calculus pour permettre la modélisation de théories éthiques et de scénarios dans lesquels les appliquer. Défini en programmation logique, il présente une méthode et une im-

TABLE 2 – Evaluations Basées sur le Droit ‘life’

	<b>Alpha</b>	<b>Beta</b>	<b>Gamma</b>
<b>pureBad</b>	Admis	Admis	Admis
<b>leastBad</b>	Admis	Inadmis	Inadmis
<b>benCosts</b>	Admis	Admis	Admis
<b>actUti</b>	Admis	Inadmis	Inadmis
<b>ruleUti</b>	Admis	Admis	Admis
<b>conduct</b>	Inadmis	Inadmis	Inadmis
<b>kant</b>	Inadmis	Inadmis	Inadmis
<b>dde</b>	Admis	Admis	Inadmis

plémentation de cette méthode. L’accent est mis sur la hiérarchie et la représentation explicite des processus de raisonnement qui déterminent la prise de décision éthique. Celles-ci permettent d’engendrer des règles avec un fort potentiel expressif qui confèrent aux agents la capacité de décider et d’expliquer leurs décisions, mais aussi de raisonner sur les actions d’autres agents. En outre, la confrontation de théories éthiques avec les contraintes logiques des langages de programmation éclaire ces théories, clarifiant leurs concepts, les relations qui les lient et les ambiguïtés potentielles qu’elles peuvent contenir. Nous envisageons un certain nombre d’avenues futures pour développer le cadre actuel. Tout d’abord, nous cherchons à modéliser l’intention, qui n’est pour l’instant traitée qu’implicitement, et de modéliser les désirs des agents. Cela permettra aux agents de gérer des scénarios plus complexes et plus réalistes. En outre, nous avons l’intention de permettre la formulation de plans éthiques d’actions dans lesquels plus d’une action peut être évaluée dans une simulation, en travaillant vers un véritable domaine de planification. Enfin, nous visons à intégrer le cadre dans un système multi-agents, afin d’exploiter plus pleinement son potentiel pour faciliter la coopération ou l’intelligence collective.

## Remerciements

Les auteurs remercient l’Agence Nationale de la Recherche (ANR) pour sa contribution financière sous la référence ANR-13-CORD-0006.

## Références

- [1] L. Alexander and M. Moore. Deontological ethics. In E. N. Zalta, editor, *The Stanford Encyclopedia of Philosophy*. 2016.
- [2] M. Anderson and S. Anderson. *Machine ethics*. Cambridge University Press, 2011.
- [3] M. Anderson, S. Anderson, and C. Armen. Medethex : a prototype medical ethics advisor. 2006.
- [4] T. Aquinas. *Summa theologica*. Xist Publishing, 2015.
- [5] R. Arkin. *Governing lethal behavior in autonomous robots*. CRC Press, 2009.
- [6] W. M. Association et al. Wma declaration of geneva. *International Journal of Person Centered Medicine*, 4(3), 2015.
- [7] T. Beauchamp and J. Childress. *Principles of Biomedical Ethics*. Principles of Biomedical Ethics. Oxford University Press, 2001.
- [8] J. Bentham. *A fragment on government*. The Lawbook Exchange, Ltd., 2001.
- [9] F. Berreby, G. Bourgne, and J.-G. Ganascia. Modeling moral reasoning and ethical responsibility with logic programming. In *Logic for Programming, Artificial Intelligence, and Reasoning*, pages 532–548. Springer, 2015.
- [10] S. Bringsjord and J. Taylor. The divine-command approach to robot ethics. *Robot Ethics : The Ethical and Social Implications of Robotics*, MIT Press, Cambridge, MA, pages 85–108, 2012.
- [11] N. Cointe, G. Bonnet, and O. Boissier. Ethical judgment of agents’ behaviors in multi-agent systems. In *Proceedings of the 2016 International Conference on Autonomous Agents & Multiagent Systems*, pages 1106–1114, 2016.
- [12] L. Dennis, M. Fisher, M. Slavkovik, and M. Webster. Formal verification of ethical choices in autonomous systems. *Robotics and Autonomous Systems*, 77 :1–14, 2016.
- [13] L. Dennis, M. Fisher, and A. Winfield. Towards verifiably ethical robot behaviour. *arXiv preprint arXiv :1504.03592*, 2015.
- [14] P. Foot. The problem of abortion and the doctrine of the double effect. *Applied Ethics : Critical Concepts in Philosophy*, 2 :187, 2002.
- [15] J.-G. Ganascia. Non-monotonic resolution of conflicts for ethical reasoning. In *A Construction Manual for Robots’ Ethical Systems*, pages 101–118. Springer, 2015.
- [16] J. Horty. Nonmonotonic foundations for deontic logic. In *Defeasible deontic logic*. Springer, 1997.
- [17] I. Kant. Groundwork of the metaphysic of morals, trans. hj paton. *New York : Harper & Row*, 1964.
- [18] C. Kluckhohn. *Values and value-orientations in the theory of action : An exploration in definition and classification*. 1951.
- [19] R. Kowalski. *Computational logic and human thinking : how to be artificially intelligent*. Cambridge University Press, 2011.
- [20] V. Lifschitz. What Is Answer Set Programming ?. In *AAAI*, volume 8, pages 1594–1597, 2008.
- [21] D. Luce and H. Ruffia. *Games and decisions*. mineola, ny, 1985.
- [22] J.-T. Mangan. Historical analysis of the principle of double effect, an. *Theological Studies*, 10, 1949.

- [23] R. Miller and M. Shanahan. Some alternative formulations of the event calculus. In *Computational logic : logic programming and beyond*. Springer, 2002.
- [24] R. Nozick. *Anarchy, state, and utopia*, 1974.
- [25] H. of Cos. *The Oath*. Loeb Classical Library, 1923.
- [26] L. Pereira and A. Saptawijaya. Modelling morality with prospective logic. In *Progress in Artificial Intelligence*, pages 99–111. Springer, 2007.
- [27] M. Shanahan. The event calculus explained. In *Artificial intelligence today*. Springer, 1999.
- [28] K. Struhl and P. Rothenberg. *Ethics in perspective : a reader*. Random House, 1975.
- [29] M. Tufiş and J.-G. Ganascia. Grafting norms onto the bdi agent model. In *A Construction Manual for Robots' Ethical Systems*. Springer, 2015.



## Approche paramétrique de bandits pour le test A/B

Emmanuelle Claeys<sup>1\*</sup> Pierre Gançarski<sup>1</sup> Myriam Maumy-Bertrand<sup>2</sup> Hubert Wassner<sup>3</sup>

<sup>1</sup> ICube – Université de Strasbourg – 67412 Illkirch – France

<sup>2</sup> IRMA – Université de Strasbourg – 67084 Strasbourg – France

<sup>3</sup> Entreprise AB Tasty – 3 Impasse de la Planchette – 75003 Paris – France

### Résumé

*Cet article propose une méthode pour optimiser l'approche de bandits en utilisant un arbre décisionnel dans le contexte du test A/B. Il rappelle les problématiques que le test A/B rencontre et propose d'anticiper le comportement d'un profil de visiteurs, afin d'orienter la stratégie (proposer la meilleure variation au visiteur) du test selon ce comportement.*

### Mots clés

Méthodes de bandits, personnalisation, clustering temporel, test A/B, allocation dynamique.

### Abstract

*This paper proposes a new architectural model for A/B testing evaluation in multiple contexts. It exposes the global problematic of A/B testing when we want to manage several types of e-commerces. The goal of this paper is to provide a method for optimize bandits approach using a decision tree.*

### Keywords

Bandit approach, personalization, time series clustering, A/B testing, exploration/exploitation dilemma.

## 1 Introduction

Pour un *e-commerçant*, c'est-à-dire un commerçant dont l'activité commerciale dépend du WEB, la qualité du site (ergonomie, iconographies, guide et parcours) est primordiale. Modifier celui-ci afin de le rendre plus efficace est donc tout aussi crucial. Néanmoins, toute modification n'est pas bonne à prendre car elle peut dégrader le site et donc impacter l'activité du *e-commerçant*. Il n'est donc pas envisageable pour le *e-commerçant* de mettre en ligne un changement sur son site sans en avoir au préalable tester la pertinence. Pour cela, de plus en plus d'entreprises utilisent des outils d'analyse pour le web marketing, comme, par exemple, les plate-forme de test A/B. Un test A/B (ou A/B testing<sup>1</sup>), consiste à créer plusieurs versions d'un

même objet, dans le but de vérifier une hypothèse marketing [1]. Par exemple, il s'agit de tester auprès d'un échantillon d'internautes, nommés *visiteurs*, durant une période donnée, différentes variations (généralement deux) d'une même page web (nommées variation A et variation B) proposée par un *e-commerçant*. Le but est d'identifier celle qui donne les meilleurs résultats par rapport à un objectif fixé en amont du test (comme par exemple l'augmentation du taux de *conversion*, le nombre de visites ou de clics sur une bannière ou encore la valeur moyenne du panier, etc.). Le test peut alors être évalué sur l'ensemble des visiteurs de la page ou uniquement sur un échantillon de visiteurs ayant une caractéristique spécifique (par exemple les visiteurs parisiens). À l'issue du test, le *e-commerçant* pourra soit conserver A, soit conserver B, soit, dans une optique de personnalisation, présenter la variation A ou B à des visiteurs en fonction de caractéristiques spécifiques, comme par exemple la géo-localisation.

Derrière cette problématique de marketing, se cachent des verrous scientifiques importants, à cause de la contrainte suivante : lors d'un test A/B, un visiteur est affecté à une variation (en général de façon équiprobable) à sa première venue sur la page, celle-ci lui est affectée jusqu'à la fin du test ainsi que pour toutes ses futures visites. De fait, une caractéristique importante des tests A/B est qu'un visiteur ne voit qu'une seule des deux variations et que de ce fait, l'ensemble des visiteurs qui ont vu la variation A est totalement disjoint de celui des visiteurs ayant vu la variation B. De par cette contrainte, il devient difficile résoudre la problématique par des algorithmes de recommandation classiques qui permettent de réitérer la proposition d'une variation à un visiteur [2].

Du fait de cette irrévocabilité de la décision d'affectation, la question sous-jacente est alors de savoir comment déterminer ce qu'aurait fait un visiteur sur une variation sachant ce qu'il a fait sur l'autre. Enfin, d'autres caractéristiques peuvent intervenir dans l'analyse du test : on ne peut pas, en général déterminer lorsqu'un visiteur ne fait pas de conversion, si cela est dû à la variation présentée ou simplement à l'absence de volonté a priori du visiteur de faire cette conversion (simple visite d'informations sur des offres du site par exemple) ; le nombre de conversions peut être très

\*claeys@unistra.fr

1. Le terme A/B étant un terme générique pour qualifier tout type de tests.

largement inférieur au nombre de non-conversions ; le délai entre la conversion et la visite de la page peut être plus ou moins long, etc.

Enfin, les tests statistiques existants [3] pour répondre à ces questions, se heurtent au problème de la *politique d'allocation dynamique*. En effet, l'identification de la meilleure variation et la caractérisation d'un ou plusieurs segments de visiteurs sont importantes pour le e-commerçant, mais sa priorité principale est avant tout de tirer des bénéfices de son site. Si l'une des variations a des conséquences désastreuses, il souhaitera limiter la phase d'*exploration* (c'est-à-dire de test) au profit de la phase d'*exploitation* de la meilleure variation courante : la politique d'allocation changera au fil de temps pour affecter prioritairement les nouveaux visiteurs sur la variation la plus rentable. C'est pour répondre à ce compromis d'exploration-exploitation que les méthodes dites « de bandits » sont très utilisées (cf. Section 2.2)

**Notre méthode.** Pour répondre à l'ensemble de ces problématiques, nous proposons une méthode combinant une *étape de pré-analyse* (hors-ligne) dans laquelle nous identifions des segments de visiteurs à partir de leurs profils temporels et de leurs centres d'intérêt liés à leurs comportements lors de leur(s) visite(s) suivie d'une *étape d'observation* (en-ligne) correspondant au test proprement dit, durant laquelle nous observons si des segments présentent des taux de conversions différents. Pour cela, pour chaque segment de visiteurs (identifié dans l'étape de pré-analyse), nous utilisons un algorithme de comparaison de variations (méthode de bandit, cf. Section 2.2) de façon totalement indépendante. Ainsi, nous proposons :

- d'identifier des segments des visiteurs en fonction de leur profils temporels et de leurs centres d'intérêts à l'aide de techniques de clustering temporel ;
- d'effectuer les comparaisons, lors de la phase d'analyse, par segment et non uniquement globalement ;
- de limiter les pertes d'exploitation dues à l'expérience en modifiant la politique d'allocation pour chacun de ces segments (donc localement et non plus globalement) en fonction de la convergence plus ou moins rapide du test vers la version optimale pour le segment (si elle existe).

## 2 Mise en œuvre de l'approche

### 2.1 Étape 1 : pré-analyse

Cette étape consiste à analyser les visiteurs pendant une période pré-définie (généralement une à deux fois la durée prévue pour le test lui-même) afin d'extraire et de caractériser des segments de visiteurs suivant leurs profils temporels et leurs centres d'intérêt.

**Profils temporels.** À l'issue de la phase de pré-analyse, à chaque visiteur est associé deux vecteurs, de taille égale à la durée de la pré-analyse. Le premier vecteur  $V_{presence}$  indique la visite ou non pour chaque intervalle de temps

(en général chaque jour) :  $V_{presence}[i] = 1$  si un visiteur a visité la page à la  $i$ -ème date, 0 sinon. Le second vecteur  $V_{pages}$  donne le nombre de pages consultées à chaque période par ce visiteur.

**Prise en compte des centres d'intérêt.** Afin de mieux caractériser les visiteurs, nous avons affecté à chacun d'eux un ensemble de centre d'intérêts en fonction des pages qu'il a traversées avant d'arriver à la page testée. Pour cela nous procédons en deux étapes : (1) recherche des différents mots clés présents pour les pages parcourues par l'ensemble des visiteurs et association à chaque page parcourue d'un ou plusieurs *sujets d'intérêt* (topics) construits à partir de ces mots clés et (2) affectation à chaque visiteur de l'ensemble des sujets associés aux pages qu'il a personnellement parcourues. À noter, qu'un topic peut être associé à plusieurs pages. Cette association a été réalisée grâce à une méthode de *topic modeling* via l'allocation de Dirichlet latente [6]. Le *topic modeling* est une méthode en traitement automatique du langage naturel permettant de modéliser des relations entre différents sujets. L'allocation de Dirichlet latente est basée sur une combinaison de modèles statistiques bayésiens, pour des données discrètes où les sujets ne sont pas corrélés. Le topic modeling est une extension du modèle allocation de Dirichlet latente (L.D.A) lorsque les mots des documents sont corrélés entre les sujets. Nous avons choisi cette méthode pour son utilisation populaire et la facilité de sa mise en œuvre. L.D.A. utilise un nombre de sujets fixé  $k$  et travaille itérativement sur chaque document  $w = (w_1, w_2, \dots, w_N)$  d'un corpus  $D$  contenant  $N$  mots d'un vocabulaire composé de  $V$  termes différents  $w_i \in 1, \dots, V$  pour  $i = 1, \dots, N$ .

**Construction des clusters de visiteurs.** Pour construire les clusters de visiteurs (qui seront utilisés pour définir nos segments) nous utilisons les deux vecteurs identifiés précédemment. Ces vecteurs peuvent-être vus comme des séries temporelles, allouées à chaque visiteur. De nombreuses méthodes existent pour réaliser des clusters de séries temporelles. L'approche que nous avons retenue est celle basée sur la métrique D.T.W. (Dynamic Time Warping) car celle-ci présente un certain nombre d'avantages [4] tels que :

- toutes les données sont utilisées ;
- l'alignement des séquences obtenu est optimal ;
- la possibilité d'extraire des comportements décalés ;
- supporte un échantillonnage irrégulier ;
- permet de comparer des séquences de longueurs différentes.

Par contre, elle s'avère assez sensible aux valeurs extrêmes et au bruit.

Nous avons donc choisi l'algorithme de partitionnement (clustering) K-means pour sa simplicité d'utilisation et son efficacité. Néanmoins, en plus d'une mesure de (dis)similarité, celui-ci nécessite un mécanisme de calcul des centres des clusters. Pour cela, nous utilisons la méthode D.B.A. (D.T.W. Barycenter Averaging) [4] sous le langage R [5], qui est une méthode itérative de calcul d'une

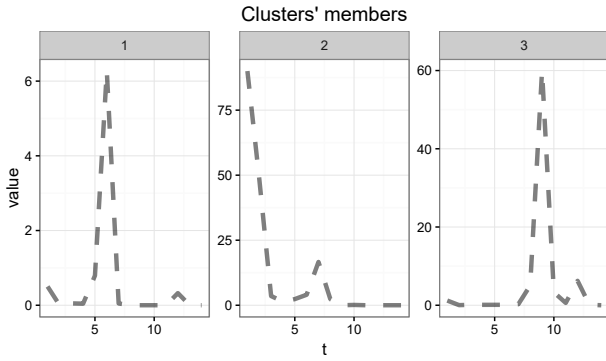


FIGURE 1 – Centres de trois clusters présentant trois évolutions du nombre de pages consultées.

moyenne. La figure 1 montre un exemple de centres pour trois clusters obtenus à partir de pages parcourues à chaque visite par les différents visiteurs et ce sur une période d’observation de 15 jours.

**Recherche de régularités dans les profils de visiteurs.** L’objectif de cette étape est de voir si certains (sous)segments de visiteurs répondent plus facilement à l’objectif de conversion.

Nous avons choisi d’utiliser un arbre de régression non paramétrique (algorithme CTREE) [8] consistant à étudier tous les profils avec une différence statistiquement significative dans le taux de conversions (test du khi-deux de Pearson, avec un niveau de seuil de confiance fixé à 95 %). Dans notre cas, cet arbre sera utilisé pour identifier les facteurs pronostiques, c’est-à-dire le lien entre les clusters et les topics construits précédemment et les différents taux de conversions.

Un tel arbre se construit de manière itérative en divisant, à chaque étape, la population en deux sous-ensembles. Pour diviser cette population, l’algorithme utilise des tests de permutation [9]. Des procédures d’essais multiples sont appliquées pour déterminer l’association significative entre les variables, via une approche récursive [7].

Pour identifier des variables explicatives discriminantes, l’arbre utilise des procédures de test spéciales comme le test de Spearman, le test de Wilcoxon-Mann-Whitney ou le test de Kruskal-Wallis ainsi que des tests de permutation basés sur des statistiques d’ANOVA ou des coefficients de corrélation. Dans notre expérimentation, pour choisir le test approprié pour vérifier le rapport de différence, nous utilisons une méthode multiple de correction de Bonferroni associée à une méthode de Monte Carlo et à un test de Kruskal-Wallis.

Ceci conclut une pré-analyse, la figure 2 présente un exemple d’arbre obtenu sur un jeu de tests *train\_set*.

## 2.2 Étape 2 : Observation

Remarque : Pour affecter un nouveau visiteur lors de l’étape d’observation (donc du test), nous considérons son

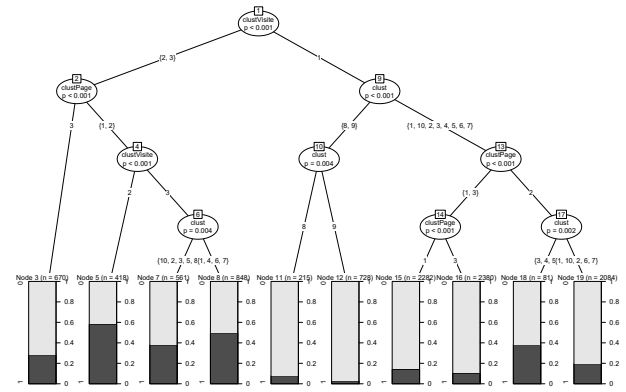


FIGURE 2 – Arbre de régression de *train\_set*

parcours depuis sa première arrivée sur le site (sachant qu’il n’a pas encore vu la page test) et nous l’affectons au cluster le plus proche (nous procédons simplement par comparaison du parcours du visiteur aux centres des clusters en utilisant D.T.W.).

Dans la littérature, il existe de nombreuses méthodes pour comparer un test sur deux échantillons, mais elles supposent que les échantillons soient choisis aléatoirement à partir d’une population homogène [3]. Ces méthodes sont capables de gérer de larges volumes de données, avec une distribution paramétrique ou non. Néanmoins, dans le cas du marketing, l’observation d’une variable (par exemple l’achat d’un produit), peut varier selon les différentes populations, cela pouvant être dû aux fluctuations de l’échantillonnage ou au contraire cela révèle des comportements différents selon les catégories. De fait, ces méthodes sont difficiles à utiliser dans le cas des tests *A/B* :

- elles n’expliquent pas pour quel profil la variation *A* ou *B* est la plus optimale.
- les faibles différences peuvent être difficiles à interpréter (convergence trop rapide vers l’exploitation).

De plus, dans le e-commerce, le taux de conversions est particulièrement bas (de l’ordre de 2 à 3 %) et dépend de la position de la page testée dans le parcours de conversion. Enfin, un enjeu important du test *A/B* est de limiter les pertes dues à l’expérience. Par exemple, si le e-commerçant veut réaliser un test hasardeux comme par exemple changer la photo du produit ou ajouter une fenêtre promotionnelle (pop-up), il veut limiter les pertes qu’engendrait la variation, dès lors qu’il est convaincu que celle-ci est mauvaise pour tous les profils. Cette approche peut être vue comme une minimisation du regret. Dans de nombreux tests *A/B*, l’échantillonnage est basé sur une proportion égale du trafic (50% du trafic a été alloué à *A* et 50% à *B*). Dans cette situation, une variation optimale devra attendre la fin du test pour être choisie.

Il est cependant possible de changer l’allocation au cours du temps via de nombreuses méthodes : les méthodes dites de *bandits* utilisent un échantillonnage adaptatif lorsque la meilleure variation est identifiée. Cette approche vue

comme un compromis exploration-exploitation, est aussi vue comme une méthode d'apprentissage par renforcement.

Un *méthode de bandit* consiste à choisir une solution, de façon itérative, parmi  $N$  solutions, chacune donnant une récompense moyenne inconnue a priori, la solution permettant de maximiser le gain [12]. Les algorithmes mis en œuvre peuvent éventuellement être paramétrés par le type ou la forme de la distribution, le seuil de confiance et la perte maximale acceptable. Une variante classique dite de *bandits multi-bras* consiste à identifier la meilleure variation parmi plusieurs variations possibles. Enfin, les *bandits contextuels* sont particulièrement populaires car ils utilisent le contexte rattaché à un visiteur pour prendre la meilleure décision. La méthode utilise pour cela un vecteur de contexte à  $d$  dimensions. Le processus d'apprentissage utilise ces vecteurs de contexte ainsi que la récompense de la variation alors affichée. Le lecteur intéressé trouvera dans la littérature un grand nombre de telles méthodes. On pourra citer par exemple :

- L'algorithme `LinUCB` : présume une dépendance linéaire (à identifier) entre le gain attendu et le vecteur de contexte [10].
- L'algorithme `UCBogram` : présume une dépendance non linéaire (à identifier) entre le gain attendu et le vecteur de contexte.
- L'algorithme `NeuralBandit` : approche neuronale où un réseau apprend le contexte et le gain associé.
- L'algorithme `KernelUCB` : version non linéaire de **LinUCB** en utilisant une matrice à noyaux. Utilisée pour le on-line learning.
- L'algorithme `Bandit Forest` : Utilise des forêts aléatoires de bandits basées sur une distribution commune des gains et du contexte. Utile lorsque l'expérience est reproductible.

En pratique ces méthodes sont efficaces mais présentent quelques lacunes ou inconvénients. Par exemple, on considère le non-achat du visiteur comme un échec dès qu'il quitte la page testée. Cependant, un visiteur peut répondre à l'objectif de conversion après plusieurs visites. Si nous considérons le nombre de visites comme un contexte, nous pouvons alors limiter la pénalisation pour le bandit.

Enfin, s'ajoute une contrainte fondamentale pour l'intégration de l'algorithme décisionnel en industrie : le temps alloué à la prise de décision. Il est inenvisageable pour le e-commerçant que le système prenne trop de temps pour décider quelle variation attribuer au visiteur : Quand un visiteur arrive sur un site web, le système doit répondre en moins de 100ms pour éviter d'impacter la visite du visiteur sur le site. Le e-marchand ignore également les informations à utiliser pour définir le contexte, encore plus lorsque les données collectées contiennent des variables potentiellement non pertinentes. Enfin, le trafic et les achats fluctuent au cours du temps, ce qui perturbe parfois l'identification de la variation la mieux adaptée.

Pour notre approche, nous mettons en œuvre une méthode de bandit binomiale [11] pour pour chaque segment de visiteurs identifiés dans notre arbre de pré-analyse. Il s'agit donc d'associer un bandit par segment. Pour chaque nouveau visiteur appartenant au segment considéré, le bandit lui alloue une des variations disponibles (ici deux :  $A$  et  $B$ ), selon la probabilité bayésienne postérieure (estimée de façon itérative) que chaque variation, dite aussi bras, soit optimale. Dans notre cas, le gain est une valeur binaire (1 : conversion ou 0 : non conversion), la fonction de gain suit une distribution de Bernoulli dont les paramètres sont estimés au cours du temps.

## 3 Expérimentation

### 3.1 Données et pré-analyse

Le jeu de données de validation est issu d'un test  $A/B$

- où la variation  $B$  consistait à changer le lien d'un bouton "retour",
- sans allocation dynamique (60% du trafic a été alloué à  $A$  et 40% à  $B$ ),
- réalisé du 11/07/16 au 25/07/16 pour un e-commerce de vêtements.

Le site de l'e-commerçant reçoit environ 150 000 visiteurs par mois, les données récoltées ici portent essentiellement sur la navigation et le comportement : nous utilisons pour chaque visiteur, l'historique de ses visites, de pages vues, de transactions éventuelles et les dates associées à ces événements. Pour simuler l'étape de pré-analyse (resp. d'observation), nous sélectionnons les visiteurs arrivés avant le 15/07/16 inclus (resp. arrivés après le 16/07/16 inclus).

**Étape de pré-analyse** Nous construisons un arbre de classification exclusivement sur les visiteurs de *train\_set* qui ont vu la version originale  $A$ . Nous cherchons à identifier les caractéristiques des populations ayant un taux de conversions différent (l'objectif ici est "Achat", "Pas d'achat"). Trois caractéristiques sont associées pour chaque visiteur.

1. Le cluster de visites basé sur la série temporelle de visites (présence ou non présence du visiteur sur le site, pour chaque jour du test).
2. Le cluster de pages basé sur la série temporelle de pages consultées (nombre de pages consultées sur le site, pour chaque jour du test).
3. Le topic associé aux pages qu'il a consulté lors de ces visites passées<sup>2</sup>.

Pour simplifier le test, nous considérerons que dans nos expériences tous les visiteurs sont parfaitement associés à leurs clusters.

La figure 2 présente l'arbre de classification issu du jeu de données de pré-analyse *train\_set* pour la variation  $A$ . Cet arbre identifie 10 profils (clusters) distincts. Il a été

<sup>2</sup> pour des raisons techniques, nous utiliserons ici uniquement la dernière page consultée avant qu'il n'arrive sur la page test.



TABLE 1 – Comparaison des méthodes

Bras	Notre méthode	Bandit Bayésien	LinUCB
Gagnant	X	B	B
Gain	1457	1252	1300
Visiteurs	11168	11168	11168
Proba	X	0.69	X

construit par l’algorithme CTREE de façon à ce que la différence entre les taux de conversions des différents profils soit significative avec un indice de confiance de 95% et en évaluant l’association entre le taux de conversions et toutes les variables explicatives via une statistique de corrélation quadratique. Selon les variables impliquées (numériques ou catégorielles), différents types de scores sont utilisés dans ces tests de corrélation. Par défaut, les  $p$ -valeurs résultantes sont corrigées par une correction de Bonferroni [13] pour des tests multiples à travers le nombre de variables utilisées pour la régression.

### 3.2 Analyse des résultats d’observation

Afin de valider notre méthode, nous la comparons avec la même méthode de bandit, mais sans contexte (un seul bandit, pas de segment particulier), ainsi qu’avec l’algorithme LinUCB très utilisé actuellement pour ce genre de test. Nous appliquons ces trois méthodes sur le même jeu de données *exp\_set*. Ainsi pour simuler la phase d’observation, nous prenons les visiteurs un à un dans le jeu de test et nous le soumettons à chacune des trois méthodes.

Pour simuler notre méthode, nous avons utilisé 10 bandits chargés d’allouer la proportion du trafic vers *A* ou *B* pour des visiteurs aux caractéristiques spécifiques (ces caractéristiques étant identifiés par les feuilles de l’arbre de pré-analyse). Ainsi, par exemple, le bandit *node3* (Figure 2) travaille sur les visiteurs associés à la fois au cluster de pages n° 3 ( $ClustPage \in \{3\}$ ) et aux clusters de visites n° 2 ou 3 ( $ClustVisite \in \{2, 3\}$ ). Il est à noter qu’un visiteur ne peut pas être associé à plusieurs bandits car l’arbre détermine des populations disjointes.

La méthode du bandit LinUCB identifie le meilleur bras selon différentes variables explicatives (ici les informations liées aux contextes sont les mêmes informations que nous utilisons pour réaliser notre arbre de pré-analyse). Lorsque LinUCB affecte ce «nouveau» visiteur *NV* à une variation différente de celle qui avait été appliquée dans la réalité, nous cherchons un visiteur *NA* aux mêmes caractéristiques, affecté à l’autre variation et qui n’a pas encore été proposé à l’algorithme. Nous inversons alors ce deux visiteurs *NV* et *NA*. Dans le cas (rare) où il n’existe plus de visiteurs *NA* dans le jeu de données, nous réalisons un tirage aléatoire parmi tous les visiteurs de même profil dans le jeu de données. Cette méthode comporte donc une part d’aléatoire. Néanmoins, les expériences ont montré que lorsque cette problématique se produit (pour les derniers visiteurs), LinUCB a en général déjà convergé vers la variation optimale.

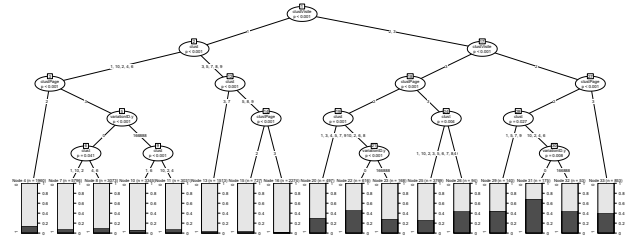


FIGURE 3 – Arbre de régression issu des données *exp\_set*

Le tableau 1 compare les trois approches suivant deux critères :

1. Le nombre de conversions (gain) obtenues.
2. La probabilité d’être le meilleur bras.

Le tableau 2 reprend les résultats de notre méthode suivant les différents segments.

Notre méthode aboutit à 6 variations optimales avec une probabilité de plus de 70%. Le gain total réalisé est de 1457 conversions. Aucun nouveau visiteur n’a été alloué au bandit 5. Certains bandits convergent rapidement.

Pour la méthode du bandit binomial sans contexte prédéfini, la probabilité d’avoir identifié le meilleur bras ne dépasse les 70% qu’à partir de 4000 visiteurs. Cette méthode réalise un gain total inférieur à 14% à notre méthode. La méthode du bandit LinUCB réalise un gain moyen de l’ordre de 1300 conversions : à aucun moment, LinUCB n’a dépassé le gain obtenu par notre méthode.

Nous avons également cherché à savoir si notre méthode retrouvait les résultats réels du jeu de données initial.

La figure 4 reprend la figure 3, soit les variables discriminantes sur l’ensemble de la population testée via l’allocation aléatoire, et présente également quels ont été les profils affectés à chaque bandit. Comme l’arbre de construction des bandits a été réalisé sur une population différente (visiteur arrivé entre le 11/07/16 et 14/07/16 contre les visiteurs du 15/07/16 au 25/07/16 pour la phase de test), cet arbre est plus précis que celui de la phase d’apprentissage. Il comprend également les variations *A* et *B* qui permettent d’observer si une population a été impactée par le test (Dans notre jeu de donnée, la variation *A* est appelée 0 et la variation *B* : 166888). La variable *Clust* est le "topic" associé à un visiteur. La variable *clustPage* est le cluster construit à partir des pages parcourues, enfin *clustVisite* est le cluster construit à partir des visites réalisées.

Dans tous les cas, la variation *A* a été meilleure que la variation *B*. D’après l’arbre de la figure 4, on distingue 3 profils impactés par le test :

- Les visiteurs à la variable  $clust \in \{1, 10, 2, 4, 6\}$ ,  $clustPage \in \{3\}$  et  $clustVisite \in \{1\}$ . Ces visiteurs ont été traités par le bandit 16. La meilleure variation était la variation *A*.
- Les visiteurs à la variable  $clust \in \{10, 2, 6, 8\}$ ,  $clustPage \in \{2\}$  et  $clustVisite \in \{3\}$ . Ces vi-

TABLE 2 – Notre méthode

Bras	3	5	7	8	11	12	15	16	18	19	TOTAL
Gagnant	A	B	A	B	A	A	X	A	A	A	X
Gain	573	50	39	51	10	9	0	628	6	91	1457
Visiteurs	1878	99	131	181	310	475	0	7380	105	609	11168
Proba	0.66	0.74	0.88	0.63	0.52	0.72	0	0.96	0.73	0.91	X

siteurs ont été traités par les bandits 7 et 8. La meilleure variation était la variation A.

- Les visiteurs à la variable  $clust \in \{10, 2, 4, 6\}$ ,  $clusPage \in \{2\}$  et  $clusVisit \in \{2\}$ . Ces visiteurs ont été traités par les bandits 7 et 8. La meilleure variation était la variation A.

Dans notre méthode

- Le bandit 7 converge vers A avec une probabilité à 88% d'être le meilleur.
- Le bandit 8 converge vers B avec une probabilité à 63% d'être le meilleur.
- Le bandit 16 converge vers A avec une probabilité à 96% d'être le meilleur.

Notre méthode est donc aussi fiable qu'une méthode fréquentielle, lorsque la variation impacte faiblement une population, mais elle apporte les avantages supplémentaires suivants :

- Identification de la variation optimale pour une segmentation particulière.
- Convergence plus rapide vers la phase d'exploitation.
- Limitation des convergences vers la mauvaise variation.
- Gain comparable à une allocation aléatoire lorsque la variation impacte une faible proportion de la population.

### 3.3 Conclusion et travaux futurs

Une des problèmes majeurs des bandits dans un trafic particulièrement volatil qu'est celui du e-commerce est que la convergence vers une mauvaise solution est difficile à contrer. En utilisant des bandits de façon indépendantes, nous limitons les erreurs d'identification de la meilleure variation. De par notre utilisation du clustering, nous tirons profit des profils issus des parcours et des centres d'intérêt des visiteurs. Ces profils permettent de limiter l'information émise pour le contexte et anticipent la création trop importante de bandits, qui seraient alors soumis à un trafic trop limité pour converger. Actuellement nous avons automatisé le nombre de sujets à définir et testé cette approche avec d'autres e-commerçants. Nous travaillons également sur la qualité des clusters, et abordons les sujets d'intérêt de visiteurs de façon évolutive. Enfin, nous souhaitons utiliser des paramètres relatifs au trafic global du site afin d'anticiper les "pics" et "creux" de trafic qui peuvent nuire à la phase d'exploration (en période de faible trafic, les différences entre les variations sont plus réduites).

## Références

- [1] S. Roukine, *Améliorer ses taux de conversions web : vers la performance des sites web au-delà du webmarketing*, Eyrolles, 2011.
- [2] E. Kaufmann, O. Cappé et A. Garivier, On the Complexity of A/B Testing, *Conference on Learning Theory, Barcelona, Spain, JMLR : Workshop and Conference Proceedings*, Vol. 35, pp 461-481, 2014.
- [3] A. Wald, Sequential tests of statistical hypotheses. *Annals of Mathematical Statistics*, Vol. 16, pp 117-186, 1945.
- [4] A. Petitjean, A. Ketterlin et P. Gançarski, A global averaging method for dynamic time warping, with applications to clustering *Pattern Recognition archive*, Vol. 44, pp 678-693, 2011.
- [5] A. Sarda-Espinosa, Time Series Clustering Along with Optimizations for the Dynamic Time Warping Distance, *dtwclust R package version 3.1.0*, 2017. <https://github.com/asardaes/dtwclust>
- [6] B. Grun et K. Hornik, An R Package for Fitting Topic Models. *Journal of Statistical Software*, Vol. 40, pp 1-30, 2011.
- [7] H. Strasser et C. Weber On the Asymptotic Theory of Permutation Statistics *Mathematical Methods of Statistics*, Vol. 8, pp 220–250, 1991
- [8] T. Hothorn, K. Hornik et A. Zeileis, Unbiased Recursive Partitioning : A Conditional Inference Framework, *Journal of Computational and Graphical Statistics*, Vol. 15, pp 651–674, 2006.
- [9] AM. Molinaro, S. Dudoit et MJ. van der Laan, Tree-Based Multivariate Regression and Density Estimation with Right-Censored Data, *Journal of Multivariate Analysis*, Vol. 90, pp 154–177, 2004.
- [10] L. Li, W. Chu, J. Langford, John et E. Schapire, A contextual-bandit approach to personalized news article recommendation, *Proceedings of the Nineteenth International Conference on the World Wide Web*, pp 661-670, 2010.
- [11] S. Scott, A Modern Bayesian Look at the Multi-armed Bandit, *Applied Stochastic Models Business and Industry*, Vol. 26, pp 639–658, 2010.
- [12] P. Auer, N. Cesa-Bianchi, Y. Freund, et R. E. Schapire, The nonstochastic multiarmed bandit problem, *SIAM Journal on Computing*, Vol. 32 pp 48–77, 2002.

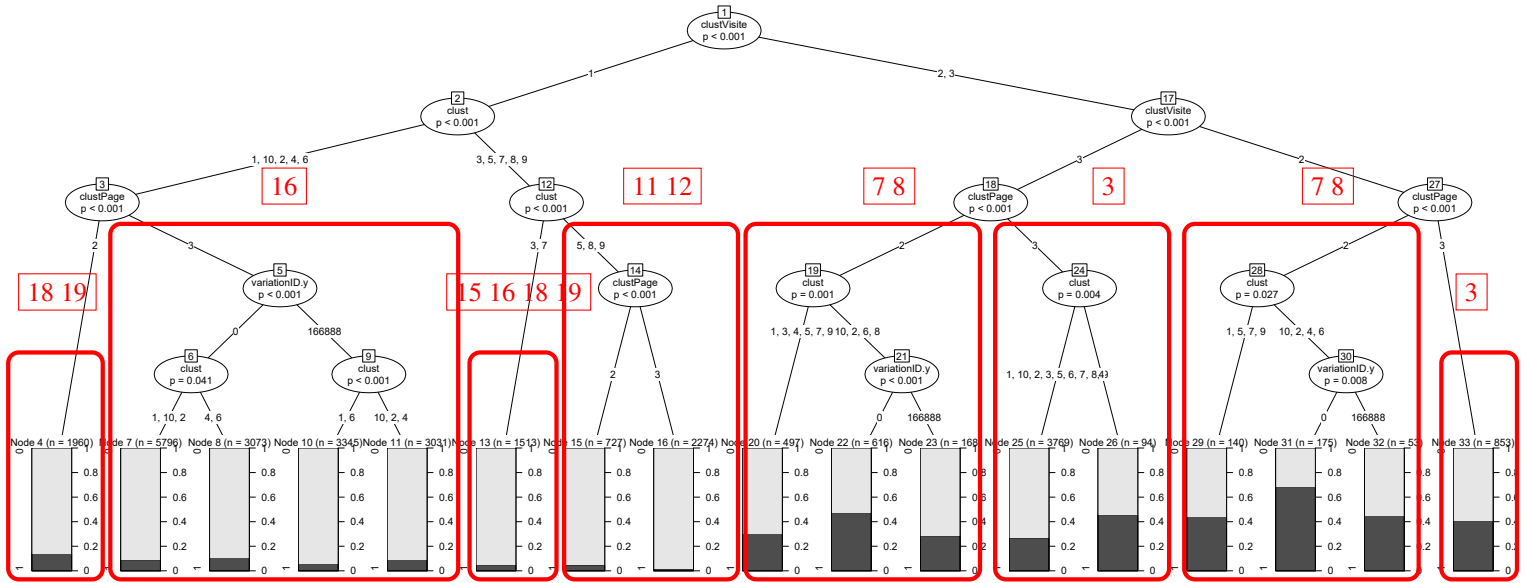


FIGURE 4 – Arbre de régression de *exp\_set* (avec bandits)

[13] W-Y. Loh Fifty Years of Classification and Regression Trees *International Statistical Review*, Vol. 82 pp 329–348, 2014.



# Que valent les stratégies probabilistes au dilemme itéré du prisonnier ?

Jean Paul Delahaye

Philippe Mathieu

CRISAL Lab, UMR 9189 CNRS, Université des Sciences et Technologies de Lille, France

{prenom.nom}@univ-lille.fr

## Résumé

*Nous menons une étude expérimentale minutieuse sur les stratégies probabilistes au dilemme des prisonniers. Nous utilisons pour cela la méthode des classes complètes associée à une approche évolutionniste. Les résultats que nous obtenons ont donc un caractère objectif et dépendent le moins possible des ensembles de stratégies mis en compétition. Les ensembles étudiés sont grands (plusieurs milliers de stratégies), homogènes, et systématiques. Nous testons la robustesse de nos résultats par diverses méthodes. Les stratégies les meilleures repérées sont pour certaines d'entre elles nouvelles en ce sens qu'elles n'ont jamais été identifiées clairement par des études antérieures, et cela malgré leur simplicité. Nous identifions un critère jusque là inconnu qui conduit à une bonne anticipation du comportement des stratégies dans des univers variés. Nous confrontons les résultats de cette étude avec ceux obtenus par les approches mathématiques de Press et Dyson. Nous confrontons aussi les nouvelles stratégies avec les meilleures stratégies connues.*

## Mots Clef

Theorie des jeux, Dilemme du prisonnier, stratégies d'agents, comportement

## Abstract

*We conduct a thorough experimental study of probabilistic strategies to the prisoner's dilemma. To do this, we use the complete class method associated with an evolutionary approach. The results we obtain are therefore objective in nature and depend as little as possible on the sets of strategies put in competition. The studied sets are large (several thousand strategies), homogeneous, and systematic. We test the robustness of our results by various methods. The best strategies identified are for some of them new in the sense that they have never been clearly identified by previous studies, despite their simplicity. We propose a criterion that leads to a good anticipation of their behavior in various contexts. We compare the results of this study with those obtained by the mathematical approaches of Press and Dyson. We also confront the new strategies with the best known strategies.*

## Keywords

Game Theory, Iterated prisoner's Dilemma, Agent's Strategy, Behaviour

## 1 Introduction

Suite à la parution de l'article de Press et Dyson [21] sur ce qu'ils ont appelé les stratégies ZD et l'extorsion, et suite aux réactions parfois critiques [1, 2, 10, 12, 11, 16, 19, 25] concernant leurs conclusions, l'attention a été portée sur les stratégies probabilistes au dilemme itéré du prisonnier [24, 8, 9, 11]. Pourtant aucune méthode de tri systématique et aussi exhaustive que possible n'a été utilisée pour savoir si des stratégies probabilistes simples égalaient ou surpassaient les meilleures stratégies connues [4, 18]. Nous menons ici une telle étude en combinant deux des méthodes que nous considérons comme les plus susceptibles de produire des résultats robustes et dépourvus de subjectivité : la méthode des compétitions évolutionnaires parfois appelée écologiques [3, 5, 4, 8, 23], et la méthode des classes complètes [6, 8, 17]. Nous utilisons en particulier des classes complètes homogènes composées de stratégies aléatoires.

Nos résultats montrent que des stratégies simples et pourtant inconnues jusqu'à présent émergent parmi les milliers de stratégies mises en compétition. Une catégorie assez large de stratégies est identifiée comme robuste et performante pour les compétitions évolutives. Un paramètre noté  $p'$  est identifié et interprété ; il est corrélé à la réussite des stratégies et semble donc fournir un critère efficace pour prévoir ce que donnera une stratégie probabiliste dans une compétition évolutive. Des variantes plus complexes, mais plus fines de ce paramètre sont recherchées par une méthode d'exploration statistique exhaustive.

Des séries systématiques de tests sont menées pour s'assurer de la robustesse des résultats obtenus. En particulier nous plongeons les nouvelles stratégies probabilistes identifiées dans des environnements de stratégies déterministes pour nous assurer qu'elles restent performantes en dehors du contexte qui a permis de les découvrir. Les nouvelles venues sont aussi confrontées aux stratégies repérées par Press et Dyson et aux stratégies les meilleures connues pour le dilemme itéré du prisonnier. Cela nous conduit à une nouvelle formulation des conclusions expérimentales générales sur les stratégies optimales connues au dilemme

du prisonnier.

## 2 Définitions et rappels

Le dilemme du prisonnier [3, 22, 23, 13] est celui auquel sont soumis deux entités ayant le choix entre coopérer (c) ou trahir (d pour “Defect”) et qui sont rétribuées par  $R$  points si chacune joue c, par  $P$  points si chacune joue d, et reçoivent respectivement  $T$  et  $S$  points si l’une joue d et l’autre c. On décrit ces règles avec les notations suivantes :

$[c, c] \rightarrow R+R$ ,  $[d, d] \rightarrow P+P$ ,  $[d, c] \rightarrow T+S$ .

Pour que la situation soit celle d’un dilemme, on impose [3] :  $T > R > P > S$  et  $T+S < 2R$ . On choisit le plus souvent les valeurs  $T=5$ ,  $R=3$ ,  $P=1$ ,  $S=0$ .

Le dilemme est itéré quand on imagine que la situation de choix entre c et d, se présente périodiquement aux deux mêmes entités. Jouer consiste alors à choisir une stratégie qui, informée du passé (donc du comportement antérieur de l’adversaire et de son propre comportement) indique comment jouer le coup suivant.

Une multitude d’études [6, 4, 13, 14, 20, 26], conduisent aux conclusions suivantes sur lesquelles un accord général semblait établi.

(a) Il n’y a pas de stratégie meilleure que toutes les autres, mais certaines sont mauvaises dans pratiquement tous des environnements possibles, alors que d’autres sont efficaces et réussissent bien (gagnent beaucoup de points) dans des tournois variés. (b) Les bonnes stratégies sont les stratégies réactives (qui répondent quand on les trahit), prennent le risque de coopérer (elles commencent par coopérer et face à un adversaire qui coopère, elles ne tentent pas de trahir), et savent être indulgentes (après une trahison de l’adversaire elles finissent par pardonner pour renouer la coopération. C’est le cas de *gradual* (voir définition section 6).

### 2.1 Simulation de l’évolution

À côté de l’évaluation des stratégies obtenues en organisant des tournois variés, il existe des méthodes de tests simulant des évolutions auxquelles ne réussissent que les stratégies robustes.

Les résultats obtenus par ces calculs modélisent la sélection naturelle. Ils confirment souvent (mais pas toujours comme on va le voir) ceux des tournois et en accroissent les contrastes. Ils conduisent de plus à une conclusion surprenante : sauf dans de très rares cas, l’arène finit par n’être occupée que par des stratégies qui ne prennent jamais l’initiative de trahir (c’est le cas de *tit\_for\_tat* de *gradual* ou de *pavlov*). Au bout de quelques générations, l’arène est donc occupée par des stratégies qui ne jouent entre elles que des coups  $[c, c]$ . L’arène se trouve donc dans un état de coopération généralisée.

### 2.2 Les classes complètes

Pour mener des tests objectifs qui ne dépendent pas des stratégies repérées bonnes ou robustes, et pour se donner des chances de découvrir de nouvelles stratégies performantes, nous utilisons la *méthode des classes complètes*

[5, 8, 18] qui consiste à regrouper systématiquement toutes les stratégies ayant des capacités équivalentes ou fonctionnant selon un principe abstrait fixé. Nous considérons en particuliers les classes  $Mem(X, Y)$  qui regroupent toutes les stratégies dont le coup  $n$  dépend de manière déterministe des  $X$  derniers coups que la stratégie a joués et des  $Y$  derniers coups que l’adversaire a joués. Tant qu’il n’y a pas  $\max(X, Y)$  coups joués, la stratégie utilise une amorce fixée une fois pour toutes. Une stratégie de  $Mem(1, 2)$  se définit donc par les deux premiers coups qu’elle joue, puis par ce qu’elle fait quand le passé est par exemple  $[d, dc]$  (elle a joué d au coup  $n-1$ , et l’adversaire a joué d au coup  $n-2$ , et c au coup  $n-1$ ). Dans ce cas il y a 8 passés possibles). Nous notons une telle stratégie par un nom du type *mem12\_cdCDDCC* avec la convention que la suite désigne les 2 premiers coups puis les répliques pour les 8 passés possibles pris dans l’ordre lexicographique  $[c cc] [c cd] [c dc] [c dd] [d cc] [d cd] [d dc] [d dd]$ . Le nombre de stratégies de  $Mem(1, 2)$  est  $1024 = 2^{10}$  (il y a dix paramètres binaires à fixer pour chacune) et plus généralement le nombre de stratégies de  $Mem(X, Y)$  est  $2^{\max(X,Y)} \cdot 2^{2^{(X+Y)}}$ .

### 2.3 Une sélection de 21 stratégies

Dans la suite de l’article nous utiliserons entre autres l’ensemble *Select* de 21 stratégies issues de [18] qui peut être considéré à la fois comme contenant les stratégies les plus simples et les stratégies les meilleures identifiées aujourd’hui (voir section 6). On trouvera Figure 1 le résultat de leur confrontation évolutionnaire. Trouver des stratégies qui se classent bien quand on les ajoute à E est un défi difficile.

## 3 Les résultats de Press et Dyson

### 3.1 Une percée théorique

L’article de William Press et Freeman Dyson [21] possède un titre provocateur : “*Iterated Prisoner’s Dilemma contains strategies that dominate any evolutionary opponent*” (“Dans le dilemme itéré du prisonnier, il existe des stratégies qui dominent tout adversaire évolutif”). L’observation quasi universelle d’une convergence vers la coopération généralisée semblait interdire qu’il existe des stratégies dominantes donc exploitant les autres. Les affirmations proposées par [21] proviennent de raisonnements et de calculs mathématiques alors que dans ce domaine, il est difficile d’obtenir des résultats démontrés puisque l’espace associé au problème est infini et discret et n’est muni d’aucune topologie naturelle.

Une question simple se pose naturellement à laquelle les arguments mathématiques ne répondent pas : parmi un ensemble aussi neutre que possible de stratégies probabilistes soumis à un processus évolutif, quelles stratégies probabilistes s’imposent ? C’est la question que nous traitons ici. Nos conclusions rejoignent d’autres conclusions déjà obtenues depuis la parution de l’article de Press et Dyson [1, 2, 10, 12, 11, 16, 19, 25] mais grâce à notre méthode

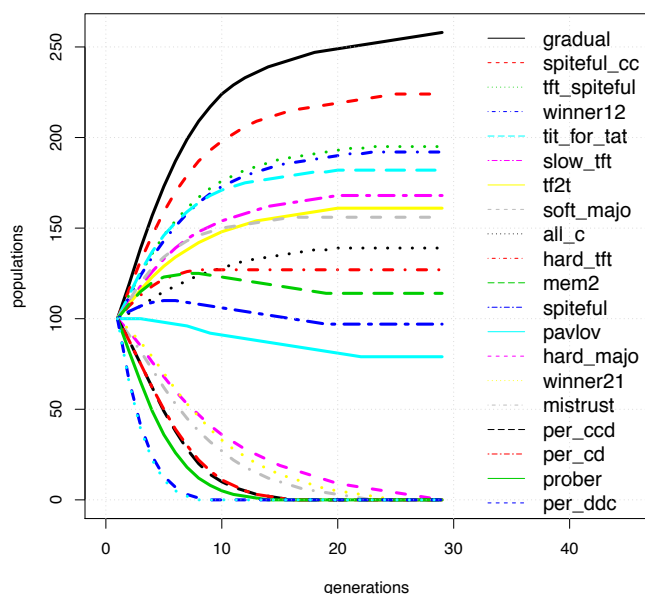


FIGURE 1 – Compétition évolutionnaire des 21 stratégies contenues dans *Select*

systématique, nous parvenons ici à mettre en avant une série de stratégies robustes et efficaces qui n'avait pas été extraites des résultats expérimentaux antérieurs et nous montrons qu'un critère simple permet de les identifier rapidement.

### 3.2 Stratégies probabilistes à mémoire d'un coup

Le travail de Press et Dyson propose deux théorèmes. Le premier théorème concerne le dilemme itéré dans une version limitée aux stratégies probabilistes à mémoire d'un coup : les stratégies *lunatique*, *tit\_for\_tat* et *pavlov* appartiennent à cette catégorie, mais pas la stratégie *gradual* qui pour décider ce qu'elle fait consulte le passé de tous les coups déjà joués.

Une stratégie à mémoire d'un coup est définie par 4 paramètres  $p_1, p_2, p_3, p_4$  qui indiquent la probabilité de jouer  $c$  lorsque le dernier coup a été  $[c\ c]$  ou  $[c\ d]$  ou  $[d\ c]$  ou  $[d\ d]$ . Notons  $\text{proba}(p_1, p_2, p_3, p_4)$  cette stratégie générale. On ne précise pas comment est jouée le premier coup, mais c'est sans importance pour le résultat mathématique qui n'en dépend pas. En pratique pour les simulations, on considérera aussi bien les stratégies dont le premiers coup est  $c$  que celles dont le premier coup est  $d$ .

Press et Dyson considèrent une classe particulière des stratégies  $\text{proba}(p_1, p_2, p_3, p_4)$  dépendant de trois paramètres  $a, b$  et  $c$  et dénommées ZD. Nous les noterons  $ZD(a, b, c)$ .

### 3.3 Les stratégies ZD.

Les équations générales reliant les paramètres  $p_1, p_2, p_3, p_4$  pour les ZD dans le cas  $R=3, S=0, T=5, P=1$  sont :

$$\begin{aligned} p_1 &= 1+3a+3b+c & p_3 &= 5a+c \\ p_2 &= 1+5b+c & p_4 &= a+b+c \end{aligned}$$

Press et Dyson démontrent que lorsqu'on oppose une stratégie ZD  $(a, b, c)$  à une stratégie probabiliste à mémoire d'un coup  $\text{proba}(p_1, p_2, p_3, p_4)$ , et que l'on note  $G_1$  le gain moyen par coup de la première et  $G_2$  le gain moyen par coup de la seconde, alors ces gains moyens vérifient :  $aG_1 + bG_2 + c = 0$ . Les deux stratégies ont des gains liés linéairement l'un à l'autre. Lorsqu'elles se rencontrent,  $\text{proba}(p_1, p_2, p_3, p_4)$  est en quelque sorte contrôlée par  $ZD(a, b, c)$ .

### 3.4 Les stratégies égaliseurs

Lorsque  $a=0$  et que  $b \neq 0$  alors  $G_2 = -c/b$ , autrement dit une stratégie probabiliste quelconque a un gain moyen indépendant des probabilités qui la définissent, gain qui ne dépend que de la stratégie ZD  $(a, b, c)$  qui lui fait face. Une telle stratégie ZD est nommée égaliseur. Les relations deviennent :

$$\begin{aligned} p_1 &= 3b+c+1 & p_3 &= c \\ p_2 &= 5b+c+1 & p_4 &= b+c \end{aligned}$$

et donc  $G_2 = -c/b$ .

Contre une telle stratégie, toutes les stratégies probabilistes à mémoire d'un coup obtiennent le même gain moyen qui est connu d'avance :  $-c/b$ . Inutile de se débattre face à une stratégie égalisateur vous gagnerez  $-c/b$  et pas plus ! Les valeurs possibles pour  $-c/b$  sont toutes les valeurs situées entre  $P$  et  $R$ . Voici les résultats des rencontres entre quelques stratégies connues et l'égaliseur *equa* qui est  $ZD(0, -1/3, 2/3)$  ou encore  $\text{proba}(2/3, 0, 2/3, 1/3)$ . On a donc  $G_2 = -c/b = 2$ . Il force donc son adversaire à avoir un gain moyen de 2.

$$\begin{aligned} \text{equa} &= 2,5 \text{ vs } \text{tit\_for\_tat} = 2 \\ \text{equa} &= 3 \text{ vs } \text{gradual} = 2 \\ \text{equa} &= 3 \text{ vs } \text{all\_c} = 2 \\ \text{equa} &= 1 \text{ vs } \text{all\_d} = 2 \end{aligned}$$

Comme on le voit, *equa* force le gain moyen de l'adversaire, mais cela se fait parfois à ses dépens, elle en sera victime lorsqu'elle jouera contre elle-même !

### 3.5 Les stratégies extorqueurs

Parmi les stratégies ZD découvertes par Press et Dyson certaines opèrent une forme d'extorsion. En effet, si  $c = -(a+b)P$  (donc  $a+b+c=0$  avec  $P=1$ ) on démontre que le gain moyen  $G_1$  de la stratégie ZD contre une autre (obtenant le gain moyen de  $G_2$ ) vérifie  $G_1 - P = X$ . ( $G_2 - P$ ) avec  $X = -b/a$ .

En clair, si la seconde veut gagner plus, donc augmenter  $(G_2 - P)$ , cela entraîne mécaniquement que la stratégie ZD augmente son gain moyen, dont l'écart à  $P$  est toujours  $X$  fois l'écart à  $P$  du gain moyen de la seconde.

Les quatre paramètres définissant ces stratégies extorqueurs sont donnés par les équations :

$$\begin{aligned} p1 &= 2a + 2b + 1 & p3 &= 4a - b \\ p2 &= 4b - a + 1 & p4 &= 0 \end{aligned}$$

Un des gros défauts des stratégies de type extorqueur est que si  $X > 1$  alors elles jouent mal contre elles-mêmes. Vouloir par exemple gagner deux fois plus contre son adversaire (par rapport à  $P$ ) entraîne que face à elle-même elles ne gagneront que  $P$ , ce qui est moins bien que  $C$ .

Quand  $X > 1$ , les extorqueurs ne vous permettent d'avoir un bon résultat que si vous leur en accordez un proportionnellement meilleur :  $G1 - P = X \cdot (G2 - P)$ . Les extorqueurs sont donc des variantes de la stratégie `all_d` : personne ne peut les battre, mais cela a pour conséquence qu'elles prennent le risque de gagner peu.

On notera que les résultats concernant les stratégies égaliseurs avaient déjà été présentées dans [7] avant [21].

### 3.6 Utilité d'une longue mémoire

Le second théorème important de l'article de Press et Dyson indique que si, dans une partie supposée infinie, la stratégie  $A$  est face à une stratégie  $B$  ayant une mémoire de  $k$  coups, il existe alors une stratégie  $A'$  qui obtient le même score moyen face à  $B$  et qui n'a qu'une mémoire de  $k$  coups. La combinaison des deux résultats mathématiques de Press et Dyson conduit à l'affirmation que face à une stratégie égaliseur ou extorqueur ce ne sont pas seulement toutes les stratégies à mémoire un coup qui se trouvent contraintes mais toutes les stratégies à mémoire finie. De là on est tenté de conclure que : *“(a) les stratégies mémorisant plus que le dernier coup sont inutiles. (b) On dispose avec les stratégies ZD de stratégies dominantes dans l'absolu pour le dilemme itéré du prisonnier”*.

Certains ont d'ailleurs interprété ainsi les théorèmes démontrés, et le titre retenu pour leur article suggère que c'est le cas de Press et Dyson. Pourtant la double affirmation sur l'inutilité des stratégies à mémoire étendue et sur la dominance absolue des stratégies ZD est fautive. Considérons d'abord l'affirmation de dominance des stratégies ZD. On sait depuis longtemps qu'au dilemme du prisonnier itéré battre son adversaire (obtenir plus de points que lui) peut se faire aux dépens de celui qui gagne et que celui-ci aurait pu obtenir plus de points en moyenne en acceptant d'être battu. La stratégie `all_d` bat tout autre stratégie (c'est une évidence) et par exemple sur une partie de 100 coups contre `tit_for_tat` obtient 104 points alors que `tit_for_tat` en gagne 99. La stratégie `all_c` ne bat pas `tit_for_tat` et gagne 300 points en 100 coups contre `tit_for_tat` qui gagne aussi 300 points. Contre `tit_for_tat`, `all_d` gagne peut-être, mais elle a tort de gagner car en faisant comme `all_c`, elle obtiendrait un bien meilleur score.

La plupart des stratégies ZD sont dans la même situation : elles ne gagnent contre leur adversaire qu'en renonçant elles-mêmes à avoir de bons scores. En un sens, ce que propose [21] est un ensemble de stratégies généralisant `all_d`. Outre l'erreur de croire que battre son adversaire

c'est gagner des points, un autre oubli a conduit à croire que les stratégies ZD étaient dominantes : pour s'imposer il faut jouer correctement contre soi-même. C'est important dans les tournois, mais plus encore dans les compétitions évolutives. En effet, si vous l'emportez lors des premières générations, l'arène se peuple de stratégies identiques à vous, et vous allez donc très fréquemment les rencontrer. Si vous jouez mal face à vous-même, cela se retourne à terme contre vous. Rien n'est faux dans les résultats mathématiques de Press et Dyson, mais en n'abordant que le problème *“qui gagne dans un combat un contre un ?”* et en oubliant le problème *“combien de points sont gagnés au total ?”* et le problème *“face à toi-même te fais-tu du tort ?”* les théorèmes démontrés ne permettent pas de conclure que les stratégies ZD sont de bonnes stratégies au sens évolutionnaire. Nos expériences de simulations montrent que les stratégies ZD sont mauvaises !

Le résultat de Press et Dyson sur l'inutilité pour une stratégie d'avoir de la mémoire est juste : Si  $A$  est face à une stratégie  $B$  ayant une mémoire de  $k$  coups, il existe une stratégie  $A'$  qui obtient le même score moyen face à  $B$  et qui n'a qu'une mémoire de  $k$  coups. Cependant cela ne signifie pas que face à deux stratégies  $B$  et  $C$  ayant une mémoire de  $k$  coups, il existe une stratégie  $A'$  à mémoire de  $k$  coups qui fasse le même score face à  $B$  et face à  $C$ . En effet, celle  $A'$  qui peut remplacer  $A$  face à  $B$ , n'est pas nécessairement la même que celle,  $A'$ , qui face à  $C$  peut remplacer  $A$ .

À nouveau les simulations confirment que les bonnes stratégies pour des environnements comportant plusieurs stratégies tirent un avantage de l'utilisation d'une large mémoire du passé. Sur les questions de la mémoire utile ou non on pourra consulter [15, 20].

## 4 Stratégies probabilistes à mémoire de un coup

Dans une première série d'expériences nous considérons des ensembles aussi grands que possible et aussi homogènes que possible de stratégies probabilistes à mémoire d'un coup, et nous les soumettons à un processus évolutionnaire.

### 4.1 Expériences évolutionnistes massives

Pour obtenir des ensembles de stratégies probabilistes de la forme `proba(p1, p2, p3, p4)`, nous fixons un pas de variation des paramètres probabilistes. Pour  $K=5$  par exemple nous faisons varier les  $p_i$  dans l'ensemble fini  $0, \frac{1}{5}, \frac{2}{5}, \frac{3}{5}, \frac{4}{5}, 1$  ce qui donne  $2 * 6^4 = 2592$  stratégies (le 2 provient des deux choix possibles pour le coup initial). Nous notons cette classe complète `ProbaCD_K=5` dont on voit l'évolution Figure 2.

Les résultats obtenus pour le tournoi sont indiqués ci-dessous. Nous avons utilisé des parties de 1000 coups et les paramètres usuels. Nous avons calculé les résultats de chaque partie un contre un en la faisant jouer 5 fois, cela de manière à limiter les effets des variations probabilistes.



1	probaD_0.0_0.0_0.0_0.2	39345809
2	probaC_0.0_0.0_0.0_0.2	39264669
3	probaD_0.2_0.0_0.0_0.2	39244163
4	probaD_0.0_0.2_0.0_0.2	39193409
5	probaC_0.2_0.0_0.0_0.2	39145486
6	probaD_0.4_0.0_0.0_0.2	39136783
7	probaC_0.0_0.2_0.0_0.2	39104825
8	probaD_0.2_0.2_0.0_0.2	39069164
9	probaC_0.4_0.0_0.0_0.2	39043576
10	probaD_0.6_0.0_0.0_0.2	38995942

La stratégie classée première est aussi notée  $\text{probaD}(0,0,0,1/5)$  ; le D indique que son premier coup est D (defect) ; l'entier en bout de ligne indique le gain en points lors du tournoi répété 5 fois.

La première stratégie ZD est 34ème, il s'agit de la stratégie équivalente à `spiteful`. Dans les 100 premières, il y a que six stratégies ZD. Sans surprise, les ZD, les extorqueurs et les égaliseurs ne réussissent pas particulièrement bien en tournoi.

Pour la compétition évolutionnaire, les dix premières stratégies avec leurs effectifs finaux sont données ici :

1	probaC_1.0_0.8_0.0_0.0	34262
2	probaC_1.0_0.6_0.0_0.0	31579
3	probaC_1.0_0.4_0.0_0.0	30550
4	probaC_1.0_0.2_0.0_0.0	28640
5	probaC_1.0_0.0_0.0_0.0	27746
6	probaC_1.0_0.0_0.0_0.2	9540
7	probaC_1.0_0.2_0.0_0.2	8893
8	probaC_1.0_0.0_0.2_0.0	8451
9	probaC_1.0_0.2_0.2_0.0	7701
10	probaC_1.0_0.4_0.2_0.0	5984

On notera que la stratégie classée cinquième correspond à `spiteful`

#### Remarque.

Le mécanisme évolutif programmé consiste à remplacer chaque génération par une nouvelle génération dont les effectifs pour un type de stratégies donné sont proportionnels au nombre de points gagnés par les stratégies de ce type lors d'un tournoi général impliquant toutes les stratégies présentes à la génération précédente : la descendance d'un type de stratégies à la génération  $n$  est proportionnelle aux points gagnés par les stratégies de ce type dans le tournoi entre stratégies de la génération  $n-1$ . Au départ, on considère que chaque effectif est 100 et on impose que l'effectif total d'une génération à l'autre reste le même (aux problèmes d'arrondis près). Nous avons aussi mené des tests quand l'effectif d'une stratégie à la génération  $n$  est obtenu en prenant  $a$  fois l'effectif de la stratégie à la génération  $n-1$ , et  $(1-a)$  fois l'effectif donné par le calcul précédent, avec le paramètre  $a$  compris entre 0 et 1. Cela revient à considérer que le remplacement d'une génération par la suivante n'est que partiel : les parents ne meurent pas tout de suite ! Les résultats (classement et effectifs finaux) obtenus sont toujours très proches (si  $a < 1$ ) de ceux qu'on obtient quand le remplacement d'une génération par une autre est total, seule la durée de convergence vers l'état de coopération généralisée s'allonge.

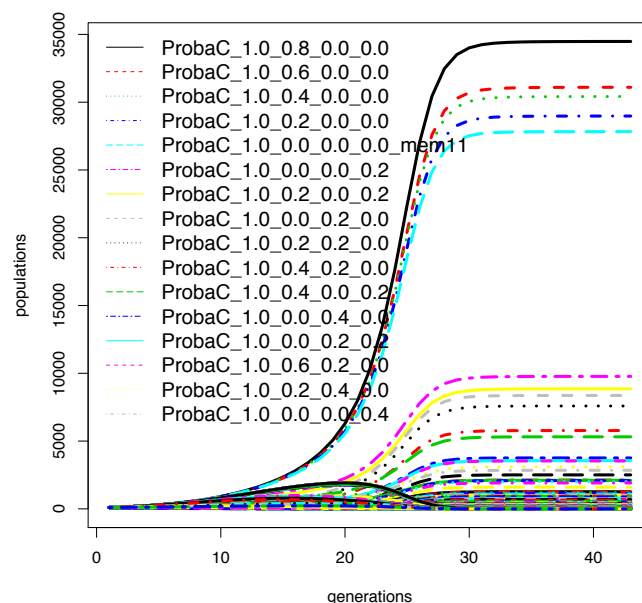


FIGURE 2 – Compétition évolutionnaire des  $\text{ProbaCD}_K=5$

Comme c'est assez souvent le cas au dilemme, les résultats d'un tournoi dans une classe complète qui comporte beaucoup de stratégies médiocres ou mauvaises ne reflètent pas du tout ce qu'on trouve comme résultat de compétitions évolutionnaires. La raison en est simple : les stratégies qui profitent de la présence de mauvaises stratégies dans la soupe initiale, sont rapidement rétrogradées, voire éliminées lorsque les mauvaises disparaissent. Pour réussir dans un processus évolutif, il faut obtenir de bons résultats avec ceux qui obtiennent de bons résultats et qui sont les seuls sur le long terme à survivre. Le résultat d'un tournoi dans une soupe comportant beaucoup de stratégies médiocres n'a guère de sens, seul celui des compétitions évolutives est pertinent.

La gagnante du tournoi est la variante de `all_d` consistant à remplacer le 4<sup>e</sup> paramètre 0.0 en 0.2 (`all_d` elle-même est classée 42ème de ce tournoi). Pour être bien classé au tournoi de la classe complète  $\text{probaCD}_K=5$ , il suffit d'exploiter les stratégies médiocres qui sont très nombreuses ; c'est très facile : il suffit de ne presque jamais coopérer. Cela ne nous apprend rien. Seule le résultat d'un processus évolutif qui commence par faire disparaître les stratégies médiocres présente de l'intérêt.

Dans le cas de la compétition évolutionnaire, ce qu'on découvre est remarquable et n'avait semble-t-il jamais été noté. La meilleure stratégie de notre large ensemble de plus de deux mille stratégies est : `probaC_1.0_0.8_0.0_0.0`. L'effectif initial de 100 est passé à 34262 quand l'état de coopération généralisé s'est installé.

Son comportement est d'une grande simplicité : c'est une

rancunière (une fois qu'elle a commencé à trahir, elle trahit toujours puisque  $p_3$  et  $p_4$  valent 0), mais c'est une rancunière qui réagit sans se précipiter quand on la trahit : en cas de coup  $[c, d]$ , elle ne se met à trahir qu'avec une probabilité de 20%. Autrement dit, dans la phase initiale du jeu, elle coopère et continue de le faire aussi longtemps que l'autre coopère, et lorsqu'elle est trahie dans cette phase initiale, elle pardonne dans 80% des cas. En revanche lorsqu'elle se décide à trahir, il n'y a plus de retour possible à la coopération. Nous nommerons ce type des stratégies des *rancunières patientes*.

Les quatre stratégies suivantes sont elles aussi des rancunières patientes, mais dont la réactivité en cas de trahison augmente : d'abord 40% pour la seconde, puis 60% pour la troisième, puis 80% pour la quatrième, puis 100% pour la cinquième, ce qui donne la rancunière habituelle (spiteful).

La sixième stratégie possède encore un comportement qui s'interprète assez facilement. C'est une rancunière (sans aucune patience puisque  $p_2=0$ ), mais qui une fois dans sa phase de punition mène des tentatives de réconciliation : lorsque le coup qui vient d'être joué est  $[d, d]$ , elle coopère dans 20% des cas, comme si elle disait à son adversaire : "*nous sommes mal partis, je tente un premier pas (dans 20% des cas) pour renouer une meilleure entente*".

Les suivantes sont, elles encore, susceptibles d'interprétations du même type, bien que de plus en plus compliquées. Nous proposons de nommer *rancunières conciliantes* toutes les stratégies de ce type. Pour être précis et mener des décomptes nous nommerons rancunières conciliantes toutes les stratégies  $\text{probaC}(p_1, p_2, p_3, p_4)$  avec  $p_1=1$  et  $p_2+p_3+p_4 \leq 1$ . On notera que *tit-for-tat* et *Pavlov* sont de ce type. Les 37 premières stratégies dans la composition finale de la soupe après stabilisation de la compétition évolutionnaire appartiennent à cette catégorie. Une autre remarque s'impose : dans la compétition évolutionnaire, seules 133 stratégies gardent des effectifs non nuls, et ce sont toutes des stratégies qui commencent par coopérer et qui vérifie  $p_1=1$ . La soupe a donc incontestablement convergé vers un état de coopération généralisé.

Les seules ZD qui survivent sont les suivantes. Elles sont indiquées avec leur classement et leur catégorie :

```
29   probaC_1.0_0.1_0_0.0_ZD_Extorq
34   probaC_1.0_0.6_0.4_0.0_ZD_Extorq
35   probaC_1.0_0.4_0.6_0.0_ZD_Extorq
72   probaC_1.0_0.2_1.0_0.4_ZD
125  probaC_1.0_0.6_0.6_0.4_ZD_Equa
```

La première ZD qui est donc 29è est en fait *tit\_for\_tat* qui est effectivement une ZD de coefficient  $\chi=1$ . Ce  $\chi=1$  signifie qu'en réalité elle n'extorque rien, mais oblige ses adversaires à gagner autant qu'elle, ni plus, ni moins. Ce type de stratégies a parfois été appelée "generous extorquer" [24] et leur capacité à survivre dans une compétition évolutive a été identifiée bien supérieure à celles ayant un  $\chi>1$ . Ce que nous trouvons confirme bien que ce type d'extorqueurs a une certaine capacité à survivre, mais ce

que nous observons est que ce ne sont ni les seules, ni les meilleures !

La seconde ZD qui est 34è est une ZD avec  $a=2/25$ ,  $b=-2/25$ ,  $c=0$ ,  $\chi=1$ . C'est une ZD de coefficient  $\chi=1$  (qui comme *tit\_for\_tat* n'extorque donc rien)

La troisième ZD qui est 35è est une ZD avec  $a=3/25$ ,  $b=-3/25$ ,  $c=0$ ,  $\chi=1$ . C'est encore une ZD de coefficient  $\chi=1$ .

La quatrième ZD, 72è du classement, n'est ni extorqueur, ni égaliseur. C'est une ZD avec  $a=2/25$ ,  $b=-7/25$ ,  $c=3/5$ .

La relation liant le gain moyen qu'elle obtient  $G_1$  et le gain moyen de son adversaire  $G_2$  est :  $aG_1+bG_2+c=0$ . Ce qui donne :  $2G_1+15=7G_2$ . La stratégie dans la situation de coopération généralisée gagne 3 en moyenne par coup comme son adversaire.

La cinquième ZD, 125è du classement, est de type égaliseur avec  $a=0$ ,  $b=-1/5$ ,  $c=3/5$ ,  $-c/b=3$ . C'est une stratégie égaliseur qui oblige son adversaire à gagner 3 points en moyenne par coup, ce qui est le cas pour elle en cas de coopération généralisée.

## 4.2 Reconnaître les stratégies efficaces ?

On le voit les seules extorqueurs ou égaliseurs qui survivent sont en réalité des stratégies qui n'extorquent rien au sens strict puisqu'elles se contentent de gagner comme leur adversaire. Il est remarquable qu'elles soient largement battues par les rancunières patientes et des rancunières conciliantes.

Comme le montre la figure 3, le classement des 133 stratégies dont l'effectif ne s'annule pas est très directement corrélé au paramètre  $p' = p_2+p_3+p_4$ .

Pour anticiper la réussite d'une stratégie probabiliste à un coup de mémoire, le double critère  $p_1=1$  et  $p' = p_2+p_3+p_4$  aussi petit que possible est très efficace.

Nous proposons maintenant une interprétation et une explication de ce double critère. Pour gagner une compétition évolutive ou seulement y réussir correctement, il faut survivre lorsque la coopération généralisée s'installe, il faut donc coopérer avec les stratégies qui coopèrent, d'où le  $p_1=1$ . Il faut aussi être réactif, c'est-à-dire ne pas se laisser faire et adopter un comportement suffisamment sévère pour inciter l'autre à coopérer. La forme de réactivité la plus dure est celle de rancunière  $p_2=p_3=p_4=0$ . Tempérer cette dureté est acceptable, si c'est modérément, et s'interprète de la manière suivante : (a) choisir une valeur non nulle pas trop grande de  $p_2$  revient à ne pas passer systématiquement dans l'état de rétorsion après un coup  $[c, d]$ , mais n'y passer qu'avec une certaine probabilité ; (b) choisir une valeur non nulle pas trop grande de  $p_3$  revient à accepter avec une certaine probabilité après un coup  $[d, d]$  de réessayer de coopérer avec un adversaire qui semble le désirer ; (c) choisir une valeur non nulle pas trop grande de  $p_4$  revient à accepter avec une certaine probabilité après un coup  $[d, d]$  de faire le premier pas dans le but de faire renaître un état de coopération mutuelle.

Combiner ces trois formes de tempérance dans une straté-

gie en adoptant des petites valeurs non nulles de  $p_2$ ,  $p_3$  et  $p_4$  n'est pas absurde à condition que le total de la température introduite dans son comportement ne soit pas trop grand, d'où le critère sur  $p' = p_2 + p_3 + p_4$ .

Dans la figure 4 les stratégies de  $\text{ProbaCD\_K}=5$  ont été groupées en 7 sous-ensembles dont on a calculé le classement moyen génération par génération. Il y a les stratégies pour lesquelles  $p_1 \neq 1$ , puis pour celles avec  $p_1 = 1$ , celles avec  $p'$  dans l'intervalle  $[0; 0,5[$  puis  $[0,5; 1[$ ,  $[1; 1,5[$ ,  $[1,5; 2[$ ,  $[2; 2,5[$ ,  $[2,5; 3]$ .

**Amélioration et variante de  $p'$ .** Si nous calculons le coefficient de corrélation de Spearman entre le rang de classement des stratégies qui terminent avec des effectifs non nuls dans la compétition évolutionnaire  $\text{ProbaCD\_K}=5$  (il y en a 133) et le paramètre  $p'$ , nous trouvons que

$$\text{Cor}(\text{rang}, p') = 0,8805$$

ce qui est une très bonne corrélation. Nous avons cherché systématiquement à améliorer ce paramètre. Le paramètre suivant, encore très simple donne un remarquable résultat.  $p'' = p_2 + p_3/2 + p_4$   $\text{Cor}(\text{rang}, p'') = 0,9411231$ . L'optimal, obtenu par analyse canonique des corrélations (CCA) permet encore une légère amélioration :  $p^* = 0,266 p_2 + 0,138 p_3 + 0,277 p_4$   $\text{Cor}(\text{rang}, p^*) = 0,9413768$ . Les figures 3 et 4 illustrent cette corrélation.

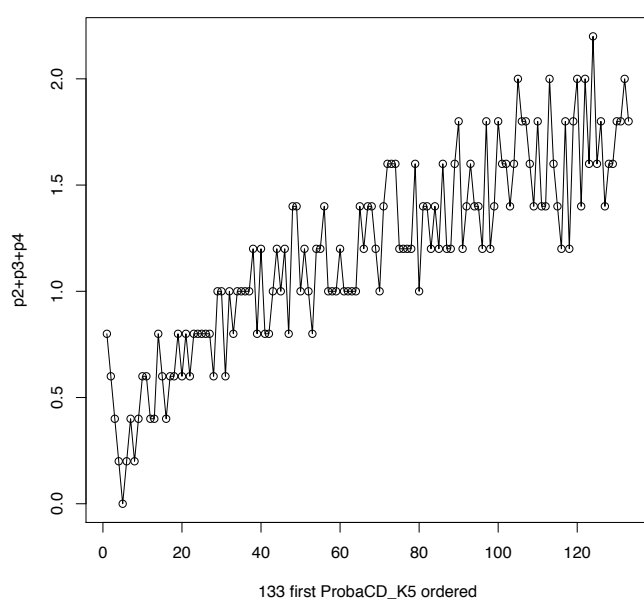


FIGURE 3 – Evolution de  $p'$  en fonction du rang des stratégies de  $\text{probaCD\_K}_5$ . Plus la stratégie est performante, plus  $p'$  est faible.

### 4.3 Robustesse des expériences

Les résultats que nous venons de commenter et analyser sont-ils robustes? Persistent-ils quand on change les paramètres précis de notre expérience avec 2592 stratégies.

C'est ce que nous étudions maintenant.

Dans une autre expérience, nous n'avons mis au départ que des stratégies qui commencent par  $c$  (il y en a donc deux fois moins exactement), ce qui correspond à la classe complète que nous notons  $\text{ProbaC\_K}=5$  le résultat est très proche : les mêmes 5 premières se retrouvent avec quelques permutations du classement.

Avec  $\text{ProbaCD\_K}=4$  nous obtenons des résultats équivalents : en tête les rancunières patientes classées par ordre de patience décroissante ( $p_2=75\%$ ,  $p_2=50\%$ ,  $p_2=25\%$ ,  $p_2=0$ ). Avec  $\text{ProbaCD\_K}=3$  c'est encore la même chose.

Une objection pourrait être faite à notre méthode : les probabilistes composant la soupe initiale sont uniformément réparties (en faisant varier les  $p_i$  par pas constant) : cette régularité pourrait entraîner des résultats spécifiques qui n'auraient donc pas de valeur générale.

Nous avons donc mené des expériences où nous avons choisi 2000 (puis 4000) stratégies au hasard parmi la classe  $\text{ProbaCD\_K}=10$  (qui comporte  $2 \cdot 11^4 = 29282$  stratégies). Les résultats obtenus confirment ceux de l'expérience principale avec  $\text{ProbaCD\_K}=5$  : les gagnantes sont à chaque fois des stratégies de type rancunière patientes ou rancunière conciliante. On y trouve la confirmation du double critère.

Voici un exemple :

- 1 `probaC_1.0_0.0_0.0_0.1`
- 2 `probaC_1.0_0.4_0.0_0.1`
- 3 `probaC_1.0_1.1_0.1_0.0_ZD_Extorq`
- 4 `probaC_1.0_0.3_0.7_0.0_ZD_Extorq`
- 5 `probaC_1.0_0.1_1.1_0.0_ZD_Extorq`
- 6 `probaC_1.0_0.3_0.8_0.0`
- 7 `probaC_1.0_0.3_0.4_0.1`
- 8 `probaC_1.0_0.1_1.1_0.1`
- 9 `probaC_1.0_0.4_0.3_0.1`
- 10 `probaC_1.0_0.0_0.3_0.2`

Les trois extorqueurs qui apparaissent en position 3, 4 et 5 ont toutes un coefficient  $X$  qui vaut 1 (ce ne sont donc pas au sens strict des extorqueurs), et surtout elles sont toutes du type rancunières patientes ou conciliantes.

Une autre objection pourrait être formulée à notre méthode : nous ne considérons pas un assez grand nombre de ZD dans notre soupe initiale. Nous avons donc pris  $\text{ProbaCD\_K}=5$  et ajouté une famille de 880 ZD (toutes celles dont les  $p_i$  sont des multiples de  $\frac{1}{32}$ ). Rien ne change pour l'essentiel : les 5 premières sont exactement les mêmes dans le même ordre que pour la soupe  $\text{ProbaCD\_K}=5$ ; la première ZD est 25ème (à l'exception de rancunière qui est 5ème) et c'est `tit_for_tat : probaC_1.0_0.0_1.0_0.0_ZD_Extorq`.

La ZD suivante dans le classement est ZD  $a=0.03125$   $b=-0.03125$   $c=0.0_ZD_Extorq$  dont le  $X$  vaut 1.

**Robustesse des stratégies nouvelles.** Nous avons voulu savoir si les meilleures stratégies déterministes connues telles que les conclusions de [18] les ont déterminées obtenaient de bons résultats dans ces classes complètes probabilistes.

Nous avons donc mené le calcul pour ProbaCD\_K=5 + Select .

Peu de choses changent concernant les positions relatives des ProbaCD mais les meilleures de Select s'intercalent et prennent de très bonnes places.

1	spiteful_cc	19286
2	tft_spiteful	19078
3	gradual	18235
4	probaC_1.0_0.8_0.0_0.0	17944
5	probaC_1.0_0.6_0.0_0.0	16922
6	probaC_1.0_0.4_0.0_0.0	15759
7	probaC_1.0_0.2_0.0_0.0	14921
8	probaC_1.0_0.0_0.0_0.0	14147
9	mem2	14109
10	spiteful	14073

On note que la huitième correspond à spiteful version probabiliste. Les deux spiteful ne sont pas côte à côte à cause fluctuations statistiques.

Toujours dans le but de tester la robustesse des stratégies identifiées, et en particulier pour voir ce qu'elles donnent quand elles se retrouvent parmi peu de probabilistes, nous avons composé une soupe avec les 20 premières de ProbaCD\_K=5, les 1024 Mem(1, 2) et les stratégies de Select. On obtient :

1	probaC_1.0_0.2_0.0_0.2	5255
2	tft_spiteful	4945
3	probaC_1.0_0.4_0.0_0.2	4877
4	probaC_1.0_0.2_0.0_0.4	4415
5	winner12	4331
6	mem12_CCCDCDDCDD	4331
7	probaC_1.0_0.6_0.0_0.2	4081
8	mem12_CCCDCDDDDD	3557
9	spiteful_cc	3557
10	probaC_1.0_0.8_0.0_0.0	3551

La sixième du classement est winner12, la huitième du classement est spitefull\_CC. La onzième commence par CC et se fâche définitivement quand l'autre trahit deux fois de suite.

Les résultats (et de nombreux autres qui les confirment) sont très clairs : les stratégies les meilleures connues sont toujours bien classées (bien qu'elles proviennent d'expériences et de processus de sélection où seules des stratégies déterministes sont impliquées). Réciproquement, les nouvelles stratégies probabilistes repérées réussissent très bien dans des environnements composés presque exclusivement des stratégies déterministes (comme pour la dernière expérience).

## 5 Conclusion

L'expérimentation approfondie et systématique dans un modèle évolutionniste général et paramétré avec des stratégies de type "probabiliste à mémoire un coup" conduit à des résultats stables et que l'analyse permet d'anticiper une fois les bons paramètres identifiés. La double condition que nous avons extraite et qui semble-t-il n'avait jamais été notée conduit à définir une classe particulière de stratégies probabilistes à mémoire d'un coup (les rancu-

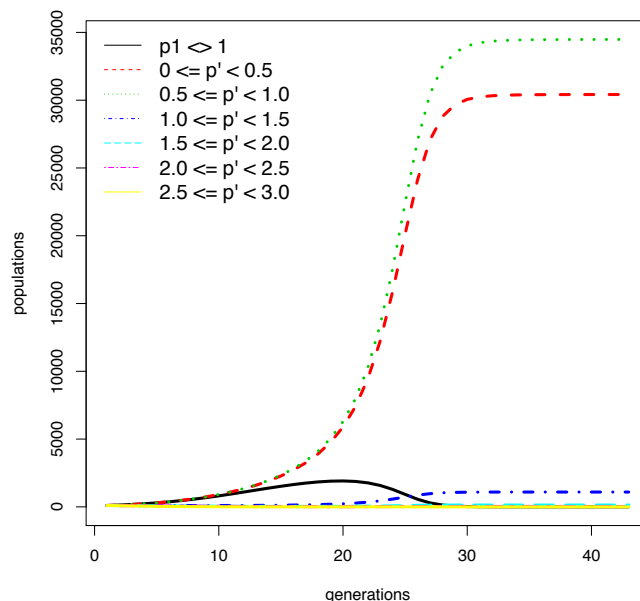


FIGURE 4 – Synthèse par catégorie

nières conciliantes) dont les membres se retrouvent systématiquement en tête de tous les classements des compétitions évolutives que l'on peut imaginer, y compris quand on varie les soupes initiales pour y mettre de nombreuses stratégies extorqueurs, égaliseur ou ZD. De plus les stratégies repérées se montrent robustes et efficaces dans des soupes évolutives composées de manière variées, par exemple ne contenant que des stratégies déterministes. De ce fait, elles rejoignent la famille des meilleures stratégies connues répertoriées dans [18].

## 6 Annexe

Liste des 21 stratégies constituant Select.

1. all\_c (gentille) : je coopère toujours.
2. all\_d (méchante) : je trahis toujours.
3. tit\_for\_tat (donnant-donnant) : je coopère au premier coup, puis je joue au coup  $n$  ce que l'autre a joué au coup  $n - 1$ .
4. spiteful (rancunière) : je coopère au premier coup et tant que l'adversaire coopère, mais dès qu'il trahit, je trahis indéfiniment.
5. soft\_majo (majorité-mou) : je coopère au premier coup et tant que mon adversaire a coopéré plus ou autant qu'il a trahi dans le passé ; sinon je trahis.
6. hard\_majo (majorité-dur) : je trahis au premier coup et tant que mon adversaire a trahi plus ou autant qu'il a coopéré dans le passé ; sinon je coopère.
7. per\_ddc (périodique-ddc) : je joue d, d, c, d, d, c, ...

8. `per_ccd` (périodique-ccd) : je joue c, c, d, c, c, d, ...
9. `mistrust` (méfiante) : je trahis au premier coup, puis je joue au coup  $n$  ce que l'autre a joué au coup  $n - 1$ .
10. `per_cd` (périodique-cd) : je joue c, d, c, d, c, d, ...
11. `pavlov` : Je coopère au premier coup, puis je joue toujours coopérer sauf quand les deux joueurs n'ont pas joué la même chose au coup précédent.
12. `tf2t` (donnant-donnant-mou) : je coopère aux deux premiers coups, puis je coopère toujours au coup  $n$ , sauf si mon adversaire a trahi au coup  $n - 1$  et au coup  $n - 2$ .
13. `hard_tft` (donnant-donnant-dur) : je coopère aux deux premiers coups, puis je coopère toujours au coup  $n$ , sauf si mon adversaire a trahi au coup  $n - 1$  ou au coup  $n - 2$ .
14. `slow_tft` (donnant-donnant-lent) : je coopère aux deux premiers coups, ensuite je me mettrais à trahir lorsque l'autre aura trahi deux fois de suite, et je ne me remettrais à coopérer dès que l'autre aura coopéré deux fois de suite.
15. `gradual` (graduelle) : je coopère au premier coup et quand la règle suivante n'est pas en train de s'appliquer : à chaque fois que l'autre me trahit, je compte le nombre,  $n$ , de ses trahisons passées et je trahis  $n$  fois consécutivement suivi de deux coopérations.
16. `prober` (sondeur) : aux trois premiers coups, je joue trahir-coopérer-coopérer (d-c-d); ensuite, si aux coups 2 et 3 l'adversaire n'a pas trahi, je trahis toujours ; sinon je joue `tit_for_tat`.
17. `mem2` : je commence par jouer deux coups comme `tit_for_tat`; ensuite je change de comportement pour deux coups en fonction des résultats des deux derniers coups, selon les règles : (A) si les deux coups précédents ont été [c, c] [c, c], je passe à `tit_for_tat`; (B) si le dernier coup a été [c, d] ou [d, c] alors je passe à `hard_tft`; (C) dans les autres cas, je passe à `all_d`. De plus, si à un moment l'autre trahit deux fois de suite, je passe définitivement à `all_d` [15].
18. `winner12` (gagnante de `Mem(1,2)`); je coopère aux deux premiers coups puis je joue la table : [C CC]->C [C CD]->D [C DC]->C [C DD]->D [D CC]->D [D CD]->C [D DC]->D [D DD]->D
19. `winner21` (gagnante de `Mem(2,1)`); aux deux premiers coups je joue D C puis je joue la table : [CC C]-> C [CC D]->D [CD C]->C [CD D]->D [DC C]->C [DC D]->D [DD C]->D [DD D]->D
20. `tft spiteful` (donnant-donnant-rancunier) : je joue `tit_for_tat`, sauf que si on me trahit deux fois de suite, je me mets à trahir indéfiniment.

21. `spiteful_cc` (rancunièreCC) : je coopère aux deux premiers coups puis je joue `spiteful`.

## Références

- [1] Christoph Adami and Arend Hintze. Evolutionary instability of zero-determinant strategies demonstrates that winning is not everything. *Nature communications*, 4, 2013.
- [2] Christoph Adami and Arend Hintze. Corrigendum : Evolutionary instability of zero-determinant strategies demonstrates that winning is not everything. *Nature communications*, 5, 2014.
- [3] Robert M Axelrod. *The evolution of cooperation*. Basic books, 2006.
- [4] Bruno Beaufils, Jean-Paul Delahaye, and Philippe Mathieu. Our meeting with gradual, a good strategy for the iterated prisoner's dilemma. In *Proceedings of the Fifth International Workshop on the Synthesis and Simulation of Living Systems, ALIFE V*, pages 202–209. The MIT Press/Bradford Books, 1997.
- [5] Bruno Beaufils, Jean-Paul Delahaye, and Philippe Mathieu. Complete classes of strategies for the classical iterated prisoner's dilemma. In *International Conference on Evolutionary Programming, EP7, Vol1447*, pages 33–41. Springer, 1998.
- [6] Bruno Beaufils and Philippe Mathieu. Cheating is not playing : Methodological issues of computational game theory. In *Proceedings of the 17th European Conference on Artificial Intelligence (ECAI'06)*, 141 :185–189, 2006.
- [7] Maarten C Boerlijst, Martin A Nowak, and Karl Sigmund. Equal pay for all prisoners. *The American mathematical monthly*, 104(4) :303–305, 1997.
- [8] Jean-Paul Delahaye and Philippe Mathieu. Métastratégies pour le dilemme itéré du prisonnier. In *24e Journées Francophones sur les Systèmes Multi-Agents (JFSMA'16)*, pages 13–22. Cépaduès, 2016.
- [9] Hao Dong, Rong Zhi-Hai, and Zhou Tao. Zero-determinant strategy : An underway revolution in game theory. *Chinese Physics B*, 23(7) :078905, 2014.
- [10] Christian Hilbe, Martin A Nowak, and Karl Sigmund. Evolution of extortion in iterated prisoner's dilemma games. *Proceedings of the National Academy of Sciences*, 110(17) :6913–6918, 2013.
- [11] Christian Hilbe, Martin A Nowak, and Arne Traulsen. Adaptive dynamics of extortion and compliance. *PloS one*, 8(11) :e77886, 2013.
- [12] Christian Hilbe, Torsten Röhl, and Manfred Milinski. Extortion subdues human players but is finally punished in the prisoner's dilemma. *Nature communications*, 5, 2014.

- [13] Graham Kendall, Xin Yao, and Siang Yew Chong. *The iterated prisoners' dilemma : 20 years on*. World Scientific Publishing Co., Inc., 2007.
- [14] Jiawei Li, Philip Hingston, and Graham Kendall. Engineering design of strategies for winning iterated prisoner's dilemma competitions. *IEEE Transactions on Computational Intelligence and AI in Games*, 3(4) :348–360, 2011.
- [15] Jiawei Li and Graham Kendall. The effect of memory size on the evolutionary stability of strategies in iterated prisoner's dilemma. *IEEE Transactions on Evolutionary Computation*, 18(6) :819–826, 2014.
- [16] Jie Liu, Y Li, C Xu, and PM Hui. Evolutionary behavior of generalized zero-determinant strategies in iterated prisoner's dilemma. *Physica A : Statistical Mechanics and its Applications*, 430 :81–92, 2015.
- [17] Philippe Mathieu, Bruno Beaufils, and Jean-Paul Delahaye. Studies on dynamics in the classical iterated prisoner's dilemma with few strategies. In *European Conference on Artificial Evolution*, pages 177–190. Springer, 1999.
- [18] Philippe Mathieu and Jean-Paul Delahaye. New winning strategies for the iterated prisoner's dilemma. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*, pages 1665–1666. International Foundation for Autonomous Agents and Multiagent Systems, 2015.
- [19] Manfred Milinski, Christian Hilbe, Dirk Semmann, Ralf Sommerfeld, and Jochem Marotzke. Humans choose representatives who enforce cooperation in social dilemmas through extortion. *Nature communications*, 7, 2016.
- [20] Colm O'Riordan et al. A forgiving strategy for the iterated prisoner's dilemma. *Journal of Artificial Societies and Social Simulation*, 3(4) :56–58, 2000.
- [21] William H Press and Freeman J Dyson. Iterated prisoner's dilemma contains strategies that dominate any evolutionary opponent. *Proceedings of the National Academy of Sciences*, 109(26) :10409–10413, 2012.
- [22] Anatol Rapoport and Albert M Chammah. *Prisoner's dilemma : A study in conflict and cooperation*, volume 165. University of Michigan press, 1965.
- [23] Karl Sigmund. *The calculus of selfishness*. Princeton University Press, 2010.
- [24] Alexander J Stewart and Joshua B Plotkin. From extortion to generosity, evolution in the iterated prisoner's dilemma. *Proceedings of the National Academy of Sciences*, 110(38) :15348–15353, 2013.
- [25] Attila Szolnoki and Matjaž Perc. Defection and extortion as unexpected catalysts of unconditional cooperation in structured populations. *Scientific reports*, 4, 2014.
- [26] Elpidia Tzafestas. Toward adaptive cooperative behavior. In *Proceedings of the Simulation of Adaptive Behavior Conference, Paris*. Citeseer, 2000.

## Coordination distribuée et hors-ligne de planifications locales

Guillaume Desquesnes

Guillaume Lozenguez

Arnaud Doniec

Éric Duviella

IMT Lille Douai, Univ. Lille, URIA, F-59000 Lille, France

prénom.nom@mines-douai.fr

### Résumé

*Une modélisation centralisée n'est pas toujours une solution viable pour les problèmes de planification, à cause de la taille exponentielle de la représentation ou bien à cause de contraintes de vie privée entre les composants du système. Dans ce papier, un modèle distribué de système basé entités est utilisé pour résoudre ces problèmes. Le but est d'optimiser chaque entité en utilisant des actions pouvant affecter d'autres entités. N'ayant qu'une connaissance d'un sous-ensemble du système, les entités affectées par ses actions, un agent doit se coordonner avec d'autres agents pour atteindre une solution localement optimale du contrôle du système. Avec la coordination hors-ligne de planifications locales, chaque agent mettra à jour son propre plan en prenant en compte ceux des autres agents, en utilisant les principes de coévolution et de détection de la terminaison introduits par le Distributed Breakout Algorithm. Une approche générique est introduite pour coordonner la planification hors-ligne de tels problèmes. Cela résulte en chaque agent ayant sa propre politique, ne nécessitant pas de communiquer durant l'exécution. Des résultats expérimentaux sur un réseau de voies navigables utilisant cette approche donnent de bons résultats et montrent une capacité à passer à l'échelle.*

### Mots Clef

Processus décisionnels markoviens, coordination et coopération, planification multi-agent/distribuée.

## 1 Introduction

Dans cet article, nous portons notre attention sur la planification d'un système observé et géré par de multiples agents, incapable d'avoir une vue globale du système pour des raisons, par exemple, de confidentialité ou de complexité. Chaque agent aura une vision locale, avec observabilité complète, limitée à sa partie du problème, avec des informations partagées avec d'autres agents. Nous nous plaçons dans le cas où les agents ne communiquent pas durant leurs exécutions.

L'approche proposée dans cet article s'intéresse en partie à résoudre des problèmes de gestion de l'eau dans des systèmes de voies navigable. Par ailleurs, en Europe, le réseau

de voies navigables est réparti sur plusieurs pays, chaque pays gérant son propre sous-réseau. Ces sous-réseaux sont faiblement connectés. Un certain nombre de biefs, sur la frontière franco-belge, sont affectés par les déplacements d'eau des deux pays donc ces biefs devront être observés par les gestionnaires des deux pays. Dans chaque pays, l'obtention d'une planification sera dépendante des actions de ses voisins. Il sera donc nécessaire de se coordonner a priori pour atteindre des buts communs sans communication durant l'exécution.

Cet article introduit des mécanismes de coordination hors-ligne pour des systèmes basés sur les entités divisées en multiples agents, inspirés par l'algorithme LID-JESP introduit pour les ND-POMDP (Nair *et al.*, 2005) et le Distributed Breakout Algorithm (Yokoo & Hirayama, 1996). Il utilise une modélisation basée sur les processus de décision markovien (MDP (Puterman, 1994), Factored-MDP (Boutilier *et al.*, 1995), ...) pour la modélisation et planification locale. L'utilisation de mécanismes de communications décentralisés et de méthodes d'évolutions liées au problème permettra à la politique jointe de converger vers une solution localement optimale. Comme chaque agent possède son propre MDP, il obtiendra une politique qui ne nécessitera pas de communication lors de son utilisation.

Ce papier est organisé de la façon suivante : tout d'abord, nous introduisons notre problématique et rappelons les bases des processus de décision markovien (MDP) et du Distributed Breakout Algorithm. Puis, nous introduisons formellement le problème et présentons l'algorithme pour la coordination de plusieurs agents. Nous montrons qu'il est possible de prouver la convergence de l'algorithme. Après coup, un exemple illustrera l'algorithme proposé. Nous terminons en montrant des résultats expérimentaux de l'application de l'algorithme pour optimiser la gestion de l'eau dans de grands réseaux de voies navigables.

## 2 Présentation du problème

De nombreuses applications de taille importante sont naturellement distribuées et donc leur système peut être divisé en un ensemble (noté  $\omega$ ) de sous-systèmes semi-indépendant, appelé entités dans cet article. Une entité est définie comme un élément du système possédant un état

$s_e$  qui appartient à un ensemble fini d'états  $S_e$ . Ici, l'évolution de l'état d'une entité ne dépend pas directement de l'état des autres entités.

Chaque entité  $e$  est affectée par un ou plusieurs points d'action du système  $\{A_{e_1}, \dots, A_{e_k}\}$ . Similairement, un point d'action  $A_i$  affectera une ou plusieurs entités. Les actions dans ces systèmes distribués ont des impacts locaux. Cependant, ces impacts locaux peuvent être chaînés empêchant donc de facilement distribuer le processus de planification. L'évolution de l'état de l'entité  $e$  ne dépend donc que de sa valeur courante  $s_e$  et de l'action jointe choisie (tuple d'actions réalisées simultanément sur chaque point d'action). Ainsi, il est possible d'exprimer la probabilité d'atteindre un état  $s'$  :

$$p(s'_e | s_e, a_{e_1}, \dots, a_{e_k}) \quad (1)$$

Toute entité  $e$  possède un ensemble d'états optimaux ainsi qu'une fonction de distance  $\delta_e$  donnant l'écart à l'état optimal le plus proche, voir équation 2. Le but est de choisir une politique jointe minimisant la distance de chaque entité relativement à ses états optimaux. À horizon 1, le but est de trouver pour chaque état  $s$  du système l'action jointe  $a$  minimisant l'équation 3.

$$\delta_e(s_e) = |s_e^* - s_e| \quad (2)$$

$$\sum_{e \in \omega} \sum_{s'_e \in S_e} \left( p(s'_e | s_e, a_{e_1}, \dots, a_{e_k}) \times \delta_e(s'_e) \right) \quad (3)$$

À cause du nombre de combinaisons sur les actions jointes, les systèmes distribués et chaînés sont difficiles à contrôler optimalement. Le problème est d'autant plus complexe en considérant les incertitudes (l'état résultant d'une action dépend de probabilités).

## 2.1 Modélisation à horizon infini

Le contrôle optimal d'un système complexe à entités ( $\omega$ ) se ramène à la résolution d'un processus décisionnel de Markov (MDP) (Puterman, 1994) (Markov Decision Process - MDP). Un MDP est un framework générique modélisant les possibilités de contrôle d'un système stochastique dynamique sous forme d'un automate probabiliste. Il est défini par un tuple  $\langle S, A, T, C \rangle$  avec  $S$  et  $A$  respectivement l'ensemble d'états et d'actions qui définissent le système et ses capacités de contrôle.  $T$  est la fonction de transition définie par  $T : S \times A \times S \rightarrow [0, 1]$ .  $T(s, a, s')$  est la probabilité d'atteindre l'état  $s'$  après avoir fait l'action  $a$  dans l'état  $s$ . La fonction de coût  $C$  est définie par  $C : S \times A \times S \rightarrow \mathbb{R}$ .  $C(s, a, s')$  donne le coût pour avoir atteint l'état  $s'$  en faisant l'action  $a$  depuis l'état  $s$ .

Ici, l'ensemble des états du système  $S_\omega$  correspond au produit cartésien des états locaux de chaque entité et les actions  $A_\omega$  au produit cartésien des actions pour les points d'actions. Les fonctions de transition et de coût s'appuient sur les probabilités de transition et l'état optimal pour

chaque entité.

$$T_\omega(s, a, s') = \prod_{e \in \omega} p(s'_e | s_e, a_{e_1}, \dots, a_{e_k}) \quad (4)$$

$$C_\omega(s, a, s') = C_\omega(s') = \sum_{e \in \omega} \delta_e(s'_e) \quad (5)$$

Résoudre un MDP consiste à construire une politique d'actions optimales. Une politique est une fonction  $\pi : S_\omega \rightarrow A_\omega$  qui assigne une action à chaque état du système. Une politique optimale  $\pi^*$  est une politique qui minimise le coût espéré, en minimisant l'équation de Bellman (Bellman, 1957) :

$$V_\omega^\pi(s) = \sum_{s' \in S} T_\omega(s, \pi(s), s') \times (C_\omega(s, \pi(s), s') + \gamma \times V_\omega^\pi(s')) \quad (6)$$

$$\pi^*(s) = \arg \min_{a \in A} \left( \sum_{s' \in S} T_\omega(s, a, s') \times (C_\omega(s, a, s') + \gamma \times V_\omega^{\pi^*}(s')) \right) \quad (7)$$

## 2.2 Une approche multi-agent

En utilisant la modélisation basée agent, un agent sera responsable d'un sous-ensemble de points d'action du système modélisé. Il observera uniquement les entités affectées par ses actions possibles pour prendre une décision. Sa politique sera calculée pour être coordonnée avec les politiques de ces voisins. Ce framework promet de résoudre façon heuristique des problèmes intraitables en utilisant des algorithmes distribués, basés sur des processus décisionnels de Markov pour la planification locale.

Le problème de coordination de plan n'est pas nouveau. Le Distributed Breakout Algorithm (Yokoo & Hirayama, 1996) (DBA) est un algorithme pour résoudre des problèmes de satisfaction de contraintes (CSP) distribués. Pour le DBA, chaque agent essaye d'optimiser son évaluation (le nombre de contraintes violées) en échangeant son évaluation actuelle et son amélioration par rapport à l'évaluation précédente à son voisinage. Deux agents voisins n'ont pas le droit de se mettre à jour en même temps afin d'éviter les changements conflictuels. Un mécanisme de communication est utilisé pour détecter la convergence globale. Cette solution a été adaptée aux Network Distributed Partially Observable MDP (Nair *et al.*, 2005) qui se basent sur des transitions ou observations indépendantes entre les agents pour résoudre des problèmes avec un nombre d'agents important. Un algorithme similaire (Chades *et al.*, 2002) a aussi été proposé pour faciliter la résolution de Dec-POMDP avec une forte contrainte d'un ensemble d'états et d'une fonction de transition globaux.

Pour des systèmes complexes à base d'entité, la somme totale des distances aux états optimaux des entités est distribuable. Les agents ont pour but de minimiser la somme des fonctions de distance de chaque entité observée vers leurs valeurs optimales. L'agent  $\alpha$  aura à minimiser sa fonction



objective locale  $V_\alpha$ , basée sur la somme des distances des entités gérées par  $\alpha$ .

La difficulté est d'estimer l'état d'une entité dont l'agent n'a qu'un contrôle partiel des actions qui l'affectent. Deux agents observant une même entité sont définis comme étant voisins. Cependant, les agents n'ont pas forcément une connaissance complète de l'évolution des entités partagées avec leurs voisins. En effet, les politiques exactes des voisins ne sont pas exprimables sur la base des seules entités observées par l'agent. En minimisant chaque fonction objective locale, le but est d'atteindre un optimum local de la fonction objective globale.

### 3 Présentation de l'algorithme local

L'algorithme local proposé, appelé OCLP (Offline Coordination of Local Planning - Coordination hors-ligne de planifications locales), a pour but de trouver une politique jointe localement optimale de manière distribuée. Il est initialement inspiré par l'algorithme LID-JESP (Nair *et al.*, 2005) pour les ND-POMDP et du DBA (Yokoo & Hirayama, 1996) pour optimiser la gestion de grands réseaux de voies navigables en utilisant des MDP discrétisés. Cet algorithme se base sur des agents avec une vision locale qui peuvent être affectés par le choix d'autres agents.

**Algorithme** Coordination hors-ligne de planifications locales pour un agent  $\alpha$

- 
- 1: Créer une politique locale initiale  $\pi_{\alpha_0} : S_\alpha \rightarrow A_\alpha$
  - 2:  $d \leftarrow$  distance maximale entre deux agents du système
  - 3:  $it \leftarrow 0$
  - 4: **répéter** :
  - 5: Adapter et échanger  $\pi_{\alpha_{it}}$  avec les voisins (voir équation 8)
  - 6: Mettre à jour  $MDP_{\alpha_{it}} = \langle S_\alpha, A_\alpha, T_{\alpha_{it}}, R_\alpha \rangle$  en  $MDP_{\alpha_{it+1}}$  grâce aux politiques reçues
  - 7: Construire  $\pi'_{\alpha_{it}}$  la politique optimale de  $MDP_{\alpha_{it+1}}$
  - 8:  $g_{\alpha_{it}} \leftarrow \text{gain}(\pi'_{\alpha_{it}}, \pi_{\alpha_{it}})$  de  $MDP_{\alpha_{it+1}}$
  - 9: Échanger  $g_{\alpha_{it}}$  avec chaque voisin ;  $G_{\alpha_{it}} \leftarrow$  gains du voisinage
  - 10:  $\pi_{\alpha_{it+1}} \leftarrow \pi'_{\alpha_{it}}$  si  $g_{\alpha_{it}} = \max \text{disponible}(G_{\alpha_{it}})$  et  $\pi_{\alpha_{it}}$  sinon
  - 11:  $counter \leftarrow d$  si  $g_{\alpha_{it}} > 0$  sinon  $counter - 1$
  - 12: Échanger  $counter$  avec chaque voisin
  - 13:  $Counters_{it} \leftarrow$  compteurs du voisinage
  - 14:  $counter \leftarrow \max(Counters_{it})$
  - 15:  $it \leftarrow it + 1$
  - 16: **jusqu'à** :  $counter = 0$
- 

e À chaque itération  $it$ , les agents échangeront des adaptations de leur politique actuelle avec chacun de leurs voisins, adaptées à la connaissance commune des deux agents (ligne 5). Une politique de  $\beta$  adaptée à  $\alpha$ , voir l'équation 8, correspond à une assignation d'un ensemble de couples action - probabilité à chaque état partagé par  $\alpha$  et  $\beta$ .

$$\pi_{\alpha}^{\beta}(s_{\alpha\beta}) = \{(\pi_{\beta}(s_{\beta}), p(s_{\beta}|s_{\alpha\beta})), \forall s_{\beta} \in S_{\beta}\} \quad (8)$$

où  $s_{\alpha\beta}$  est une assignation d'états aux entités supervisées à la fois par  $\alpha$  et  $\beta$ .

Lorsqu'un agent obtient toutes les estimations de politique de son voisinage, il sera capable de changer son modèle  $MDP_{it}$ . Ainsi, il prendra en compte le nouvel impact du voisinage sur ses entités partagées via la mise à jour de sa fonction de transition de façon à refléter les actions probables du voisinage (ligne 6).

Toute itération pourra mettre à jour les parties du modèle affectant les entités partagées de chaque agent. L'utilisation de ce nouveau modèle,  $MDP_{it+1}$ , permettra aux agents d'obtenir une nouvelle politique optimale  $\pi'_{it}$  (ligne 7). L'amélioration de la politique optimale ( $\pi'_{it}$ ) par rapport à la politique actuelle ( $\pi_{it}$ ) dans le nouveau modèle ( $MDP_{it+1}$ ) est évaluée grâce à une fonction heuristique (ligne 8). Cette fonction heuristique a pour but de guider l'exploration de la résolution et de détecter la convergence. Seul l'agent avec la plus grande amélioration par rapport à son voisinage gardera sa nouvelle politique, tandis que les autres agents l'ignoreront (ligne 10). Afin d'éviter de n'avoir qu'un seul agent gardant sa politique à chaque itération, la valeur d'amélioration d'agents bloqués par d'autres voisinages est ignorée. Ne garder qu'une seule nouvelle politique par voisinage permet d'éviter l'adoption des changements conflictuels lors d'une même itération.

**Proposition** L'algorithme terminera, au plus en  $d = \max_{\alpha, \beta \in Ag^2} \text{pathSize}(\alpha, \beta)$  itérations si et seulement si tous les agents sont dans un optimum local.  $\text{pathSize}$  est défini comme :

$$\text{pathSize}(\alpha, \beta) = \begin{cases} 1 & \text{si } \beta \in N(\alpha) \\ 1 + \min_{\gamma \in N(\alpha)} \text{pathSize}(\gamma, \beta) & \text{sinon} \end{cases} \quad (9)$$

avec  $N(\alpha)$  l'ensemble des agents voisins de  $\alpha$  et  $Ag$  l'ensemble des agents.

Pour garantir que tous les agents s'arrêtent de calculer en même temps dès qu'ils atteignent un optimum local, un compteur est utilisé. Chaque agent initialisera son compteur à la même valeur :  $d > 0$ . La valeur de  $d$  est assignée une fois pour toutes au début de l'algorithme par un agent externe et correspond à la taille du chemin maximal entre deux agents du système. À la fin d'une itération, si une nouvelle politique, différente de la politique actuelle est disponible, alors le compteur de l'agent sera réinitialisé à  $d$ . Autrement, le compteur sera diminué de 1 (ligne 11). Ensuite, les agents échangeront leur compteur avec leurs voisins et ne garderont que le plus grand (ligne 14). Un agent aura terminé sa résolution lorsque son compteur atteindra 0 (ligne 16).

Pour cet algorithme, les communications sont synchronisées puisque la continuation de l'algorithme après l'échange des politiques, compteurs et gains dépend des réponses du voisinage.

### 4 Preuve

La coordination hors-ligne de planifications locales possède certaines garanties notamment sur la convergence vers une solution localement optimale ainsi que

sur la terminaison de l'algorithme après avoir fait  $\max_{\alpha, \beta \in Ag^2} pathSize(\alpha, \beta)$  itérations une fois qu'une solution localement optimale est atteinte.

#### 4.1 Preuve de terminaison

Supposons que l'agent  $\alpha$  ne commence pas l'itération  $it$ , car il a terminé sa résolution à l'itération  $i - 1$  ( $counter_\alpha(it - 1) = 0$ ), mais que d'autres agents ne sont pas dans un optimum local. Cela implique qu'à l'itération  $i - d$ , il existe au moins un agent  $\beta$  qui peut améliorer sa politique et ainsi réinitialiser son compteur ( $counter_\beta(it - d) = d$ ). Les compteurs étant décrémentés d'au plus 1 à chaque itération et étant propagés à travers le voisinage, alors à l'itération  $it - d + pathSize(\alpha, \beta)$  la borne minimale du compteur de l'agent  $\alpha$  peut être définie par  $it - d + pathSize(\alpha, \beta)$ . De ce fait,  $counter_\alpha(it - 1) \geq d - pathSize(\alpha, \beta) + 1 - d + pathSize(\alpha, \beta) = 1$ . Donc l'agent  $\alpha$  aura besoin d'effectuer l'itération  $i$ . Par contradiction, si l'algorithme termine pour un agent, alors tous les agents sont dans un optimum local.

À l'inverse, si tous les agents ont atteint un optimum local, les compteurs ne seront jamais réinitialisés à  $d$  et diminueront de 1 à chaque itération. Donc après  $d$  itérations,  $\forall \alpha, counter_\alpha = 0$  et les agents s'arrêtent.

#### 4.2 Preuve de convergence

Pour prouver la convergence de l'algorithme, il va être montré par récurrence que la somme des coûts espérés pour chaque état du système diminue à chaque itération :

$$\sum_{s \in S_\omega} V_\omega^{\pi_\omega^{it}}(s) \geq \sum_{s \in S_\omega} V_\omega^{\pi_\omega^{it+1}}(s) \quad (10)$$

**Cas mono-agent.** Si le système est modélisé par un seul agent, alors  $\alpha = \omega$  et le MDP local de l'agent  $\alpha$  est complet donc par définition :

$$\sum_{s \in S_\omega} V_\omega^{\pi_\omega^{it}}(s) \geq \sum_{s \in S_\omega} V_\omega^{\pi_\omega^{it+1}}(s) \quad (11)$$

En l'occurrence, l'agent trouvera sa politique optimale globale lors de la première et unique itération de l'OCLP puisque la coordination n'est pas requise.

**Cas multi-agent.** En supposant qu'un ensemble de  $n$  agents soit garantie de converger, il est possible de prouver que l'ajout d'un agent supplémentaire, avec un certain ensemble d'entités à contrôler, gardera cette garantie. Dans la preuve suivante, le  $(n + 1)^{\text{ème}}$  agent est dénoté  $\alpha$  tandis que l'ensemble de  $n$  agents est représenté par un méta-agent  $\beta$ .  $\omega$  correspond donc à l'ensemble des entités supervisées par  $\alpha$  et  $\beta$ .

**Si  $\alpha$  met à jour sa politique** Dans le cas où  $\alpha$  met à jour sa politique, lorsque son gain est le maximum de son voisinage (ligne 10 de l'OCLP), à l'itération  $it$  ( $\pi_\alpha^{it+1} = \pi_\alpha^{it}$ ), alors comme aucun agent voisin avec  $\alpha$  ne changera de politique durant cette itération :

$$\sum_{s \in S_\alpha} V_\alpha^{\pi_\alpha^{it}}(s) \geq \sum_{s \in S_\alpha} V_\alpha^{\pi_\alpha^{it+1}}(s) \quad (12)$$

Le méta-agent  $\beta$  possède des agents qui seront influencés par le changement de politique de  $\alpha$ , il est donc impossible d'établir une relation entre la somme des coûts espérés aux itérations  $it$  et  $it + 1$ . Néanmoins, grâce à l'hypothèse de convergence du méta-agent  $\beta$ , il est possible d'affirmer que le nouvel ensemble de politiques du méta-agent diminue la somme des coûts espérés à l'itération  $it$  :

$$\sum_{s \in S_\beta} V_\beta^{\pi_\beta^{it}}(s) \geq \sum_{s \in S_\beta} V_\beta^{\pi_\beta^{it+1}}(s) \quad (13)$$

L'évolution d'une entité ne dépendant que de son état et du choix d'action, cela rend la fonction de valeur décomposable, d'où :

$$V_\mu^{\pi_\mu^{it}}(s) = \sum_{e \in \mu} V_e^{\pi_\mu^{it}}(s) \quad \forall \mu \in Ag, \forall s \in S_\mu \quad (14)$$

Soit  $\alpha \cap \beta$  l'ensemble des entités partagées par  $\alpha$  et  $\beta$ . Puisque l'agent  $\alpha$  garde sa politique à l'itération  $it$  alors tout agent du méta-agent  $\beta$  supervisant au moins une des entités de  $\alpha \cap \beta$  ne changera pas de plan durant cette itération. Ceci implique :

$$\sum_{e \in \alpha \cap \beta} V_e^{\pi_\alpha^{it}}(s) = \sum_{e \in \alpha \cap \beta} V_e^{\pi_\alpha^{it+1}}(s), \quad \forall s \in S_\beta \quad (15)$$

$$\sum_{e \in \beta \setminus \alpha} V_e^{\pi_\alpha^{it}}(s) = \sum_{e \in \beta \setminus \alpha} V_e^{\pi_\alpha^{it+1}}(s), \quad \forall s \in S_\beta \quad (16)$$

d'où

$$V_\beta^{\pi_\beta^{it}}(s) = \sum_{e \in \beta \setminus \alpha} V_e^{\pi_\beta^{it+1}}(s) + \sum_{e \in \alpha \cap \beta} V_e^{\pi_\beta^{it}}(s), \quad \forall s \in S_\beta \quad (17)$$

donc en combinant l'équation 17 à l'équation 13 :

$$\sum_{s \in S_\beta} V_\beta^{\pi_\beta^{it}}(s) \geq \sum_{s \in S_\beta} V_\beta^{\pi_\beta^{it+1}}(s) \quad (18)$$

$$\sum_{s \in S_\beta} \left( \sum_{e \in \alpha \cap \beta} V_e^{\pi_\beta^{it}}(s) + \sum_{e \in \beta \setminus \alpha} V_e^{\pi_\beta^{it}}(s) \right) \geq \sum_{s \in S_\beta} \left( \sum_{e \in \alpha \cap \beta} V_e^{\pi_\beta^{it+1}}(s) + \sum_{e \in \beta \setminus \alpha} V_e^{\pi_\beta^{it+1}}(s) \right) \quad (19)$$

$$\sum_{s \in S_\beta} \left( \sum_{e \in \beta \setminus \alpha} V_e^{\pi_\beta^{it}}(s) \right) \geq \sum_{s \in S_\beta} \left( \sum_{e \in \beta \setminus \alpha} V_e^{\pi_\beta^{it+1}}(s) \right) \quad (20)$$

L'addition des équations 12 et 13 peut alors se simplifier

en :

$$\begin{aligned} \sum_{s \in S_\alpha} V_\alpha^{\pi_\alpha^{it}}(s) + \sum_{s \in S_\beta} V_\beta^{\pi_\beta^{it}}(s) \\ \geq \sum_{s \in S_\alpha} V_\alpha^{\pi_\alpha^{it+1}}(s) + \sum_{s \in S_\beta} V_\beta^{\pi_\beta^{it+1}}(s) \end{aligned} \quad (21)$$

$$\begin{aligned} \sum_{s \in S_\alpha} \left( \sum_{e \in \alpha} V_e^{\pi_\alpha^{it}}(s) \right) + \sum_{s \in S_\beta} \left( \sum_{e \in \beta \setminus \alpha} V_e^{\pi_\beta^{it}}(s) \right) \\ \geq \sum_{s \in S_\alpha} \left( \sum_{e \in \alpha} V_e^{\pi_\alpha^{it+1}}(s) \right) + \sum_{s \in S_\beta} \left( \sum_{e \in \beta \setminus \alpha} V_e^{\pi_\beta^{it+1}}(s) \right) \end{aligned} \quad (22)$$

et donc, grâce à l'indépendance entre les entités :

$$\begin{aligned} \sum_{s \in S_\omega} \left( \sum_{e \in \alpha} V_e^{\pi_\alpha^{it}}(s) + \sum_{e \in \beta \setminus \alpha} V_e^{\pi_\beta^{it}}(s) \right) \\ \geq \sum_{s \in S_\omega} \left( \sum_{e \in \alpha} V_e^{\pi_\alpha^{it+1}}(s) + \sum_{e \in \beta \setminus \alpha} V_e^{\pi_\beta^{it+1}}(s) \right) \end{aligned} \quad (23)$$

$$\sum_{s \in S_\omega} \left( \sum_{e \in \omega} V_e^{\pi_\omega^{it}}(s) \right) \geq \sum_{s \in S_\omega} \left( \sum_{e \in \omega} V_e^{\pi_\omega^{it+1}}(s) \right) \quad (24)$$

$$\sum_{s \in S_\omega} V_\omega^{\pi_\omega^{it}}(s) \geq \sum_{s \in S_\omega} V_\omega^{\pi_\omega^{it+1}}(s) \quad (25)$$

□

**Si  $\alpha$  ne met pas à jour sa politique** Lorsque l'agent  $\alpha$  ne change pas sa politique lors de l'itération  $it$ , alors par définition :

$$\sum_{s \in S_\beta} V_\beta^{\pi_\beta^{it}}(s) \geq \sum_{s \in S_\beta} V_\beta^{\pi_\beta^{it+1}}(s) \quad (26)$$

$$\sum_{s \in S_\alpha} V_\alpha^{\pi_\alpha^{it}}(s) = \sum_{s \in S_\alpha} V_\alpha^{\pi_\alpha^{it+1}}(s) \quad (27)$$

d'où

$$V_e^{\pi_\alpha^{it}}(s) = V_e^{\pi_\alpha^{it+1}}(s) \quad \forall e \in \alpha \setminus \beta, \forall s \in S_\alpha \quad (28)$$

donc en suivant un raisonnement similaire à  $\alpha$  changeant de politique :

$$\begin{aligned} \sum_{s \in S_\alpha} \left( \sum_{e \in \alpha \setminus \beta} V_e^{\pi_\alpha^{it}}(s) \right) + \sum_{s \in S_\beta} \left( \sum_{e \in \beta} V_e^{\pi_\beta^{it}}(s) \right) \\ \geq \sum_{s \in S_\alpha} \left( \sum_{e \in \alpha \setminus \beta} V_e^{\pi_\alpha^{it+1}}(s) \right) + \sum_{s \in S_\beta} \left( \sum_{e \in \beta} V_e^{\pi_\beta^{it+1}}(s) \right) \end{aligned} \quad (29)$$

$$\sum_{s \in S_\omega} \left( \sum_{e \in \omega} V_e^{\pi_\omega^{it}}(s) \right) \geq \sum_{s \in S_\omega} \left( \sum_{e \in \omega} V_e^{\pi_\omega^{it+1}}(s) \right) \quad (30)$$

$$\sum_{s \in S_\omega} V_\omega^{\pi_\omega^{it}}(s) \geq \sum_{s \in S_\omega} V_\omega^{\pi_\omega^{it+1}}(s) \quad (31)$$

□

Lorsque  $n$  agents arrivent à converger entre eux, l'ajout d'un agent supplémentaire maintiendra cette propriété que ce dernier change ou garde sa politique.

**Conclusion de la preuve de convergence.** Il a été prouvé qu'un agent seul convergera toujours vers l'optimum global. De plus, ajouter un agent à un groupe d'agents pouvant converger vers un optimum local maintiendra cette propriété, alors par récurrence l'algorithme OCLP convergera toujours vers un optimum local.

## 5 Exemple d'utilisation de l'algorithme basée sur un CSP

Cet algorithme a été introduit pour résoudre des problèmes de partage de ressources entre plusieurs entités utilisant des actions distribuées. Un exemple illustratif est ici présenté, modélisant une répartition d'objets identiques entre plusieurs chariots, les entités du système, représentés par des carrés sur la figure 1. Chaque chariot a une évaluation différente de sa charge, voir équation 32, définissant donc une valeur optimale. Chaque chariot a un remplissage initial et une capacité maximale (indiqués dans chaque carré par « initial/maximale »). Il n'est possible de déplacer les objets que de chariot en chariot, uniquement dans le sens indiqué par les flèches et dans les capacités définies par les intervalles à proximité. Dans cet exemple, les actions sont considérées comme instantanées et simultanées.

Trois agents,  $\alpha, \beta, \gamma$ , vont gérer les transferts, supposés simultanés d'objets entre les différents chariots qu'ils contrôlent. L'agent  $\alpha$  contrôle le transfert  $\{o\}$ , ce qui lui donne la vision des chariots  $\{0, 1\}$ .  $\beta$  et  $\gamma$  contrôlent respectivement  $\{p, q\}$  et  $\{r\}$  avec vision respectivement de  $\{1, 2, 3\}$  et de  $\{2, 4\}$ . Le chemin maximal entre deux agents dans ce système est de taille  $d = 2$ .

Un agent ne peut déplacer un objet d'un chariot à un autre que s'il sait que le chariot source n'est pas vide et que le chariot destination n'est pas rempli. Chaque agent a donc un simple problème sous contraintes, visant à minimiser la somme des distances des entités qu'il supervise, voir équation 32, qui sera utilisé en combinaison de l'algorithme présenté comme outil de planification à la place d'un MDP, dans un souci de simplification.

$$\begin{aligned} s_0^* = 4 & & s_2^* = 3 & & s_3^* = 0 & & s_4^* = 0 \\ \delta_0(s_0) & = & 4 - s_0 & & \delta_3(s_3) & = & |0 - s_3| \\ \delta_1(s_1) & = & 0 & & \delta_4(s_4) & = & (0 - s_4)^2 \\ \delta_2(s_2) & = & |3 - s_2| & & & & \end{aligned} \quad (32)$$

Le but étant donc de minimiser la somme  $C_\omega$  de ces distances. Dans ce scénario, le résultat des actions est déterministe.

L'état initial du système est  $\langle 0, 0, 3, 1, 3 \rangle$ . Dans cet exemple, chaque agent choisit son plan initial en supposant la non-existence de ses voisins. Cela donne les plans

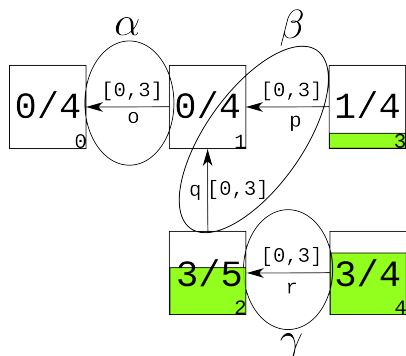


FIGURE 1 – Réseau de l'exemple composé de 5 entités (carrés), 4 actionneurs (arcs) et 3 agents ( $\{\alpha, \beta, \gamma\}$ ) dans son état initial

suivants :

$$\alpha : (o = 0) \quad \beta : (p = 1, q = 0) \quad \gamma : (r = 2)$$

Ces plans initiaux rendent le système dans l'état suivant :

$$(0, 1, 5, 0, 1) \text{ avec comme valeur globale } (C_\omega = 7).$$

Dans cet exemple, les plans des agents donnent les actions qui seront faites peu importe l'état du système. L'adaptation pour le voisinage consistera donc uniquement à envoyer les actions affectant les entités partagées.

L'agent  $\alpha$  enverra son action ( $o = 0$ ) à  $\beta$ . Similairement,  $\gamma$  enverra ( $r = 2$ ) à  $\beta$ . Ce dernier communiquera ( $p = 1, q = 0$ ) à  $\alpha$  et ( $q = 0$ ) à  $\gamma$ . Ainsi, chaque agent pourra modifier sa vision de son environnement en prenant en compte les choix annoncés de leurs voisins. Après un nouveau calcul de plan, les nouveaux plans des agents sont :

$$\alpha : (o = 1) \quad \beta : (p = 1, q = 2) \quad \gamma : (r = 2)$$

Ce qui donne les gains suivants, ici la différence d'évaluation du monde après l'application du plan actuel et du nouveau plan que les agents viennent de calculer :

$$g_\alpha = 1 \quad \text{et} \quad g_\beta = 2 \quad \text{et} \quad g_\gamma = 0$$

Le gain de  $\gamma$  est nul, signifiant qu'il possède déjà un plan localement optimal, aucun changement ne sera donc appliqué à cet agent. L'agent  $\beta$  possède le gain maximum de son voisinage (2) ce qui implique qu'il gardera son nouveau plan. Un agent du voisinage d' $\alpha$  garde déjà son plan ( $\beta$ ),  $\alpha$  ne gardera donc pas le plan qu'il vient de calculer. Le nouveau plan joint résultant de cette itération amènera au système suivant :

$$\alpha : (o = 0) \quad \beta : (p = 1, q = 2) \quad \gamma : (r = 2)$$

$$(0, 3, 3, 0, 1) \text{ avec comme valeur globale } (C_\omega = 5).$$

Les agents ayant trouvé un plan meilleur que leur plan actuel ( $gain > 0$ ) réinitialisent leur compteur ( $d_\alpha = 2$  et  $d_\beta = 2$ ) tandis que les autres agents le décrémentent de

1 ( $d_\gamma = 1$ ). Puis les agents échangent leur compteur avec leur voisinage et ne gardent que le maximum ( $d_\gamma = 2$ ). Comme les compteurs ne sont pas nuls, une nouvelle itération amène aux plans et gains suivants :

$$\alpha : (o = 3) \quad \beta : (p = 1, q = 2) \quad \gamma : (r = 3)$$

$$g_\alpha = 3 \quad \text{et} \quad g_\beta = 0 \quad \text{et} \quad g_\gamma = 1$$

À cette itération, le gain de  $\beta$  est nul, il ne garde donc pas son plan.  $\alpha$  et  $\gamma$  ont les gains maximums de leur voisinage respectif, ils garderont donc leur nouveau plan. Le nouveau plan joint sera donc :

$$\alpha : (o = 3) \quad \beta : (p = 1, q = 2) \quad \gamma : (r = 3)$$

$$(3, 0, 4, 0, 0) \text{ avec comme valeur globale } (C_\omega = 2).$$

Après les réductions, les nouveaux compteurs sont :  $d_\alpha = 2$ ,  $d_\beta = 1$  et  $d_\gamma = 2$ .  $\beta$  a un voisin dont la valeur du compteur est supérieure à la sienne, il la prendra donc ( $d_\beta = 2$ ). Une nouvelle itération sera donc effectuée, menant à :

$$\alpha : (o = 3) \quad \beta : (p = 1, q = 3) \quad \gamma : (r = 3)$$

$$g_\alpha = 0 \quad \text{et} \quad g_\beta = 1 \quad \text{et} \quad g_\gamma = 0$$

Seul  $\beta$  possède un gain non nul, il gardera donc son nouveau plan. Les deux autres agents resteront avec leur plan précédent. Le nouveau plan joint sera donc :

$$\alpha : (o = 3) \quad \beta : (p = 1, q = 3) \quad \gamma : (r = 3)$$

$$(3, 1, 3, 0, 0) \text{ avec comme valeur globale } (C_\omega = 1).$$

Les compteurs après réductions et échanges seront tous égaux à 2, l'algorithme devra donc continuer. À ce point, l'algorithme a atteint un optimum local. Néanmoins, il faudra encore deux itérations pour que la convergence soit détectée par les agents. Durant ces deux itérations, aucun agent ne pourra améliorer son plan, résultant en une réduction progressive des compteurs jusqu'à ce qu'ils atteignent tous la valeur 0 de façon simultanée, ce qui signifiera l'arrêt de la résolution.

## 6 Expérimentation sur des réseaux de voies navigables

Une modélisation stochastique des réseaux de voies navigables en utilisant des MDPs a été initiée dans (Desquesnes *et al.*, 2016) et adaptée à l'OCLP dans (Desquesnes *et al.*, 2017) avec de bons résultats. Un réseau de voies navigables est un système à grande échelle composé de canaux artificiels et des rivières canalisées, divisés par des écluses. Toute partie du réseau entre deux écluses est appelée bief. La préoccupation principale des gestionnaires des voies navigables est de maintenir un certain niveau d'eau sur tous les biefs pour permettre la navigation. La navigation n'est autorisée sur un bief que si et seulement si le niveau d'eau sur tous les biefs est compris entre le niveau maximum de navigation (Highest Navigation Level – HNL) et le niveau

minimum de navigation (Lowest Navigation Level – LNL). Le niveau d'eau optimal est le niveau normal de navigation (Normal Navigation Level – NNL). Le niveau d'eau est perturbé par l'ouverture des écluses imposée par la navigation, des échanges incontrôlés tels que des échanges avec les nappes phréatiques ou encore la météo et, par des déplacements contrôlés résultants de l'ouverture des points de transferts (pompes et portes) entre deux biefs, par les gestionnaires. Avec l'OCLP, les entités du modèle seront les biefs, avec une discrétisation des différents niveaux d'eau comme ensemble d'états de l'entité. Similairement, les transferts d'eau discrétisés sont les actions du réseau. Dans cette application, les agents contrôlent un ensemble de points de transfert et donc observent tous les biefs qu'ils affectent. Le but est de minimiser la distance quadratique du niveau des biefs à leur NNL. Le réseau expérimental, voir figure 2, est composé de 7 biefs, représentés par des carrés, et de 14 points de transfert, représentés par des arêtes, et est planifié sur une durée arbitraire de 8 pas de temps.

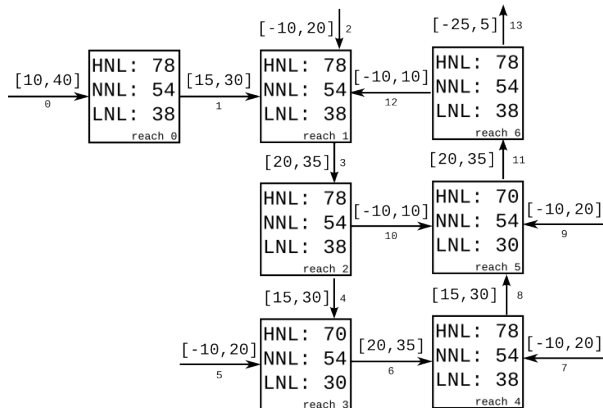


FIGURE 2 – Réseau de voies navigables expérimental composé de 7 biefs (carrés) et de 14 points de transfert (arcs)

Des décompositions, de tailles variables, de ce réseau en agents ont été proposées par un algorithme heuristique spécifique à cette application, visant à réduire la taille des fonctions de transition de tous les agents. Pour chaque décomposition, la somme des tailles des fonctions de transition des agents (en nombre de valeurs), le temps pour converger, le pourcentage du temps qu'un bief soit hors de son rectangle de navigation (*out*) et la distance relative moyenne du bief à son NNL au cours du temps (*avg*) sont observés. Les deux dernières valeurs (*out* et *avg*) correspondent à une moyenne sur 50000 simulations des politiques produites. Ceci est dû au fait que les actions, dans ces expérimentations, correspondent à des intervalles de volumes à transférer. Le simulateur a donc choisi une valeur aléatoirement dans l'intervalle obtenue par la politique. Dans toutes les simulations, les biefs commencent à leur NNL. L'exécution de l'algorithme s'est effectuée sur un cluster, avec un agent par machine, en utilisant le framework JADE (Bellifemine *et al.*, 1999).

	<i>taille</i>	<i>durée</i> (s)	<i>out</i> (%)	<i>avg</i> (%)
6 agents	$5.8 \times 10^6$	591	0.000	17.13
7 agents	$2.9 \times 10^6$	167	0.001	15.38
8 agents	$4.8 \times 10^5$	68	0.000	16.45
9 agents	$4.3 \times 10^5$	35	0.000	17.27
10 agents	$3.5 \times 10^5$	32	0.021	17.16
11 agents	$2.9 \times 10^5$	30	0.023	17.83
12 agents	$2.3 \times 10^5$	31	0.016	17.27
13 agents	$1.7 \times 10^5$	30	0.009	16.64
14 agents	$1.1 \times 10^5$	26	0.007	15.57

TABLE 1 – Scénarios de sept biefs

Les résultats, voir table 1, montrent une réduction évidente de la taille des fonctions de transition et du temps requis pour converger. En effet, la taille du modèle d'un agent croît exponentiellement en fonction du nombre d'entités observées et du nombre de points de transfert contrôlés. L'évaluation produit de bons résultats, puisque le pourcentage du temps hors du rectangle de navigation est infime, tout en étant légèrement influencé par le facteur aléatoire lors de l'application des politiques. Les distances relatives à l'optimal sont aussi correctes au vu de la discrétisation choisie des volumes des biefs et des volumes transférés en intervalles. Pour comparaison, la taille d'un intervalle de volume d'un bief est équivalente à une distance relative de 20% au volume optimal. À cause de la taille du problème et de la limitation des ressources de calcul à notre disposition, trouver une solution optimale de façon centralisée pour cette modélisation n'était pas possible de même pour toute décomposition de taille inférieure à 6.

## 7 Conclusion et travaux futurs

### 7.1 Conclusion

Dans cet article, la coordination hors-ligne de planifications locales (Offline Coordination of Local Planning – OCLP) a été présentée pour des systèmes à base d'entités. L'algorithme proposé permet de résoudre des grands problèmes en distribuant le modèle sur plusieurs agents connectés. Ces agents possèdent des connaissances limitées du système et construisent leur propre politique locale grâce à une coordination inter-agents durant le calcul des politiques. Comme chaque agent est défini avec sa propre politique locale, il n'y a pas besoin de communication durant l'exécution des agents.

L'approche OCLP est prouvée de terminer lorsqu'une solution localement optimale est atteinte et de converger vers cet optimum local. Elle a été utilisée avec succès pour optimiser la gestion de l'eau dans un grand réseau de voies navigables sur plusieurs pas de temps. Ces résultats expérimentaux montrent une diminution significative de la taille du modèle joint et du temps de résolution selon la décomposition du réseau en agents.

Comme cette approche se base sur la notion d'entités parta-

gées, il sera difficile de modéliser des systèmes fortement connectés. Chaque connexion entre deux agents implique des informations redondantes qui nécessiteront plus de modélisation et de communications. Néanmoins, cela pourrait améliorer la qualité des modèles locaux en augmentant leur connaissance du monde extérieur.

## 7.2 Travaux futurs

L'impact de la décomposition de la structure en agents sur la qualité des solutions obtenues pourra être étudié. L'influence du degré d'interconnexion entre les entités et entre les agents sur la capacité de passage à l'échelle et sur les résultats obtenus pourra aussi être examinée. Une fonction de coût multicritères correspondant à une somme de coûts indépendante sur les états et les actions semble posséder les mêmes garanties de convergence que celle utilisée dans cet article et pourrait donc être étudiée. Nous allons pouvoir étudier l'intérêt et la possibilité d'utiliser des agents avec des algorithmes hétérogènes de planification locale en utilisant un langage partagé pour échanger les politiques adaptés. Finalement, cet algorithme pourrait être appliqué à d'autres formalismes de modélisation et de planification avec pas ou peu de modifications et pourraient donc être explorés.

## Références

- BELLIFEMINE F., POGGI A. & RIMASSA G. (1999). Jade—a fipa-compliant agent framework. In *Proceedings of PAAM*, volume 99, p. 33– : London.
- BELLMAN R. (1957). A Markovian Decision Process. *Journal of Mathematics and Mechanics*, **6**(4), 679–684.
- BOUTILIER C., DEARDEN R., GOLDSZMIDT M. & OTHERS (1995). Exploiting structure in policy construction. In *IJCAI*, volume 14, p. 1104–1113.
- CHADES I., SCHERRER B. & CHARPILLET F. (2002). A Heuristic Approach for Solving Decentralized-POMDP : Assessment on the Pursuit Problem. In *SAC '02 : Proceedings of the 2002 ACM symposium on Applied computing*, p. 57–62, Madrid, Spain : ACM.
- DESQUESNES G., LOZENGUEZ G., DONIEC A. & DU-  
VIELLA E. (2016). Dealing with large mdps, case study of waterway networks supervision. In *Advances in Practical Applications of Scalable Multi-agent Systems. The PAAMS Collection*, p. 48–59. Springer.
- DESQUESNES G., LOZENGUEZ G., DONIEC A. & DU-  
VIELLA E. (2017). Distributed MDP for water resources planning and management in inland waterway. *IFAC 2017 World Congress, Toulouse, France, 9-14 July (submitted : notification of acceptance 20 february)*.
- NAIR R., VARAKANTHAM P., TAMBE M. & YOKOO M. (2005). Networked Distributed POMDPs : A Synthesis of Distributed Constraint Optimization and POMDPs. In *National Conference on Artificial Intelligence*, p. 7–.
- PUTERMAN M. L. (1994). *Markov Decision Processes : Discrete Stochastic Dynamic Programming*. John Wiley & Sons, Inc.
- YOKOO M. & HIRAYAMA K. (1996). Distributed breakout algorithm for solving distributed constraint satisfaction problems. In *Proceedings of the Second International Conference on Multi-Agent Systems*, p. 401–408.

# Représentation de termes complexes dans un espace vectoriel relié à une ontologie pour une tâche de catégorisation

A. Ferré<sup>1,2</sup>

<sup>1</sup> MaIAGE, INRA, Université Paris-Saclay, 78350 Jouy-en-Josas, France

<sup>2</sup> LIMSI, CNRS, Université Paris-Saclay, 91405 Orsay, France

INRA de Jouy en Josas, Domaine de Vilvert, 78352 Jouy-en-Josas

arnaud.ferre@universite-paris-saclay.fr

## Résumé

Nous proposons dans cet article une méthode semi-supervisée pour étiqueter des termes de textes par les concepts d'une ontologie. La méthode génère des représentations vectorielles continues des termes complexes dans un espace sémantiquement structuré par l'ontologie du domaine. La méthode proposée s'appuie sur une approche de sémantique distributionnelle, qui génère des vecteurs initiaux pour chacun des termes extraits. Ces vecteurs sont alors plongés dans l'espace vectoriel construit à partir de la structure de l'ontologie. Ce plongement s'effectue par entraînement d'un modèle linéaire. Un calcul de distance est enfin effectué pour déterminer la proximité entre vecteurs de termes et vecteurs de concepts pour déterminer l'étiquetage ontologique des termes. Nous avons évalué la qualité de ces représentations pour une tâche de catégorisation en utilisant les concepts d'une ontologie comme étiquettes sémantiques. La performance de cette méthode atteint l'état de l'art pour cette tâche de catégorisation ouvrant d'encourageantes perspectives.

## Mots Clef

TAL, extraction d'information, étiquetage de texte par une ontologie, espace vectoriel, sémantique distributionnelle, modèle linéaire

## Abstract

We propose in this paper a semi-supervised method for labeling terms of texts by concepts of an ontology. The method generates continuous vector representations of complex terms in a semantically space structured by the ontology of the domain. The proposed method relies on a distributional semantics approach, which generates initial vectors for each of the extracted terms. Then these vectors are embedded in the vector space constructed from the structure of the ontology. This embedding is carried out by training a linear model. Finally, we apply a distance calculation to determine the proximity between vector of terms and vector of concepts and thus to determine the ontological tags of terms. We have evaluated the quality of these representations for a categorization task by using the concepts of an ontology as semantic labels. The performance of this method for the state of the art for this task of standardization opening up encouraging prospects.

## Keywords

NLP, information extraction, ontology-based text tagging, vector space, distributional semantics, multivariate linear regression

## 1 Introduction

Beaucoup des connaissances du domaine biomédical ou biologique sont sous une forme non-structurée, comme celle exprimée dans les articles scientifiques [1]. Pour les experts de ces domaines, l'augmentation conséquente de la littérature spécialisée a créé un besoin important en méthodes automatiques d'extraction d'information [2]. La tâche de catégorisation est une des tâches principales pour répondre à ce besoin.

La catégorisation consiste à annoter des termes (mono- ou multi-mots) des textes avec une ou plusieurs catégories sémantiques (e.g. un terme extrait d'un corpus tel que « *children greater than 9 years of age who had lower respiratory illness* » pourrait être catégorisé par une catégorie sémantique ayant pour label « *pediatric patient* » et/ou « *patient with disease* »). Les concepts d'une ontologie peuvent être utilisés pour représenter ces catégories sémantiques de façon formelle et structurée. La catégorisation rencontre plusieurs difficultés, comme la variabilité importante de la morphologie des termes, qu'ils soient représentés par un mot ou par plusieurs [3]. Les termes multi-mots qui présentent des structures morphosyntaxiques variées et des imbrications complexes, tels que les groupes nominaux complexes (*complex noun phrases*) sont particulièrement difficiles à étiqueter par des catégories. Or, dans les textes de la littérature spécialisée, tels que les articles scientifiques en science du vivant, les groupes nominaux complexes sont abondant [4]. Une approche basée sur la similarité morphologique entre terme et étiquette sémantique apparaît limitée pour effectuer cette tâche [5], parce que la morphologie des labels des concepts n'est pas nécessairement proche de la morphologie des termes à annoter. Une autre difficulté vient du nombre important de catégories sémantiques utilisées, rendant une approche par classification supervisée coûteuse en annotation manuelle (plus de 2000 catégories par exemple dans l'ontologie des habitats bactériens OntoBiotop [6]). Une alternative consiste à calculer la proximité sémantique entre des termes par sémantique distributionnelle. C'est une approche fondée sur la corrélation entre la similarité de

sens et la similarité de distribution des unités sémantiques (mot, combinaison de mots, phrase, documents, ...) [7], [8]. Une unité sémantique peut être représentée par un vecteur construit à partir de la distribution des informations de contexte dans lesquels elle est trouvée. La proximité des vecteurs dans cet espace est alors transposable à une proximité sémantique [9]. Il existe aujourd'hui de nombreuses méthodes de génération de tels espaces vectoriels, tel que Word2Vec [10], mais celles-ci se concentrent habituellement sur les jeux de données massifs [11] dans lesquels l'information est relativement répétée.

La question qui nous intéresse ici est : comment utiliser la sémantique distributionnelle pour catégoriser les termes par une ontologie, autrement dit comment relier l'information distributionnelle aux catégories d'une ontologie. Dans le cadre de la littérature spécialisée qui nous intéresse ici, la question se focalise sur des relativement petits corpus et un grand nombre de catégories sémantiques.

Nous proposons une méthode originale dans laquelle nous représentons des termes complexes basés sur un « plongement de mots » (*word embedding*), en représentant une ontologie sous forme d'espace vectoriel et en entraînant une transformation de vecteurs de termes en vecteurs de concepts. Ensuite, cette transformation est utilisée pour déterminer le concept le plus approprié pour chaque terme extrait.

## 2 Matériel

Les données utilisées sont celles de la tâche de catégorisation Bacteria Biotope (tâche 3) de BioNLP Shared Task en 2016 [12]. Les documents sont des références de la MEDLINE [13], composées de titres et de résumés d'articles scientifiques dans le domaine de la biologie. La tâche consiste, étant donné les entités du corpus dénotant les habitats bactériens, à leur assigner une catégorie de l'ontologie OntoBiotope. Le corpus (noté BB dans la suite) est divisé en trois : le corpus d'entraînement, le corpus de développement et le corpus de test. Dans les corpus d'entraînement et de développement les catégories des termes sont données : elles nous ont servi à entraîner notre méthode. Le corpus de test est celui pour lequel les catégories sont à prédire : il nous sert à évaluer notre méthode pour la tâche de catégorisation. Les entités de chacun de ces corpus ont été annotées manuellement. Voici un résumé de leurs caractéristiques :

	Entraîn.	Dév.	Test	Total
Documents	71	36	54	<b>161</b>
Mots	16 295	8 890	13 797	<b>38 982</b>
Entités	747	454	720	<b>1 921</b>
Entités distinctes	476	267	478	<b>1 125</b>
Cat. sémantiques	825	535	861	<b>2 221</b>
Cat. distinctes	210	122	177	<b>329</b>

TABLE 1 : Statistiques descriptives du corpus BB

En plus de ce corpus, nous avons utilisé un corpus élargi du même domaine pour générer des représentations

vectorelles de chaque mot. Il est composé de 100 000 phrases venant de titres et de résumés d'articles scientifiques dans le domaine de la biologie disponibles sur PubMed. Cela représente un corpus de taille relativement petit, qui contient une majorité de mots non-outils avec une faible fréquence d'apparition (cf. TABLE 2).

Répétés >2	72412	35%
Répétés 2 fois	31569	15%
Non répétés	105364	50%
Total mots non-outils	209345	100%

TABLE 2 : Statistiques descriptives du corpus élargi

## 3 Méthode

### 3.1 Génération de vecteurs de mots

L'espace vectoriel des termes (EVT) est obtenu en générant un vecteur pour chacun des mots du corpus élargi qui comprend également les corpus d'entraînement et de développement, mais pas le corpus test. Pour cela, nous avons utilisé l'outil Word2Vec [10] en prenant pour contexte d'un mot, tous les mots contenus dans la phrase. Pour avoir suffisamment de données d'entraînement pour la génération de vecteurs de mots, et aussi pour éviter de prendre en compte des fautes de frappes ou des erreurs, il est habituellement conseillé d'utiliser Word2Vec sans les mots peu fréquents, n'apparaissant qu'une ou deux fois dans tout le corpus. Notre corpus contenant beaucoup de mots d'intérêt à faible fréquence, nous avons fait le choix de ne pas appliquer de seuil de fréquence. Après quelques tests de performance, la dimension 200 a été choisie pour les vecteurs de sorties (cf. FIGURE 1A), ce qui est du même ordre de grandeur que ce qui est conseillé habituellement [10].

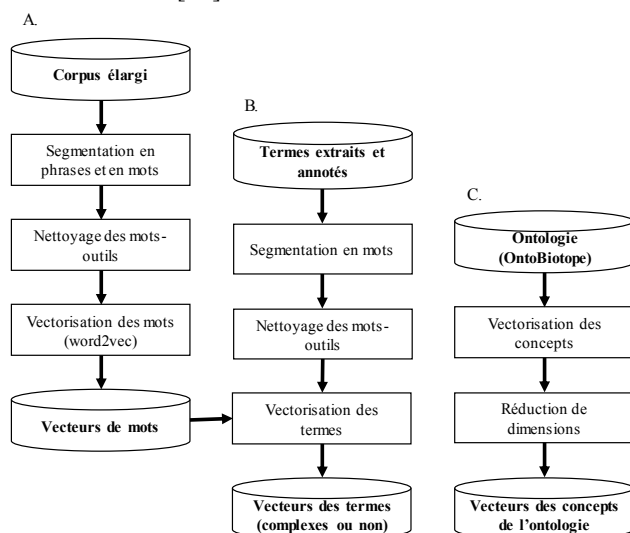


FIGURE 1 : A. Processus de création de vecteurs de mots. B. Processus de création des vecteurs de termes. C. Processus de création de vecteurs de concepts

### 3.2 Génération de vecteurs de termes complexes

Pour calculer les représentations vectorielles des termes composés extraits des corpus (cf. FIGURE 1B), on



commence par les segmenter en mots. Pour chaque mot non-outil, on utilise le vecteur calculé par Word2Vec. Le vecteur du terme composé est obtenu par la moyenne des vecteurs des mots qui le composent :

$$v_{t_k} = \sum_{i=1}^{n_k} v_{m_i^k} / n_k \quad (1)$$

Où  $v_{t_k}$  est le vecteur associé au terme  $t_k$ ,  $n_k$  est le nombre de mots non-outils du terme  $t_k$ ,  $v_{m_i^k}$  est le vecteur du mot  $m_i^k$  issu de Word2Vec, et le terme  $t_k$  est tel que :

$$\forall i \in [1, n_k], m_i^k \in t_k \quad (2)$$

### 3.3 Génération de vecteurs de concepts

Pour construire les vecteurs de concepts et donc un espace vectoriel ontologique (EVO), on initialise des vecteurs nuls possédant autant de dimension que de concepts dans l'ontologie. Chaque valeur du vecteur correspond donc à un des concepts de l'ontologie. La valeur est à 1 si le concept est un ancêtre du concept considéré, à 0 sinon :

$$v_{c_k} = (w_{c_k}^0, \dots, w_{c_k}^i, \dots, w_{c_k}^n) \quad (3)$$

Où  $v_{c_k}$  est le vecteur associé au concept  $c_k$ ,  $c_k$  est relié à la  $k^{\text{ème}}$  dimension des vecteurs (i.e.  $w_{c_k}^k$ ),  $n$  est le nombre de concepts dans l'ontologie et  $w_{c_k}^i$  est la valeur du vecteur  $v_{c_k}$  pour la dimension  $i$ , tel que :

$$w_{c_k}^i = \begin{cases} 1 & \text{si } i = k \\ 1 & \text{si } c_i \text{ parent (direct ou non) de } c_k \\ 0 & \text{sinon} \end{cases} \quad (4)$$

Cette représentation a comme intérêt de conserver les distances (distance cosinus) attendues entre les concepts (cf. FIGURE 2 et TABLE 3) : un concept est plus proche de ses fils, puis de ses parents.

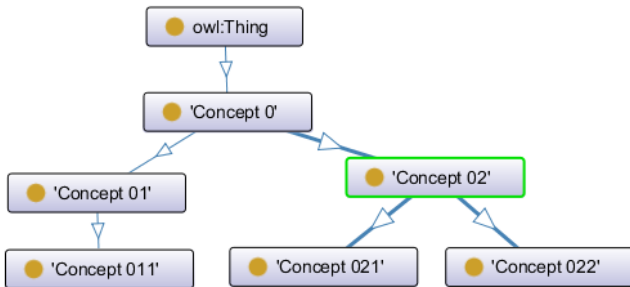


FIGURE 2 : Représentation conceptuelle d'une ontologie

Concept 02	distance
Concept 02	1,0000
Concept 021	0,8165
Concept 022	0,8165
Concept 0	0,7071
Concept 01	0,5000
Concept 011	0,4082

TABLE 3 : Similarité cosinus du concept 02 avec ses concepts voisins

On remarque que la dimension de l'EVO généré a alors pour taille le nombre de concepts de l'ontologie (soit plus de 2000 pour l'ontologie OntoBiotope). Pour décomposer la transformation EVT->EVO recherchée, on peut commencer par réduire le nombre de dimension de l'EVO

pour atteindre le même nombre de dimensions que l'EVT généré précédemment (cf. FIGURE 1C). Une analyse en composantes principales (ACP) et un positionnement multidimensionnel (MDS) ont été testés.

### 3.4 Entraînement (modèle linéaire)

L'objectif de la phase d'entraînement est de déterminer une transformation de l'EVT vers l'EVO qui minimise la distance entre les vecteurs de termes issus de cette transformation et les vecteurs des concepts associés. Nous avons choisi de nous limiter à une transformation linéaire et avons entraîné un algorithme de type modèle linéaire avec les corpus annotés d'entraînement et de développement (cf. FIGURE 3).

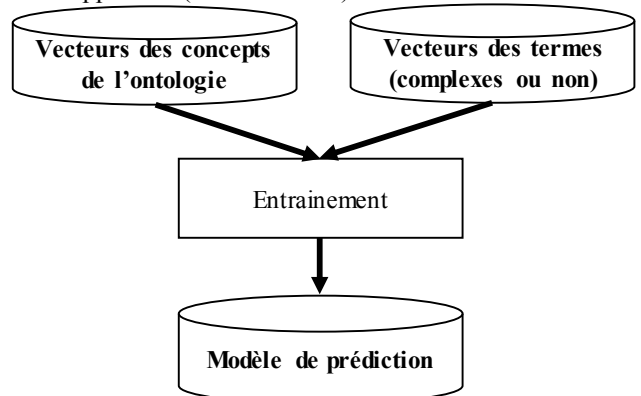


FIGURE 3 : Processus d'entraînement pour déterminer une transformation de l'EVT vers l'EVO

Les matrices de transformations obtenues permettent alors de prédire de nouveaux vecteurs associés aux termes du corpus de test exprimés dans l'EVO. Pour répondre à la tâche d'évaluation, on recherche le vecteur de concept le plus proche en terme de distance cosinus. Le concept ainsi trouvé est celui qui est assigné au terme (cf. FIGURE 4).

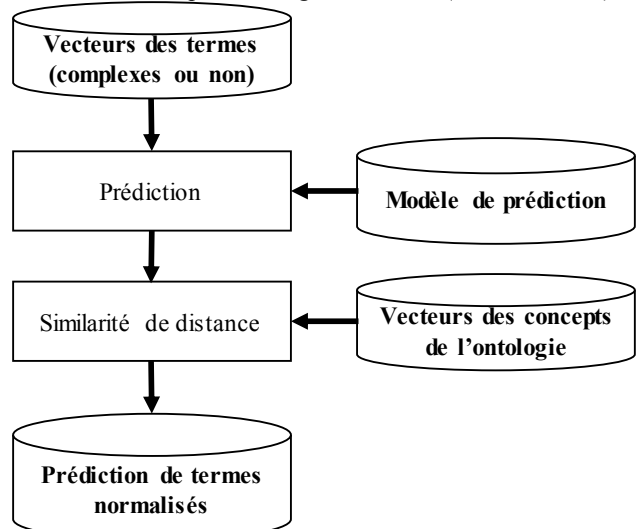


FIGURE 4 : Processus de prédiction des catégories sémantiques associées aux termes extraits

L'objectif de la phase d'entraînement est de déterminer une transformation de l'EVT vers l'EVO qui minimise toutes les distances entre les vecteurs des termes résultants dans

l'EVO et les vecteurs des concepts associés. Dans cet article, une transformation linéaire est étudiée car nous faisons l'hypothèse qu'il y a une certaine similitude de répartition entre les vecteurs de termes dans l'EVT et les vecteurs de concepts associés dans l'EVO. Autrement dit, une transformation non-linéaire pourrait fortement déformer la répartition des vecteurs de termes dans l'EVO pour s'adapter aux données d'entraînement peu nombreuses et ne recouvrant qu'une faible partie des annotations pouvant être détectées dans les textes ciblés. Cet entraînement vise à obtenir les meilleurs paramètres pour approximer l'équation matricielle suivante :

$$Y = X.B + U \quad (5)$$

Où Y est une matrice formée d'une série de vecteurs de concepts, X est une matrice formée d'une série de vecteurs de termes (où la  $i^{\text{ème}}$  ligne de X représente le vecteur d'un terme qui est annoté par un concept qui a pour vecteur la  $i^{\text{ème}}$  ligne de Y), B est la matrice contenant les paramètres qui sont à estimer, et U est une matrice contenant une distribution gaussienne multivariée. Cet entraînement est réalisé sur les corpus d'entraînement et de développement (cf. FIGURE 3).

La matrice obtenue nous permet de concevoir une fonction de transformation linéaire, afin de permettre de prédire de nouveaux vecteurs associés aux termes du corpus d'essai exprimé dans l'EVO :

$$f: \left( \begin{array}{c} \text{EVT} \rightarrow \text{EVO} \\ v_{\text{term}} \rightarrow v'_{\text{term}} = f(v_{\text{term}}) \end{array} \right) \quad (6)$$

Où  $v_{\text{term}}$  est un vecteur de terme dans l'EVT et  $v'_{\text{term}}$  est le vecteur résultant du même terme projeté dans l'EVO. Pour satisfaire aux exigences de la tâche d'évaluation, le vecteur de concept le plus proche (en terme de distance cosinus) de  $v'_{\text{term}}$  est choisi pour le terme annoté (cf. la FIGURE 4).

## 4 Résultats

### 4.1 Catégorisation

Pour évaluer la performance des systèmes participants à la tâche 3 de BB, une mesure de similarité sémantique est implémentée sur le site du challenge BioNLP-ST 2016. La mesure utilisée est celle définie par Wang et al. en 2007 [14], avec le paramètre de poids à 0.65. Avec cette mesure, nous pouvons calculer une *baseline* en attribuant à tous les termes le concept « bacteria habitat », qui est la racine de la hiérarchie de l'ontologie OntoBiotope. La mesure trouvée est alors de 32.17%.

Catégorisation	Score final de similarité
BOUN	0.62
<b>CONTES</b>	<b>0.60</b>
LIMSI	0.44

TABLE 4 : Résultats de la tâche de catégorisation de BioNLP-ST 2016

Deux équipes avaient participé à cette tâche de BioNLP-ST 2016 et avaient obtenu les résultats rapportés dans la TABLE 4. Notre méthode (CONTES - CONcept-TErm System) a obtenu un résultat de 60%, tout à fait comparable à celui de la première équipe et

significativement au-dessus de la méthode du LIMSI qui s'appuyait sur une approche morphologique.

### 4.1 Vecteurs de termes extraits

En dépit de la faible fréquence d'apparition des mots du corpus élargi (cf. TABLE 2), les vecteurs de mots obtenus présentent des proximités relativement satisfaisantes du point de vue de la similarité sémantique des termes associés. De plus, la méthode utilisée pour former des vecteurs pour les termes complexes semblent elle aussi satisfaisante comme le montre l'exemple suivant :

cell	Similarité
<i>HCE cell</i>	0,99
<i>13C-labeled cell</i>	0,99
<i>parietal cell</i>	0,99
<i>Schwann cell</i>	0,99
<i>CD8+ T cell</i>	0,98
<i>PMN cell</i>	0,97
<i>macrophage cell</i>	0,95

TABLE 5 : Termes à proximité du terme "cell"

Il semble également que des différences morphologiques n'empêchent pas l'agglomération de vecteurs de termes proches (cf. TABLE 6 et TABLE 7), ce qui était une des propriétés recherchées.

younger ones	Similarité
<i>children less than five years of age</i>	0,81
<i>children less than 2 years of age</i>	0,81
<i>children less than two years of age</i>	0,80

TABLE 6 : Termes à proximité du terme "younger ones"

seawater	Similarité
<i>sediments</i>	0,77
<i>sediment sample from a disease-free fish farm</i>	0,75
<i>fish farm sediments</i>	0,73
<i>subterranean brine</i>	0,73
<i>lagoon on the outskirts of the city of Cagliari</i>	0,71
<i>petroleum reservoir</i>	0,71
<i>marine environments</i>	0,71
<i>marine bivalves</i>	0,69
<i>sediment samples from diseased farms</i>	0,69
<i>urine sediments</i>	0,68
<i>petroleum</i>	0,66
<i>subterranean environment</i>	0,65
<i>fresh water</i>	0,65
<i>fresh water supply</i>	0,64
<i>Seafood</i>	0,64
<i>marine</i>	0,64

TABLE 7 : Termes à proximité du terme "seawater"

Néanmoins, la cooccurrence de certains mots semble agglomérer certains termes de catégorie différente. 2 mots apparaissant fréquemment dans des contextes communs se

retrouvent alors avec des vecteurs similaires. Cette similarité persiste alors également lors du calcul des vecteurs de termes. C'est par exemple le cas pour les termes relatifs au poisson et ceux relatifs aux fermes d'élevage de poissons (cf. TABLE 8). Ces représentations sont moins satisfaisantes car elles ne permettent pas de différencier les catégories sémantiques sous-jacentes.

fish	Similarité
fish farming	0,98
fish farm	0,92
disease-free fish farm	0,91
fish farm sediments	0,87
healthy fish	0,81

TABLE 8 : Termes les plus proches du terme "fish"

### 4.2 Vecteurs de concepts de l'ontologie

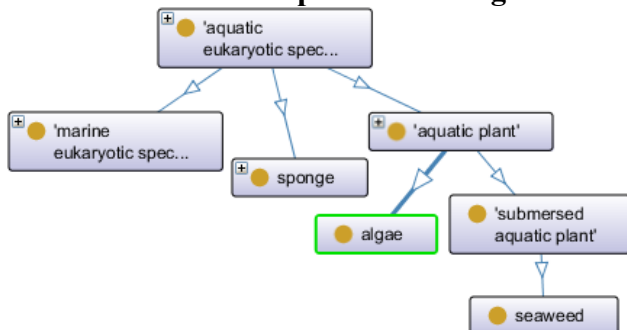


FIGURE 5 : Taxonomie des concepts autour du concept "algae" (visualisé avec le logiciel Protégé)

On peut estimer la qualité des vecteurs de concepts créés en observant la cohérence entre la proximité vecteur/vecteur et leur sens (cf. FIGURE 5 et TABLE 9). Une réduction de dimension semble détériorer progressivement l'EVO et impacte parfois ces distances de manière suffisante pour empêcher la distinction entre concepts parents, concepts fils et concepts frères (cf. FIGURE 5 et TABLE 10). Comme conséquence, à partir de l'EVO réduit, il semble difficile de restituer parfaitement la structure de l'ontologie. Prenons l'exemple du concept 'algae' :

<OBT:001922: algae> sans ACP	distance
<OBT:001777: aquatic plant>	0,9258
<OBT:001895: submersed aquatic plant>	0,8571
<OBT:001967: seaweed>	0,8018

TABLE 9 : Similarité pour le concept algue de l'ontologie (sans ACP)

<OBT:001922: algae> avec ACP	Similarité
<OBT:001777: aquatic plant>	0,99
<OBT:001895: submersed aquatic plant>	0,99
<OBT:001967: seaweed>	0,99
<OBT:000372: sponge>	0,93
<OBT:000269: marine eukaryotic species>	0,93

TABLE 10 : Similarité pour le concept algue de l'ontologie (avec ACP de dimension finale 300)

En comparant plusieurs exemples, il semble que l'ACP ne modifie pas l'ordre de proximité des concepts, mais on peut observer des augmentations de densité de ces vecteurs (cf. comparaison entre le TABLE 9 et TABLE 10). Cela semble cohérent du fait de la diminution de l'espace concerné. En effet, l'ordre des plus proches voisins d'un vecteur de concept ne semble pas modifié.

### 4.3 Influence de la dimension de l'EVT

Word2Vec permet l'utilisation de 2 architectures différentes pour générer des vecteurs associés à des mots d'un corpus : Continuous Bag Of Words (CBOW) et Skip-Gram. L'architecture CBOW peut être défini par l'objectif de prédire un mot en fonction de son contexte, alors que le Skip-Gram est de prédire le contexte en fonction d'un mot en entrée. Quelque soit l'architecture utilisée, Word2Vec permet de générer des vecteurs associés à chaque mot du corpus. Nous avons testé les 2 architectures sur des dimensions de vecteurs de sortie différentes (cf. FIGURE 6). Pour des espaces vectoriels générés avec une dimension entre 100 et 250, les scores finaux semblent relativement stables, particulièrement pour le CBOW. De même, l'écart de score entre les 2 architectures restent en dessous des 3%. Au dessus d'une dimension de 250, on assiste à une diminution du score pour les 2 architectures, avec une pente plus importante avec le CBOW.

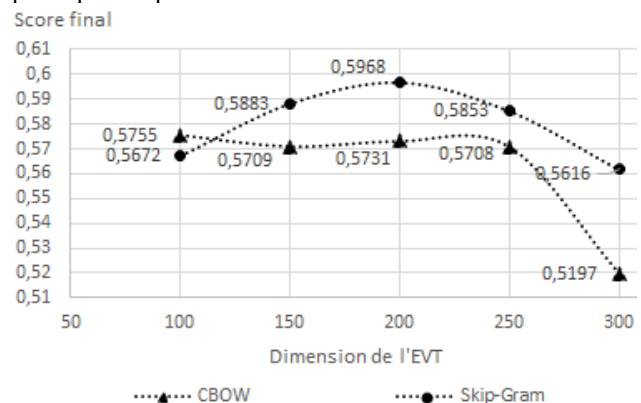


FIGURE 6 : Comparaison des architectures CBOW et Skip-Gram sur différentes dimensions d'EVT

### 4.4 Influence d'une réduction de dimension de l'EVO sur la catégorisation

L'EVO possède une dimension importante par rapport à l'information spécifique qui y est codée (i.e. la structure de l'ontologie). Cela peut poser des difficultés d'ordre combinatoire mais aussi théorique : une projection linéaire de l'EVT sur l'EVO (de dimension plus grande que l'EVT) ne devrait alors se faire que sur un sous-espace de l'EVO, limitant ainsi les résultats. Il était donc intéressant d'étudier l'impact d'une réduction de dimension de l'EVO sur le score final. On peut alors observer qu'une réduction par analyse en composantes principales - ACP (avec des résultats similaires avec un positionnement multidimensionnel - MDS) diminue systématiquement le score obtenu par rapport à l'utilisation de l'EVO non-réduit (cf. FIGURE 7). Néanmoins, on observe également un

palier d'une certaine performance (moins de 3% en dessous du score sans réduction) jusqu'à un certain point où le score diminue drastiquement.

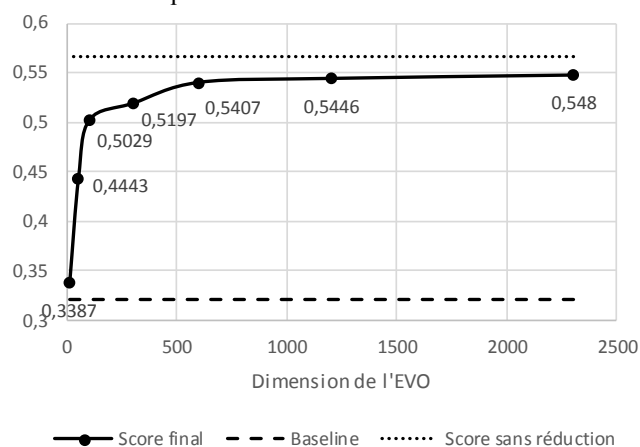


FIGURE 7 : Impact de la réduction de l'EVO (pour un EVT généré de dimension 100 avec Skip-Gram)

## 5 Discussions

Pour étendre les interprétations issues d'exemples, il serait intéressant d'évaluer la qualité globale des espaces vectoriels générés : espace vectoriel des mots, des termes, des concepts ainsi que l'espace finale contenant les transformations des vecteurs des termes.

Un des plafonds de la méthode présentée dans cet article est dû au fait que, pour cette tâche de catégorisation, un terme peut être catégorisé par plusieurs concepts de l'ontologie (ex : le terme « school age children with wheezing illness » devrait être catégorisé par le concept <OBT:002307: pediatric patient> ainsi que le concept <OBT:002187: patient with disease>), mais c'est également le cas des autres systèmes participants.

## 6 Perspectives

Pour de futurs travaux, il serait pertinent d'appliquer des méthodes d'évaluation globale de la qualité des espaces vectoriels générés. En particulier, cela permettrait d'évaluer plus exhaustivement les processus intermédiaires et d'observer avec plus de précision l'impact des modifications sur leurs paramètres internes. De nouvelles méthodes plus élaborées pourraient alors être envisagées pour améliorer les résultats. Par exemple, il serait certainement positif d'utiliser une méthode de représentation vectorielle d'une ontologie qui générerait un espace possédant une dimension plus faible tout en conservant la possibilité de discerner la structure initiale de l'ontologie. De même, la méthode utilisée ici pour générer les vecteurs de l'EVT pourrait être améliorée pour prendre en compte le contexte syntaxique des termes. Cela pourrait résoudre les problèmes de similarité sémantique entre 'fish' et 'fish farm' (cf. TABLE 8).

Il arrive fréquemment que des termes doivent être annotés par plusieurs concepts de l'ontologie cible (par exemple : 'children greater than 9 years of age who had lower respiratory illness' devait ici être annoté par le concept <OBT:002307: pediatric patient> et par le concept

<OBT:002187: patient with disease>). Avoir à disposition une ontologie complètement définie possédant explicitement tous les concepts résultant de l'intersection possible de ses autres concepts (par exemple : un concept 'pediatric patient with disease' qui est un sous-ensemble de <OBT:002307: pediatric patient> et de <OBT:002187: patient with disease>) devrait améliorer les résultats. Si de telles ontologies semblent relativement rares dans le domaine biologique, il pourrait être intéressant de commencer par générer automatiquement tous les concepts équivalents à l'intersection des concepts non-disjoints pour répondre à ce problème. Néanmoins, si les concepts partagent de nombreuses intersections entre eux ou que le caractère disjoint n'a pas été formalisé, la taille de l'ontologie générée risque de poser des difficultés combinatoires.

Il est relativement peu commun d'avoir pour données initiales les termes extraits (ainsi que leur appartenance à des entités nommées d'intérêt). Les méthodes d'extraction terminologique possédant des performances relativement acceptables, il serait intéressant d'en utiliser en amont de la tâche actuelle.

Enfin, malgré la limitation inhérente des méthodes de catégorisation basées sur la morphologie des mots, celles-ci pourraient néanmoins être utilisées pour effectuer une pré-catégorisation du corpus. En conséquence, on pourrait envisager d'utiliser ces annotations pour entraîner la méthode au lieu d'utiliser une annotation manuelle. Ainsi, cela transformerait cette méthode en une méthode non-supervisée.

## 7 Conclusion

L'objectif de cet article était de proposer une approche pour la création de représentations vectorielles pour des termes (complexes ou non) dans un espace sémantique. De plus, il visait à proposer une méthode capable de s'adapter à un corpus spécialisé de petite taille où les termes d'intérêts apparaissent avec une fréquence relativement faible. Les méthodes les plus utilisées actuellement génèrent des espaces vectoriels dont il est difficile d'interpréter le sens autrement qu'en terme de proximité spatiale/similarité sémantique. Notre méthode semble montrer qu'en combinant des approches relativement classiques, il est possible d'utiliser une ontologie pour générer des vecteurs dans un espace vectoriel plus interprétable. Les résultats comparables à ceux de l'état de l'art, semble ouvrir d'encourageantes perspectives. Au-delà de la tâche de catégorisation, de nouvelles méthodes performantes de génération d'espace vectoriel interprétables pourraient également répondre à de nombreuses problématiques.

## Remerciements

This work is supported by the "IDI 2015" project funded by the IDEX Paris-Saclay, ANR-11-IDEX-0003-02.

## Bibliographie

- [1] N. Kang, B. Singh, Z. Afzal, E. M. van Mulligen, and J. A. Kors, "Using rule-based natural language processing to improve disease normalization in biomedical text," *J. Am. Med. Inform. Assoc.*, vol. 20, no. 5, pp. 876–881, Sep. 2013.
- [2] S. Ananiadou and J. McNaught, Eds., *Text mining for biology and biomedicine*. Boston: Artech House, 2006.
- [3] A. Nazarenko, C. Nédellec, E. Alphonse, S. Aubin, T. Hamon, and A.-P. Manine, "Semantic annotation in the alvis project," in *International Workshop on Intelligent Information Access (IIA)*, 2006, p. 5–pages.
- [4] F. Maniez, "Prémodification et coordination : quelques problèmes de traduction des groupes nominaux complexes en anglais médical," *ASp*, no. 51–52, pp. 71–94, Dec. 2007.
- [5] W. Golik, P. Warnier, and C. Nédellec, "Corpus-based extension of termino-ontology by linguistic analysis: a use case in biomedical event extraction," in *WS 2 Workshop Extended Abstracts, 9th International Conference on Terminology and Artificial Intelligence*, 2011, pp. 37–39.
- [6] R. Bossy, W. Golik, Z. Ratkovic, D. Valsamou, P. Bessières, and C. Nédellec, "Overview of the gene regulation network and the bacteria biotope tasks in BioNLP'13 shared task," *BMC Bioinformatics*, vol. 16, no. 10, p. S1, 2015.
- [7] J. R. Firth, "The technique of semantics," *Oxford University Press*, London, 1957.
- [8] Z. S. Harris, "Distributional Structure," *WORD*, vol. 10, no. 2–3, pp. 146–162, Aug. 1954.
- [9] C. Fabre and A. Lenci, "Distributional Semantics Today Introduction to the special issue," *Trait. Autom. Lang.*, vol. 56, no. 2, pp. 7–20, 2015.
- [10] T. Mikolov, K. Chen, G. Corrado, and J. Dean, "Efficient estimation of word representations in vector space," *ArXiv Prepr. ArXiv13013781*, 2013.
- [11] C. Fabre, N. Hathout, F. Sajous, and L. Tanguy, "Ajuster l'analyse distributionnelle à un corpus spécialisé de petite taille," in *21e Conférence sur le Traitement Automatique des Langues Naturelles (TALN 2014)*, 2014, pp. 266–279.
- [12] L. Deléger, E. Chaix, M. Ba, A. Ferré, P. Bessières, and C. Nédellec, "Overview of the Bacteria Biotope Task at BioNLP Shared Task 2016," 2016.
- [13] MEDLINE, "Base de données bibliographique MEDLINE : <https://www.ncbi.nlm.nih.gov/pubmed>." .
- [14] J. Z. Wang, Z. Du, R. Payattakool, P. S. Yu, and C.-F. Chen, "A new method to measure the semantic similarity of GO terms," *Bioinformatics*, vol. 23, no. 10, pp. 1274–1281, May 2007.



## Extraction de connaissances à partir de l'usage des outils professionnels de communication et de collaboration

Antoine Flepp<sup>1,2</sup>

Fabrice Bourge<sup>1</sup>

Julie Dugdale<sup>2,3</sup>

Tiphaine Marie<sup>1</sup>

<sup>1</sup>Orange Labs – France

<sup>2</sup>Univ. Grenoble Alpes, CNRS, Grenoble INP\*, LIG, F-38000 Grenoble France

<sup>3</sup>Université de Adger – Norvège

Orange Labs P&S

42 rue des Coutures – BP 6243

14066 Caen Cedex 4

antoine.flepp@orange.com

### Résumé

*La multiplication des outils de communication et de collaboration (média sociaux, e-mail, conférence à distance, etc.) a entraîné ces dernières années une explosion du volume de données non-structurées. Or, ces données (contenus et usages des outils) recèlent une connaissance explicite et tacite de valeur, souvent inexploitée. Cette connaissance apporte un contexte essentiel pour comprendre les documents de référence existants – plus structurés – qui ont été produits de manière collaborative, permettant à tout employé de donner un sens à leur contenu et favoriser ainsi la collaboration. Après avoir décrit différentes classifications de connaissances, les outils de communication et de collaboration sont catégorisés sur la base de travaux précédents dans le domaine du travail collaboratif, puis au travers de théories sur les médias. Une approche bottom-up est proposée pour extraire la connaissance en utilisant des méthodes quantitatives et des méthodes qualitatives (i.e. des entretiens, questionnaires, pop-ups, fichiers de traces, etc.). Cette connaissance sera utilisée dans le but de développer un modèle qui servira de base pour des collaborations plus efficaces et plus efficaces.*

### Mots Clef

Outils de communication et de collaboration, connaissance explicite, connaissance tacite, connaissance contextuelle, sense-making, CSCW.

### Abstract

*The multiplication of communication and collaboration tools (social media, e-mail, conferencing, etc.) in recent years has led to an explosion of the quantity of unstructured data. Buried within this data – both content and the way we use these tools – lies valuable explicit and tacit knowledge, often untapped. This knowledge provides essential context for understanding the existing, more*

*structured reference documents that were produced in a collaborative manner, allowing people to make sense of their contents and further aiding collaboration. After discussing various classifications of knowledge, communication and collaboration tools are categorised based on previous work in the domain of collaborative work as well as media theories. A bottom-up approach is proposed to extract the knowledge using both quantitative and qualitative methods (i.e. interviews, questionnaires, pop-ups, log files, etc.). This knowledge will be used in order to develop a model that will serve as a basis for more efficient and effective collaborations.*

### Keywords

Communication and collaboration tools, explicit knowledge, tacit knowledge, contextual knowledge, sense-making, CSCW.

### 1 Introduction

Chaque jour la grande majorité des nouvelles données générées et stockées au sein des entreprises sont non-structurées, c'est-à-dire qu'elles ne sont pas organisées selon des modèles prédéfinis. Comme on constate depuis de nombreuses années une croissance de la multiplicité des outils de communication et de collaboration (C&C) – e.g. les outils de messagerie, de conférence à distance, les média sociaux – il est raisonnable de penser qu'une partie significative de ces données non structurées réside dans ces outils : courriels, documents, blogs, contenus audio ou vidéo, etc. Or, ces données constituent potentiellement un gisement prodigieux d'information et de connaissance encore très peu exploité aujourd'hui. Aussi nous avons l'ambition de proposer un cadre de recherche exploratoire visant à extraire, formaliser et modéliser la connaissance non seulement à partir des contenus manipulés ou échangés via ces outils de C&C, mais aussi à partir des usages de ceux-ci. Au-delà de l'intérêt pratique au sein des entreprises, l'apport scientifique est d'extraire et de rendre exploitable des connaissances tacites et latentes.

\* Institute of Engineering Univ. Grenoble Alpes

Nos travaux ayant débuté récemment, cet article se concentre principalement sur un état de l'art. Ainsi, la partie 2 présente une synthèse de plusieurs travaux sur les connaissances tacites et explicites, afin de définir un cadre d'analyse sur le type de connaissance que les données extraites pourront apporter. Ce cadre doit permettre *in fine* de modéliser et de distinguer les connaissances obtenues par les données de contenu, de celles obtenues par les données d'usage. Ensuite nous présentons diverses études sur les outils de C&C, afin d'identifier les données pertinentes qui peuvent être extraites à partir de ces outils. Il y est ainsi décrit le caractère formel ou informel des échanges que ces outils permettent, des caractéristiques importantes de classifications des outils, puis est abordé le potentiel de ces outils à pouvoir compléter ou remplacer les échanges en face-à-face. Dans la partie 3 nous proposons notre hypothèse de travail. Cette dernière stipule qu'à partir du recueil et de l'analyse des données (de contenus et d'usages) issues des outils de C&C – utilisés lors d'un travail collaboratif – nous devrions être en mesure d'en extraire de la connaissance contextuelle (i.e. relative au contexte du travail collaboratif), pour ensuite rendre cette connaissance utile et accessible pour toute personne extérieure à ce contexte. La partie 4 précisera l'approche suivie, ainsi que les éléments de méthodologie associés. Enfin, dans la conclusion, nous évoquerons brièvement le contenu de cet article ainsi que les étapes futures de ces travaux de recherches.

## 2 Contexte

### 2.1 Classifications de la connaissance

De nombreuses études se sont succédées suite aux travaux de Michael Polanyi sur les dimensions tacite et explicite de la connaissance [24]. Deux types d'approches semblent être adoptés : une approche descriptive qui tend à éclaircir les ambiguïtés entre les concepts de connaissance tacite et explicite [15,20], et une approche opérationnelle cherchant plutôt à décrire les processus par lesquels on passe de connaissances tacites à des connaissances explicites et inversement [6,7,13,23]. Ainsi, Polanyi [24] a d'abord décrit les connaissances tacites et les connaissances explicites en deux catégories distinctes. Les connaissances explicites seraient celles codifiées sous une forme structurée (e.g. graphique, tableur, base de données, rapport d'activité). Alors que les connaissances tacites seraient les connaissances non encore codifiées ni structurées que les individus possèdent dans leur cerveau et qui s'acquerraient en majeure partie par l'expérience, l'observation ou encore l'échange d'idées [20]. De fait, la connaissance explicite semble facilement – et à moindre coût – accessible à un grand nombre de personnes, mais elle ne représente probablement qu'une partie de l'ensemble de la connaissance [24]. Ensuite, ce concept de connaissance tacite et de connaissance explicite a été repris – en termes de processus – au travers du modèle SECI de Nonaka et Takeuchi [23], et plus récemment dans le modèle conceptuel de PKM 2.0 de Cheong et

Tsuis [6]. Plus précisément, le modèle SECI – qui tire son nom de : Socialisation, Externalisation, Combinaison, Internalisation – décrit l'interaction entre connaissances tacites et explicites, tant individuelles que collectives, comme un processus en spirale. Le modèle de Cheong, quant à lui, se focalise sur les connaissances individuelles uniquement, d'où le nom de PKM pour *Personal Knowledge Management* (gestion des connaissances personnelles). Globalement, l'auteur propose dans son modèle quatre composants principaux : *Personal Information Management* (gestion des informations personnelles), *Personal Knowledge Internalisation* (internalisation des connaissances personnelles), *Personal Wisdom Creation* (création d'une sagesse personnelle) et *Inter-Personal Knowledge Transferring* (transfert de connaissances interpersonnelles). De fait, ces modèles décrivent les connaissances tacites et les connaissances explicites en termes de processus individuel ou collectif, plutôt qu'en termes de catégories distinctes. C'est pourquoi des auteurs comme Jasimuddin [15] suggèrent finalement qu'il faut concevoir ce concept de connaissance comme un continuum dont la connaissance tacite et la connaissance explicite en seraient les deux extrêmes. Ceci se confirme et s'étaye par des travaux [7,13] qui se poursuivent sur l'articulation des connaissances individuelles et collectives avec les connaissances tacites et explicites. En outre, Grundstein [13] propose que les connaissances individuelles tacites (e.g. talents, habiletés, expérience) amèneraient à des connaissances collectives tacites ; comme, par exemple, des routines défensives (qui feraient obstacles au changement) et des routines offensives (qui favoriseraient l'innovation et le changement). Plus récemment, Collins [7] distingue la connaissance tacite relationnelle (les choses que l'on peut décrire en principe si on s'en donne la peine), la connaissance tacite somatique (les choses que notre corps est capable de faire mais que nous ne pouvons pas décrire) et la connaissance tacite collective (la connaissance que nous tirons du fait de faire partie de la société, comme les règles de langage). De fait, aucun consensus n'apparaît quant à ces différents types de connaissances (tacite individuelle / tacite collective / explicite individuelle / explicite collective). Cependant, il est important de soulever un point que ces deux derniers auteurs considèrent : la connaissance tacite peut être explicitable ou non (*explicable knowledge*). Et c'est en particulier Collins [7 : p.81] qui traite de ce propos, en distinguant différents processus d'explicitation de la connaissance : par élaboration, par transformation, par mécanisation ou par explication (voire par *induction*). Ainsi, même si comme Lejeune [20] l'affirme : « *Dès qu'il est formalisé, le savoir n'est plus tacite, il s'est transformé* » ; il s'avère – au vu des recherches présentées – qu'il y a des connaissances tacites plus difficiles à formaliser que d'autres (*explicable knowledge*), et qu'à en croire Collins, il existe aussi différentes manières de les formuler.



## 2.2 Les outils de communication et de collaboration

Au fil du temps, l'entreprise est passée d'une organisation hiérarchique à une organisation fonctionnelle caractérisée par une formalisation de ses règles de fonctionnement. Ce modèle vise une interdépendance la plus faible possible entre les entités spécialisées qui la composent afin, notamment, de limiter les flux considérés comme un mal nécessaire ; chaque entité s'inscrit dans la réalisation de processus définis [3]. Néanmoins, du fait de la complexification de l'environnement de travail et de la multiplication des rôles assumés par les individus, Bertin et Tran suggèrent que l'on vit actuellement une phase de transition vers un modèle multipolaire dans lequel l'entreprise est composée d'entités de plus en plus multifonctionnelles. Celles-ci tendent à s'auto-organiser pour réaliser des activités pour des groupes – plus ou moins temporaires – qui ne dépendent pas nécessairement des mêmes entités hiérarchiques. Une des conséquences est l'augmentation des flux transverses d'information et de connaissance, ainsi que la nécessité de collaborer davantage en dehors des processus définis, ce qui induit moins de formalisme. Le foisonnement des flux de communication et de coordination est renforcé par la multiplication des outils de C&C qui semble suivre un « effet millefeuille » [17], à savoir un empilement successif des outils les uns sur les autres, sans réorganisation des processus de gestion associés.

De nombreuses classifications des outils de C&C existent de par le monde. En 1988, dans le cadre du CSCW (Computer-Supported Cooperative Work) Johansen [16] propose de catégoriser les outils selon leur contexte d'utilisation via une matrice spatio-temporelle permettant de distinguer si un travail collectif est réalisé en un seul ou plusieurs lieux géographiques, et en même temps (synchrone) ou à des moments différents (asynchrone). Diverses extensions ont été proposées depuis, par exemple, pour prendre en compte la notion de processus : Andriessen [2] retient les processus liés aux échanges entre personnes, aux tâches et aux activités de groupe ; alors que Grudin et Poltrock [12] préfèrent distinguer les processus de communication, de partage d'information et de coordination. Plus récemment, Alimam et al. [1] proposent une typologie de divers outils de collaboration sociale utilisés en entreprise tels que les principaux outils de télécommunication, les outils de collaboration (CSCW, groupware...) et les outils sociaux (Entreprise 2.0). Après l'évaluation de ces outils au regard de treize critères, une typologie est produite en retenant trois des critères indépendants les plus significatifs : diffusion de l'information (modèle expéditeur-destinataire(s) ou modèle *publish-subscribe* – littéralement, publier-s'abonner) ; conception de l'interface utilisateur (classique – i.e. conforme à un modèle communément utilisé ce qui en facilite l'usage et l'acceptabilité – ou spécifique – dont l'usage et l'acceptabilité sont difficiles à

évaluer) ; fonction principale de l'outil (i.e. communication, partage ou recherche d'information, édition ou organisation). La première catégorie de cette typologie regroupe les outils qui permettent une interaction selon le modèle expéditeur-destinataire(s), dont la conception est classique, et qui partagent la même fonction principale de communication et de partage d'information. Ces outils, que les auteurs qualifient de canaux classiques de communication privée, sont : le téléphone, la messagerie vocale, le *chat*, le SMS et l'e-mail. La seconde catégorie de cette typologie regroupe les outils qui suivent le modèle *publish-subscribe*. On distingue alors deux sous-catégories. La première rassemble les outils de conception classique ; leur fonction principale concerne l'organisation (marque-page social) et la gestion de l'information (moteur de recherche, flux d'activité et RSS, streaming / wiki / blog / micro-blog, tableau blanc, forum / communauté / groupe de discussion). La seconde sous-catégorie regroupe les outils de conception spécifique ; leurs fonctions principales sont celles de communication et de partage d'information (conférence à distance), de gestion d'information (partage de fichier, recherche sociale, planification collaborative, banque d'idées, profil social) ou d'organisation (agenda et gestion de calendrier, suivi et localisation, gestion de contenu, application composite ( *mash-up* en anglais), application métier).

Outre ces typologies, il existe aussi la théorie de la richesse des médias – proposée par Daft et Lengel [9]. Cette dernière décrit les outils de C&C selon leur potentiel à transmettre une information riche entre des interlocuteurs (e.g. le téléphone ne transmet pas les expressions du visage, au contraire de la visioconférence). Cette théorie a ensuite été approfondie à plusieurs reprises. On trouve par exemple la théorie de la naturalité des médias introduite par Kock [19], qui stipule que le potentiel d'un média à reproduire l'information transmise dépendrait de son degré de similarité avec une communication face-à-face. L'auteur en conclut qu'il y aurait cinq éléments qui décrivent la naturalité d'un média : la co-implantation (i.e. être au même endroit), la synchronicité (i.e. au même moment) et la capacité à transmettre des expressions langagières, faciales et corporelles. Ainsi, le manque de naturalité d'un média entraînerait chez les interlocuteurs plus d'ambiguïté dans leur communication, un effort cognitif supplémentaire, ainsi qu'une survenue plus rapide de l'ennui dans l'interaction. L'auteur a pu confirmer – avec des études empiriques [18] – ces effets négatifs d'un manque de naturalité d'un média. Néanmoins, cette même étude a aussi montré que la qualité de la tâche demandée n'avait pas été affectée, laissant suggérer ce que l'auteur appelle une adaptation compensatoire (*compensatory adaptation*). Cependant, on trouve aussi d'autres alternatives à la théorie de la richesse des médias, telles que la théorie de la synchronicité des médias de Dennis [10]. Cette théorie – contrairement à la *naturalité des médias* – suggère que

la communication face-à-face n'est pas toujours le médium le plus « riche », mais que le meilleur média ou ensemble de médias dépend de cinq dimensions selon une situation donnée, qui sont : l'immédiateté de la rétroaction, le parallélisme, la variété des symboles, la capacité à peaufiner un message avant de l'envoyer (*rehearsability*) et à rejouer plusieurs fois un message après réception (*reprocessability*). Concrètement, par exemple, le téléphone a un haut degré d'immédiateté de la rétroaction mais un faible degré de *reprocessability* contrairement au courriel. Outre ces théories sur différentes typologies et potentiels des outils de C&C, on trouve aussi des théories centrées *utilisateur*. Par exemple, selon Haythornthwaite [14], plus des pairs auront des liens forts (e.g. liens sociaux ou professionnels) et plus ils utiliseront de médias de communication. Au contraire, des pairs qui auraient des liens faibles, utiliseraient alors peu de médias pour communiquer entre eux. Néanmoins, cette théorie peut être nuancée selon les propos de Conein [8] qui suggèrent qu'« à partir d'un certain seuil un groupement n'est plus viable et se fragmente. Les contraintes cognitives concernent nos capacités limitées de traitement des informations qui portent sur des relations sociales ». Enfin, on peut aussi citer la théorie de l'expansion des canaux de Carlson et Zmud [5], qui aborde notre perception des canaux de communications selon l'expérience utilisateur. Les auteurs proposent quatre types d'expériences utilisateurs pertinentes : expérience du canal de transmission, expérience du sujet du message, expérience du contexte organisationnel et expérience avec les coparticipants. Ainsi, comme l'ont en partie suggéré Kock et les auteurs de cette dernière étude, les utilisateurs s'adaptent aux médias de manière compensatoire. En d'autres termes, avec une plus grande expérience dans l'un des quatre domaines décrits par Carlson et Zmud (canal, sujet, contexte, partenaires), les utilisateurs de médias auraient une meilleure capacité de communiquer dans divers contextes situationnels et percevront ainsi le canal utilisé comme étant plus riche.

### 3 Propositions

Nous tenons à préciser que nous serons amenés dans le futur à distinguer plusieurs types de connaissances tacites et explicites. Pour autant, dans cette première phase de nos recherches, nous considérons uniquement l'hypothèse selon laquelle les contenus échangés avec les outils de C&C sont plutôt sources de connaissance explicite (donc plus facilement extractible) et les usages, plutôt sources de connaissance tacite (plus difficilement extractible). Les outils de C&C semblent incontournables pour échanger et collaborer au sein d'une entreprise. Ainsi, le travail individuel pour la production de données structurées (e.g. un *document de référence* : communiqué de presse, livrable) tend à faire place au travail collaboratif. En outre, comme l'illustre la figure 1 avec la zone de *Production* (en pointillé), le *travail collaboratif* qui se faisait principalement en face-à-face entre des

*contributeurs* peut se faire – à présent – de plus en plus au travers des outils de C&C (bulle orange). Par ailleurs, des études comme celles de Yuan [25] ont montré que souvent les *documents de référence* (correspondant à de la connaissance explicite), ne se suffisent pas à eux-mêmes pour qu'ils soient compréhensibles par le plus grand nombre. Concrètement, lors de l'exploitation de *documents de référence* par des *non-contributeurs* (N-C), ces derniers se retrouvent confrontés à un manque d'informations pour appréhender ces *documents de référence* et doivent contacter directement un *contributeur* du document, voire en discuter avec d'autres *non-contributeurs*, pour pouvoir achever leur processus de compréhension. Au sein d'Orange, par exemple, cette situation se présente à chaque fois qu'une personne intègre un projet déjà en cours, et qu'elle doit s'approprier le contexte au travers de la lecture des documents déjà produits par certains groupes de travail. De même, tout nouvel arrivant dans l'entreprise (apprenti, stagiaire...) est confronté à cette situation puisqu'il doit monter en compétence le plus rapidement possible via la lecture de multiples documents. On peut noter que ce processus de compréhension – qui sera à approfondir dans les recherches futures – est appelé *sense-making* pour des auteurs comme Maurel [21] ou *brainfiltering* par d'autres [4]. Notre hypothèse de travail est que les informations transmises via les outils de C&C – que ce soit au niveau des *contributeurs* entre eux, entre *contributeurs* et *non-contributeurs* ou *a posteriori* au niveau des *non-contributeurs* entre eux – peuvent être automatiquement détectables et exploitables pour en extraire de la connaissance qui sera utile au sein de toute l'entreprise. De plus, différentes études [11,19,22,25] démontrent l'importance du contexte dans la recherche d'informations et de connaissances. Et bien que des auteurs aient mis l'accent sur certaines propriétés contextuelles de la connaissance comme les différences culturelles [25], la complémentarité des outils (e.g. usage de la messagerie

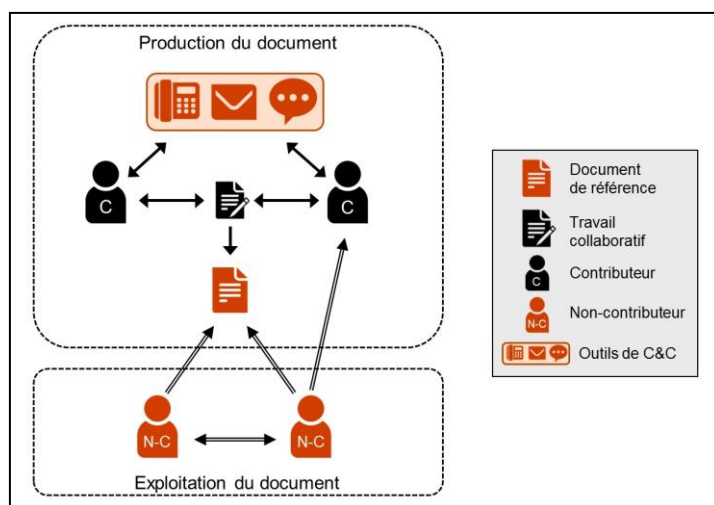


Figure 1: Modélisation d'extraction de connaissances à partir des outils de C&C

instantanée si la personne à contacter est en ligne / usage du téléphone si c'est urgent / usage du courriel si ce n'est pas urgent [22]), le contexte social [19], etc. ; il sera sans doute possible de considérer, dans nos futurs travaux, la *connaissance contextuelle* comme facteur de l'*awareness of expertise distribution* de Yuan [25], voire plus généralement de la *situation awareness* [11].

## 4 Méthodologie

Une approche *top-down* consisterait d'abord à modéliser toute la connaissance d'une entreprise puis à essayer de valider ce modèle sur le terrain. Cependant, il peut y avoir incompatibilité entre le modèle et les données du terrain, d'où une validation difficile. C'est pourquoi nous préférons suivre une approche *bottom-up*, plus pragmatique se décomposant en trois phases : la collecte des données, l'analyse de ces données et la transformation des résultats en connaissance (production d'un modèle de représentation de connaissance). Dans cet article, on se focalise sur la première phase.

### 4.1 Outils ciblés

Il est important d'étudier l'usage de plusieurs outils en parallèle car, d'une part, ceux-ci présentent différentes formes d'interaction (plus ou moins intrusive, formelle/informelle, synchrone/asynchrone, etc.). D'autre part, les personnes utilisent ces outils souvent de manière combinée pour répondre à leur besoin de partage de la connaissance. Ainsi, les outils que nous ciblons a priori pour la collecte de données sont ceux qui sont les plus utilisés au sein d'Orange. Trois de ces outils correspondent au modèle expéditeur-destinataire(s) : l'e-mail, la messagerie instantanée, la conférence à distance. Et un quatrième outil, mis en œuvre plus récemment, correspond au modèle *publish-subscribe* : le réseau social d'entreprise.

### 4.2 Données ciblées

Comme indiqué dans la partie 3, nous souhaitons analyser les contenus échangés et partagés lors d'un travail collaboratif, ainsi que les données d'usages de ces outils de C&C utilisés pour ce travail. Par contenus, nous entendons les contenus proprement dit (i.e. ce qui est manipulé par les utilisateurs) et les données qui les accompagnent (e.g. nom de l'émetteur, date d'envoi). Les données d'usage, quant à elles, correspondent aux informations relatives aux actions des utilisateurs sur les outils (e.g. envoyer un message, ouvrir un outil, chercher un contact). C'est la complémentarité de ces deux types de données qui devrait nous permettre d'extraire la connaissance contextuelle.

### 4.3 Mise en œuvre

Pour collecter des données d'usage, nous mènerons dès que possible des entretiens et soumettrons des questionnaires auprès de quelques personnes qui collaborent fréquemment (soit parce qu'elles font partie de la même équipe, soit parce qu'elles contribuent au même projet) de manière à favoriser la dimension de

partage de connaissances. Lors de cette étape, nous collecterons également auprès de ces mêmes personnes des contenus qu'ils auront générés ou manipulés. Parallèlement, nous récupérerons des fichiers de traces au niveau des serveurs internes de l'entreprise quand ce sera possible. Aussi nous devrions disposer rapidement d'un ensemble de données porteuses de valeur de par leur diversité et complémentarité. Le croisement de ces contenus avec les données d'usages devrait nous permettre d'en déduire des informations quant à l'activité des utilisateurs en fonction des contenus échangés, des personnes avec qui ils interagissent et du moment de la journée. L'objectif est de comprendre le pourquoi et le comment des actions de collaboration. En outre, ceci devrait nous permettre d'affiner la liste des outils qui avaient été ciblés a priori.

Dans un deuxième temps, il nous semble important d'automatiser la collecte des données afin de disposer de volumes significatifs. Ainsi, nous développerons un logiciel – installé sur le poste de travail – pour récupérer certains contenus échangés et leurs données associées, et pour générer des fichiers de traces en local. L'intérêt de ces derniers est de pouvoir tracer finement la cohérence des différentes actions des utilisateurs sur l'ensemble de leurs outils, ce qui est difficile à obtenir à partir des seuls fichiers de traces au niveau des serveurs. Par la suite, nous ferons évoluer ce logiciel afin de recueillir auprès des utilisateurs des données qualitatives sur leurs activités collaboratives. Le recueil de ces avis doit être fait au fil de l'eau, afin que l'utilisateur soit le plus spontané et objectif possible. Ceci de manière à éviter ainsi d'éventuels biais dus au délai et au changement d'état d'esprit entre l'action et l'avis que l'on y porte. Une étape de validation de cet outil sera mise en œuvre afin, notamment, de s'assurer qu'il ne soit pas trop intrusif.

### 4.4. Évaluation des travaux

Il nous semble difficile de pouvoir évaluer nos travaux à l'issue de cette première phase de collecte des données dans la mesure où celles-ci seront porteuses de valeur latente uniquement. Néanmoins, les différentes techniques mises en œuvre ont pour but de maximiser les chances de collecter des données porteuses de valeur pour les phases suivantes.

C'est une fois qu'on aura transformé en connaissances les résultats de l'analyse des données collectées qu'il nous sera possible d'évaluer nos travaux. Moyennant d'identifier les méthodes théoriques ou expérimentales adaptées, deux pistes s'offrent à nous : d'une part évaluer la pertinence et la complémentarité des informations venant « enrichir » un document de référence et, d'autre part, mesurer quantitativement et qualitativement le gain d'efficacité et d'efficience d'un utilisateur lorsqu'il bénéficie d'informations complémentaires pour l'aider à comprendre des documents de référence.

## 5 Discussions

Dans la suite de nos travaux, il sera nécessaire de comparer et d'appliquer différentes méthodes d'extraction de connaissances. Ces dernières devront répondre à différents points importants d'évaluation, tels que la valeur de la connaissance extraite, non seulement du point de vue de leur pertinence, mais aussi en termes de fidélité par rapport aux échanges lors de la collaboration via les outils de C&C. Compte tenu de l'approche bottom-up proposée, les techniques sémantiques de traitement automatique de la langue naturelle – et plus précisément terminologiques – nous semblent a priori plus adaptées pour répondre à notre problématique, bien qu'elles seront sans doute à combiner avec des méthodes statistiques.

## 6 Conclusion

Les outils de C&C font de plus en plus partie intégrante de notre activité professionnelle quotidienne. C'est pourquoi leurs contenus et leurs usages constituent – pour l'entreprise – une source d'information de grande valeur, encore peu exploitée jusqu'à présent. De fait, nous proposons d'extraire de la connaissance à partir du contenu des outils, ainsi qu'à partir des données de leurs usages. Plus précisément, notre objectif est de générer de la connaissance relative au contexte d'un travail collaboratif (*connaissance contextuelle*) – réalisé via des outils de C&C – afin d'en faire bénéficier tout collaborateur extérieur. L'état de l'art présenté permet d'obtenir un premier cadre exploratoire de recherche, que ce soit sur les connaissances, les outils de C&C ou encore la relation qu'il peut y avoir entre la connaissance et ces outils, à savoir que : les données de contenus des outils de C&C seront sources de connaissances explicites et les données d'usages de ces outils, sources de connaissances tacites. Ainsi, nos travaux se porteront dans un premier temps sur la collecte des données qui ont de la valeur a priori à travers les contenus et les usages des outils de C&C. Puis, les étapes suivantes – que nous présenterons à l'occasion d'articles futurs – porteront à la fois sur la mise en œuvre d'outils pour analyser les données collectées, ainsi que sur la formalisation et la modélisation de la connaissance ainsi générée.

## Bibliographie

- [1] M. Alimam, E. Bertin et N. Crespi, "Enterprise Social Systems: The what, the why, and the how," in *17th IEEE Conf. Bus. Informatics (CBI)*, 2015, pp. 9-17.
- [2] J.H.E. Andriessen, *Working with groupware: understanding and evaluating collaboration technology*, Springer Science & Business Media, 2003.
- [3] E. Bertin et S. Tran, "La ville, le SI l'entreprise: du fonctionnel au multipolaire," *Management & Avenir*, 2014, n° 2, pp. 54-72.
- [4] E. Benitez et D. Pauleen, "Brainfiltering: The Missing Link Between PKM and PIM ?," dans *Amer. Conf. on Inform. Syst. (AMCIS) 2009 Proc.*, 13.
- [5] J. R. Carlson et R. W. Zmud, "Channel expansion theory and the experiential nature of media richness perceptions," in *Academy of management journal*, 1999, vol. 42, n° 2, pp. 153-170.
- [6] R. Cheong et E. Tsui, "From Skills and Competencies to Outcome-based Collaborative Work : Tracking a Decade's Development of Personal Knowledge Management (PKM) Models," *Knowledge and Process Management*, 2011, vol. 18, n° 3, pp. 175-193.
- [7] H. Collins, *Tacit and explicit knowledge*, University of Chicago Press, 2010.
- [8] B. Conein, "Cognition distribuée, groupe social et technologie cognitive," *Réseaux*, 2004, n° 2, pp. 53-79.
- [9] R.L. Daft et R.H. Lengel, "Organizational information requirements, media richness and structural design," *Management science*, 1986, vol. 32, n° 5, pp. 554-571.
- [10] A. R. Dennis et J. S. Valacich, "Rethinking media richness: Towards a theory of media synchronicity," in *Proc. 32nd Annu. Hawaii Int. Conf. on Syst. Sci.*, IEEE, 1999.
- [11] M. R. Endsley, "Toward a theory of situation awareness in dynamic systems," *Human Factors: The Journal of the Human Factors and Ergonomics Society*, 1995, vol. 37, n° 1, pp. 32-64.
- [12] J. Grudin et S. Poltrock, "Computer Supported Cooperative Work," in *The Encyclopedia of Human-Computer Interaction*, 2nd Ed., Danemark, 2013.
- [13] M. Grundstein. De la capitalisation des connaissances au management des connaissances dans l'entreprise, les fondamentaux du knowledge management. *INT-Entreprises*, 2013, vol. 3, p. 256-272.
- [14] C. Haythornthwaite, "Social networks and Internet connectivity effects," *Information, Community & Society*, 2005, vol. 8, n° 2, pp. 125-147.
- [15] S. M. Jasimuddin, J. H. Klein, et C. Connell, "The paradox of using tacit and explicit knowledge: strategies to face dilemmas," *Management decision*, 2005, vol. 43, n° 1, pp. 102-112.
- [16] R. Johansen, *Groupware: Computer support for business teams*, The Free Press, 1988.
- [17] M. Kalika, N. B. Charki et H. Isaac, "La théorie du millefeuille et l'usage des TIC dans l'entreprise," *Revue française de gestion*, 2007, n° 3, pp. 117-129.
- [18] N. Kock, "Compensatory adaptation to media obstacles: An experimental study of process redesign dyads," *Information resources management journal*, 2005, vol. 18, n° 2, p. 41.
- [19] N. Kock, "Media richness or media naturalness? The evolution of our biological communication apparatus and its influence on our behavior toward e-communication tools," *IEEE Trans. on Prof. Commun.*, 2005, vol. 48, no 2, p. 117-130.
- [20] M. Lejeune, "Transmettre l'informel en milieu de travail : proximité, posture et modalités organisationnelles," dans *Formations et professionnalisations : à l'épreuve de la complexité*, Paris, L'Harmattan, 2012, chapitre V, pp. 107-115.
- [21] D. Maurel, "Sense-making : un modèle de construction de la réalité et d'appréhension de l'information par les individus et les groupes," *Études de communication*, 2010, n° 2, pp. 31-46.
- [22] T. Niinimäki, A. Piri, C. Lassenius et al, "Reflecting the choice and usage of communication tools in GSD projects with media synchronicity theory," in : *5th IEEE Int. Conf. Global Software Eng. (ICGSE)*, 2010, pp. 3-12.
- [23] I. Nonaka et H. Takeuchi, *The knowledge creating company: how Japanese companies create the dynamics of innovation*, New York : Oxford University Press, 1995.
- [24] M. Polanyi. *The tacit dimension*. First published Doubleday & Co, 1966.
- [25] C. Yuan, X. Zhao, Q. Liao, et al, "The use of different information and communication technologies to support knowledge sharing in organizations: From e-mail to micro-blogging," *Journal of the American Society for Information Science and Technology*, 2013, vol. 64, n° 8, p. 1659-1670.

# Détection de la fumée et du feu par réseau de neurones convolutifs

Sébastien Frizzi<sup>1</sup> Rabeb Kaabi<sup>2,3,4</sup> Moez Bouchouicha<sup>2,3</sup> Jean-Marc Ginoux<sup>2,3</sup> Farhat Fnaiech<sup>4</sup> Eric Moreau<sup>2,3</sup>

<sup>1</sup>Département Génie Biologie- IUT de Toulon, 83957 La Garde, France

<sup>2</sup>Aix Marseille Université, CNRS, ENSAM, LSIS, UMR 7296, 13397 Marseille, France

<sup>3</sup>Université de Toulon, CNRS, LSIS, UMR 7296, 83957 La Garde, France

<sup>4</sup>Université de Tunis, ENSIT, LR13ES03, SIME, 1008, Montfleury, Tunisie

## Résumé

La recherche dans le domaine de la détection du feu et de la fumée devient un sujet de plus en plus traité. Les algorithmes conventionnels utilisent exclusivement des modèles basés sur des vecteurs de caractéristiques. Ces vecteurs restent difficiles à définir et dépendent largement du type de feu observé. Ces méthodes donnent des résultats avec un faible taux de détection et un haut taux de fausse classification. Une approche innovante pour résoudre ce problème est d'utiliser un algorithme permettant de déterminer automatiquement les caractéristiques utiles pour classifier le feu et la fumée. Dans cet article, nous proposons un réseau de neurones convolutif pour identifier le feu et la fumée d'une vidéo en temps réel. Les réseaux de neurones convolutifs ont montré leur grande performance dans le domaine de la classification des objets. Testée sur des séquences vidéos réelles, l'approche proposée atteint une meilleure performance de classification que les méthodes conventionnelles. Ces résultats indiquent clairement que l'utilisation des réseaux de neurones convolutifs pour la détection du feu et de la fumée est très prometteuse.

## Mots Clef

Feu, fumé, détection, réseau de neurones convolutif, apprentissage profond, cartes de caractéristiques, dropout, maxpooling.

## Abstract

Research on video analysis for fire detection has become a hot topic in computer vision. However, the conventional algorithms use exclusively rule-based models and features vector to classify whether a frame is fire or not. These features are difficult to define and depend largely on the kind of fire observed. The outcome leads to low detection rate and high false-alarm rate. A different approach for this problem is to use a learning algorithm to extract the useful features instead of using an expert to build them. In this paper, we propose a convolutional neural network (CNN) for identifying fire in videos. Convolutional neural network are shown to

*perform very well in the area of object classification. This network has the ability to perform feature extraction and classification within the same architecture. Tested on real video sequences, the proposed approach achieves better classification performance as some of relevant conventional video fire detection methods and indicates that using CNN to detect fire in videos is very promising. It's the English version of the abstract.*

## Keywords

Fire and smoke detection, deep learning, convolutional neural network, feature maps, max pooling, dropout

## 1 Introduction

La détection du feu et de la fumée est cruciale pour la protection des personnes. Des dispositifs ont été développés pour prévenir les dommages provoqués par un feu. La majorité de ces dispositifs sont basés sur des capteurs non optiques et se limitent à la détection dans les locaux d'habitations ou commerciaux. La détection de la fumée par ionisation requiert une grande proximité avec le feu. Par conséquent il est impossible de couvrir une large zone et de donner des informations sur la localisation de la naissance d'un incendie, la direction de propagation, l'importance du feu, etc.

La rapide évolution des caméras numériques a permis aux techniques de détections vidéos via l'évolution des ordinateurs de remplacer les méthodes standard de détection du feu et de la fumée. Les techniques de détection vidéo de la fumée et du feu sont particulièrement adaptées en espace ouvert. De plus, ces systèmes offrent la possibilité d'analyser le comportement et la localisation en 3 dimensions du feu. Actuellement les circuits de surveillance vidéos tant en intérieur qu'en extérieur se développent rapidement. Dans ces circonstances, il serait judicieux de développer un dispositif de détection du feu et de la fumée pouvant utiliser ces équipements sans frais.

La recherche dans ce domaine a commencé dans les années 1990. On dénote quelques algorithmes vidéos de détection du feu et des flammes dans la littérature. La majorité de ces algorithmes se concentrent sur les

caractéristiques de couleurs et de formes du feu et de la fumée [1],[2],[3],[4],[5], [6] et [7]. Ensuite, l'objectif fut de construire un vecteur de caractéristiques multidimensionnel utilisé en entrée d'un algorithme conventionnel de classification : SVM, Réseau de neurones, Adaboost, etc. Par conséquent, des experts deviennent indispensables pour créer les règles de base permettant de déterminer les vecteurs de caractéristiques propres à la détection recherchée. Une approche différente et innovante réside dans le fait de ne pas utiliser d'expert pour la construction du vecteur de caractéristique en faisant extraire automatiquement celui-ci par un algorithme d'apprentissage. Les algorithmes d'apprentissage profond possèdent la capacité d'apprendre les caractéristiques utiles à une classification du feu et de la fumée. Les réseaux de neurones convolutifs (CNN) sont de variantes des réseaux profonds et permettent d'extraire les propriétés topologiques d'une image.

Notre approche est conceptuellement simple. Nous utilisons un réseau de neurones convolutifs comme un puissant détecteur du feu et de la fumée dans des images d'une vidéo. L'avantage de cette approche est que cette méthode est insensible au mouvement de la caméra. Le Réseau convolutif opère directement sur les images brutes en RGB dans le spectre du visible. Le réseau convolutif apprend automatique par un ensemble d'images de la base d'entraînement. L'article est organisé de la manière suivante : le paragraphe suivant relate les travaux connexes. Le paragraphe 3 introduit brièvement les réseaux de neurones convolutifs. Le paragraphe 4 quant à lui décrit en détail l'architecture de notre réseau convolutif. Les résultats expérimentaux et la performance se trouvent dans le paragraphe 5. Enfin, le dernier paragraphe discute des limitations de notre méthode et conclut cet article.

## 2 Travaux connexes

Le nombre de documents traitant de la détection de feu vidéo dans la littérature augmente de façon exponentielle. Plusieurs chercheurs ont joué un rôle important dans le développement d'algorithmes de détection de feu par traitement vidéo [6]. Verstockt [1] a proposé un détecteur d'incendie à capteurs multiples qui fusionne les caractéristiques de flammes visuelles et non-visuelles des objets en mouvement. Il a utilisé des images thermiques ordinaires vidéo et infrarouge à ondes longues (LWIR). Tout d'abord, il exploite une soustraction de fond dynamique pour extraire des objets en mouvement. De plus, les objets en mouvement LWIR sont filtrés par segmentation d'objets à chaud basée sur l'histogramme. Un ensemble de caractéristiques de flamme analyse ces objets en mouvement en mettant l'accent sur les caractéristiques géométriques, temporelles

et spatiales distinctives des désordres des régions de flamme. Ensuite, une probabilité de flamme LWIR est calculée en combinant la probabilité du désordre de la fenêtre de délimitation, du désordre d'orientation principal et de la rugosité de l'histogramme des objets en mouvement chaud dans LWIR. Toreyin [2] utilise un algorithme de détection vidéo à quatre étapes. D'abord, il a estimé les pixels et les régions en mouvement en utilisant une méthode de fond hybride: une opération de différenciation à trois images est effectuée pour déterminer des régions de mouvement légitime, suivies d'une soustraction de fond adaptative pour extraire toute la région en mouvement. Deuxièmement, il a utilisé un modèle de mélange gaussien dans l'espace de couleur RVB pour détecter les pixels de couleur de feu. La répartition des couleurs du feu est obtenue à partir d'images échantillonnées contenant des régions de feu. Dans la troisième étape, une transformation en ondelettes temporelles est effectuée pour analyser le scintillement de la flamme. Enfin, une analyse spatiale en ondelettes de régions en mouvement contenant des pixels de masque de feu pour évaluer des variations de couleur en valeurs de pixel est effectuée. Des variations spatiales significatives présupposent la région du feu. Celik [3] a développé deux modèles: un pour la détection d'incendie et l'autre pour la détection de fumée. Un modèle de logique floue basé sur des règles a été utilisé à la place des règles heuristiques existantes. Ce choix a rendu la classification plus robuste en distinguant efficacement le feu et le feu comme des objets colorés. Pour la détection des fumées, une analyse statistique a été effectuée sur la base de l'idée que la fumée présente une couleur grisâtre avec un éclairage différent. Borges [4] a utilisé un vecteur de caractéristiques multidimensionnel comme entrée à un classificateur Bayes. Les caractéristiques sont: la rugosité limite des régions de feu potentiel, le moment statistique du troisième ordre des régions de feu potentiel qui définit l'asymétrie, la variance et enfin la quantité de feu d'un cadre à l'autre (varie en raison de la flamme). Il existe des travaux plus récents inspirés par des recherches antérieures connexes [5], [6]. Comme dit précédemment, toutes ces méthodes sont basées sur la règle ou exigent la construction de caractéristiques discriminantes pour détecter le feu. Une autre solution serait l'utilisation de l'algorithme d'apprentissage en profondeur. Dans la section suivante, nous présentons une variante de l'apprentissage profond: Réseau de neurones convolutifs.

## 3 Réseau de neurones convolutifs

Les réseaux convolutifs ont été introduits pour la première fois par Fukushima [8], il a dérivé une architecture de réseau nerveux hiérarchique inspirée par

le travail de recherche de Hubel [9]. Lecun [10] les a généralisés pour classer les chiffres avec succès et pour reconnaître les numéros de contrôle manuscrit par LeNet-5 qui est montré à la Fig. 1. Ciresan [11] a utilisé les réseaux convolutifs et a réalisé les meilleures performances dans la littérature pour la reconnaissance d'objets multiples pour des bases de données d'images multiples: MNIST, NORB, HWDB1.0, CIFAR10 et l'ensemble de données ImageNet.

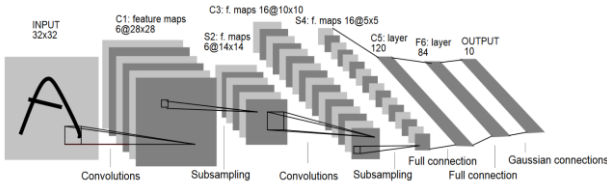


Figure 1. LeNet-5, un réseau neuronal convolutif pour la reconnaissance des chiffres

Un réseau neuronal convolutif se compose de plusieurs couches. Figure. 2, montre ces différentes couches.

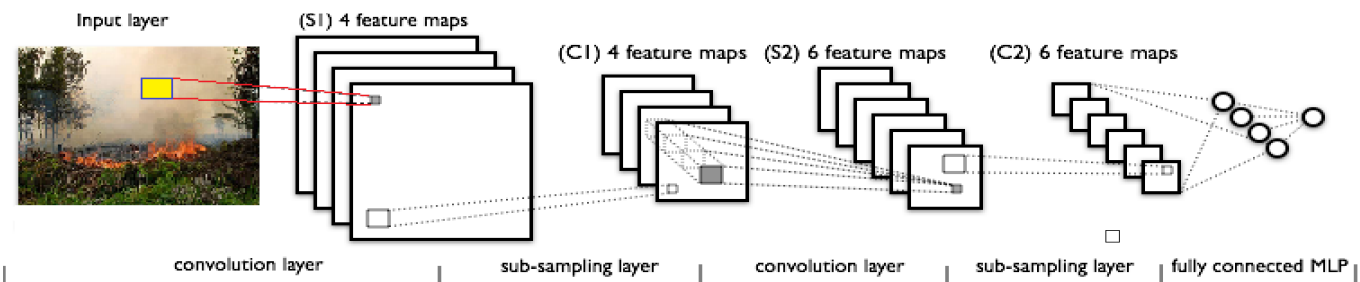


Fig. 2. CNN layers

### 3.1 Couches convolutives

Les couches convolutives constituent le noyau du réseau convolutif. Ces couches se composent d'une grille rectangulaire de neurones qui ont un petit champ réceptif étendu à travers toute la profondeur du volume d'entrée. Ainsi, la couche convolutionnelle est juste une convolution d'image de la couche précédente, où les poids spécifient le filtre de convolution.

### 3.2 Couches de pooling

Après chaque couche convolutive, il peut y avoir une couche de pooling. La couche de pooling sous-échantillonne leur entrée. Il y a plusieurs façons de faire cette mise en commun, comme prendre la moyenne ou le maximum, ou une combinaison linéaire apprise des neurones dans le bloc. Par exemple, la Fig. 3. montre max pooling sur une fenêtre  $2 \times 2$ .

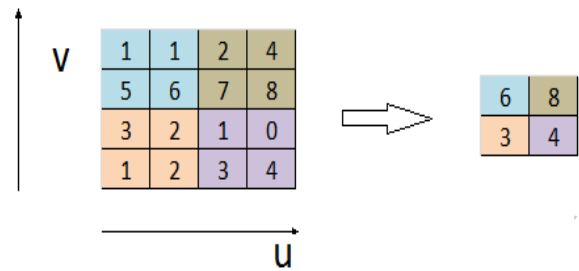


Fig. 3. Max pooling

### 3.3 Couches totalement connectées

Enfin, après plusieurs couches de convolution et pooling, le raisonnement de haut niveau dans le réseau neuronal se fait via des couches totalement connectées.

Dans les réseaux de neurones convolutifs, chaque couche agit comme un filtre de détection pour la présence de caractéristiques spécifiques ou de motifs présents dans les données d'origine. Les premières couches d'un convolutif détectent des caractéristiques qui peuvent être reconnues

et interprétées relativement facilement. Les couches ultérieures détectent de plus en plus des caractéristiques plus abstraites. La dernière couche du réseau convolutif est capable de faire une classification ultra-spécifique en combinant toutes les caractéristiques spécifiques détectées par les couches précédentes dans les données d'entrée. Dans la section suivante, l'architecture proposée du réseau convolutif pour le feu vidéo et la détection de fumée est présentée.

## 4 Réseau de neurones convolutifs pour la détection du feu et de la fumée dans une vidéo

### 4.1 La structure

Notre architecture de classification est classique [17], elle combine convolution et Max pooling. Cependant, pour obtenir une classification rapide permettant une classification et localisation en temps réel, nous avons choisi un réseau léger. Figure. 4 montre les neuf couches de notre réseau convolutif.

Une image couleur RVB dans le spectre visible passe successivement par deux opérations convolutives avec noyau de taille 3x3. La même structure est appliquée après la troisième couche. Un Max pooling 3x3 avec le pas 2 suit les couches convolutives deux et cinq. Les couches un à quatre possèdent 16 cartes de caractéristiques. Les couches cinq et six n'ont qu'une seule carte de caractéristiques. La couche sept et huit sont entièrement connectées. La sortie de la dernière couche totalement connectée alimente un Softmax 3 voies produisant une distribution sur 3 classes.

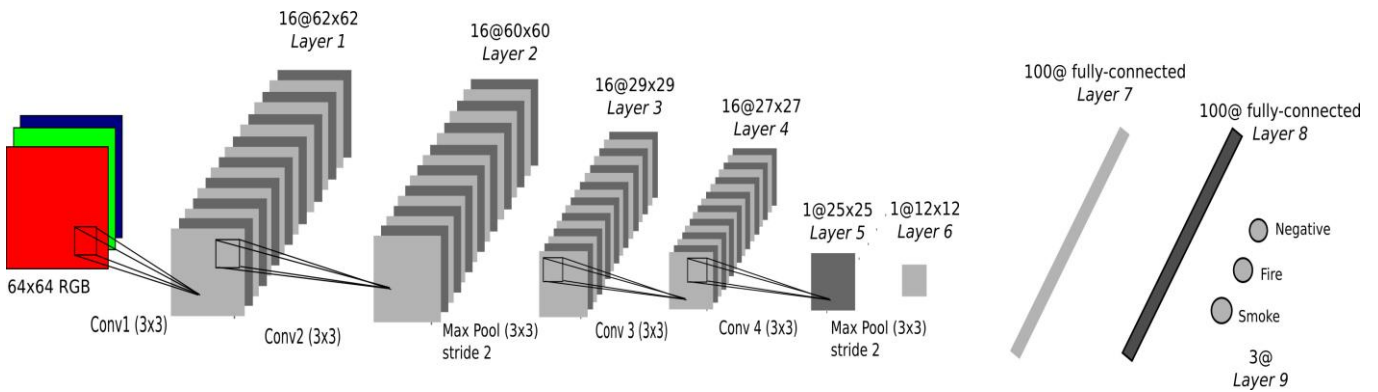


Fig. 4. Architecture de notre réseau convolutionnel

Nous avons choisi [12] et [13] pour les couches convolutives et une entièrement connectée avec une fonction d'activation Leaky ReLu avec coefficient  $a = 1/3$  (voir figure 5).

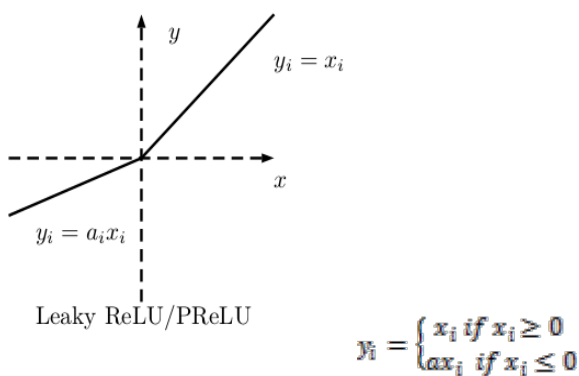


Fig. 5. Leaky ReLu

### 4.2 Entraînement du réseau

L'objectif de notre classification est de décider si une image contient du feu et / ou de la fumée. Pour résoudre ce problème, l'apprentissage du classifieur est réalisé à partir d'une collection d'images dans le spectre du visible en 3 canaux (RVB) étiquetées. En outre, nous voulons

localiser la position du feu et de la fumée dans une image ou dans une vidéo. L'ensemble d'entraînement est composé de 27919 images RVB étiquetées de taille 64x64 pixels. 8915 pour la fumée, 7257 pour le feu et 11752 négatives (pas de feu ou de fumée). Nous créons 3 sous-ensembles: training 60% des images, validation 20% et test 20%. L'entraînement a été réalisé avec un ordinateur composé d'un microprocesseur Intel Xeon (fréquence CPU 3,1Ghz, RAM 16Go) et une carte graphique GTX 980 Ti (2816 cœurs, 6 Go de mémoire). Nous avons utilisé une descente de gradient stochastique (SGD) avec des mini-lots de taille 100. Le poids dans le réseau est initialisé aléatoirement. Le taux

d'apprentissage initial est de 0,01 et le momentum de 0,9. Le taux d'apprentissage diminue d'un facteur 0,95 chaque 5 itérations (epochs). En revanche, le momentum augmente pour atteindre 0.9999. Pour obtenir la meilleure précision possible pour ces paramètres, plusieurs essais ont été effectués. Nous avons implémenté CNN avec Theano [14], [15] et Lasagne [16]. Nous utilisons un dropout de 0,5 dans les deux couches entièrement connectées pour éviter le surapprentissage. L'apprentissage des paramètres du réseau est réalisé sur 100 cycles ou itérations.

### 5 Résultats

La précision de classification sur l'ensemble test est de 97,9%. L'ensemble d'essais est composé de 1427 images d'incendie, 1758 images de fumée et 2399 images négatives, donc 5584 images. Les images de fumée représentent plus une texture que l'objet « fumée » en lui-même. Les tableaux 1 à 3 donnent la matrice de confusion pour chaque classe. Sur la matrice de confusion du feu, les faux négatifs et les faux positifs ne contiennent pas d'images de fumée. De la même manière, la matrice de confusion des fumées ne contient pas d'image de feu pour les faux négatifs et les faux positifs. Nous pouvons conclure que les paramètres de notre modèle de réseau convolutif permettent une bonne distinction de classification entre le feu et la fumée. En



outre, les surfaces sous la courbe ROC (Receiver Operating Characteristic) de la Fig. 6, pour les trois classes sont proches de l'unité, indiquant une bonne classification sur l'ensemble d'essai. La courbe ROC feu a une aire plus grande que les autres, ce qui indique une meilleure classification des performances pour le feu.

TABLE I. CONFUSION MATRIX POUR LE FEU

Fire	True class		
	True	False	
Hypothesis class	True	1400	3 <sup>a</sup>
	False	27 <sup>a</sup>	4154

<sup>a</sup>not smoke images

TABLE II. CONFUSION MATRIX POUR LA FUMÉE

Smoke	True class		
	True	False	
Hypothesis class	True	1698	26 <sup>a</sup>
	False	60 <sup>a</sup>	3800

<sup>a</sup>not fire images

TABLE III. CONFUSION MATRIX POUR LES NEGATIFS (NI FIRE/SMOKE)

No Fire/Smoke	True class		
	True	False	
Hypothesis class	True	2370	87 <sup>a</sup>
	False	29 <sup>b</sup>	3098

<sup>a</sup>Image Fire 27 – image Smoke 60  
<sup>b</sup>Image Fire 3 – image Smoke 26

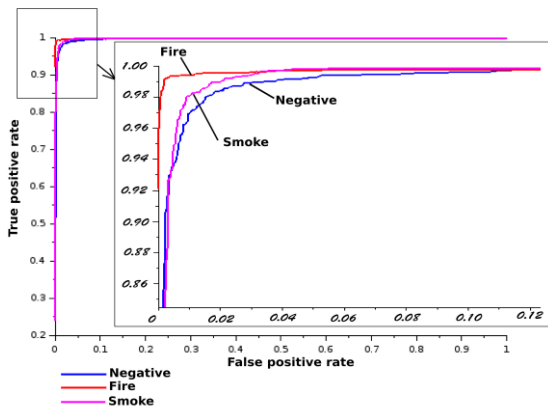


Fig. 6. Courbe ROC pour les 3 classes: fumée, feu et négative

Notre objectif est de détecter en temps réel le démarrage du feu ou de caractériser un incendie sur une vidéo. Le temps de traitement pour la détection est un facteur clé avec la précision. Par conséquent, nous décidons d'utiliser la "structure légère" décrite dans la Fig. 4. Les méthodes actuelles utilisent des fenêtres coulissantes pour détecter et classer l'objet sur des images originales

ou redimensionnées. Chaque fenêtre passe par le réseau de neurones convolutif et les couches totalement connectées pour finir par le classifieur. Pour analyser l'image entière d'un cadre vidéo, la position de la fenêtre doit changer et aller de nouveau à travers le réseau neuronal convolutif. Notre approche est tout à fait différente, au lieu de glisser une fenêtre de 64x64 pixels dans l'image d'origine RGB pour localiser le feu et la fumée; Nous décidons de travailler sur la dernière carte de caractéristiques. Nous divisons le réseau en deux parties. La première partie est composée de 6 couches: 1 à 6 (couches convolutives et Max pooling), la deuxième partie est composée uniquement des entièrement connectées Fig. La figure 7 montre cette architecture CNN.

En utilisant la première partie du réseau, nous évaluons la dernière carte de caractéristiques (couche 6) de l'image entière. Nous savons que sur la base de la structure du réseau de la CNN, une fenêtre coulissante de taille 64x64 pixels dans l'image RVB correspond à une taille de fenêtre de 12x12 pixels dans la dernière carte de caractéristiques. Fig 8.

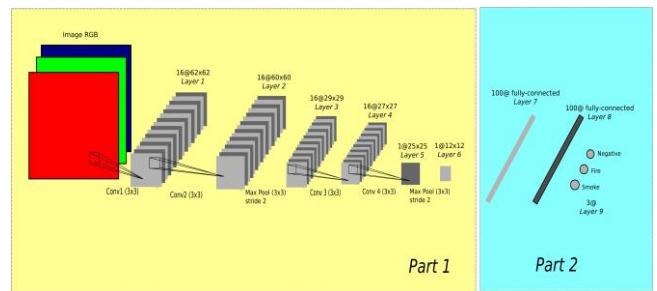


Figure. 7. Deux parties de notre réseau convolutif. Partie 1: 6 couches (convolutives, maxpooling). Partie 2: Deux couches totalement connectées et la couche de sortie.

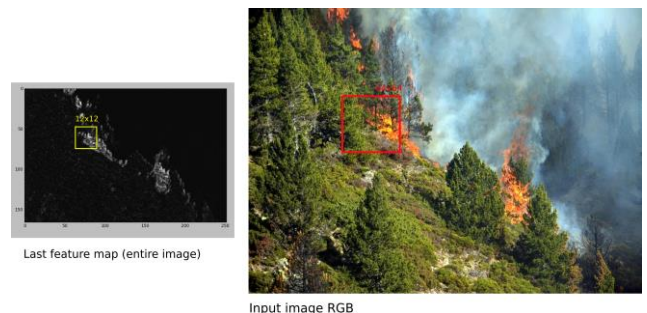


Figure. 8. Fenêtre glissante: Image originale vs Carte des caractéristiques

Pour détecter le feu et la fumée dans l'image d'une vidéo, nous appliquons une fenêtre glissante de taille 12x12

pixels sur la dernière carte de caractéristiques (voir la figure 8). Pour accélérer la prédiction pour chaque fenêtre 12x12 de l'image entière, nous réalisons un tenseur 12x12x1xN (N: nombre de fenêtres permettant la détection du feu ou de la fumée) à partir de la dernière carte de caractéristiques et nous utilisons le GPU de la carte graphique. Avec cette méthode, la précision semble inchangée et la vitesse de détection et de prédiction augmentent en fonction de la taille d'origine de l'image et du nombre de fenêtres à prédire comme représenté sur la Fig. 9. et Fig. 10.

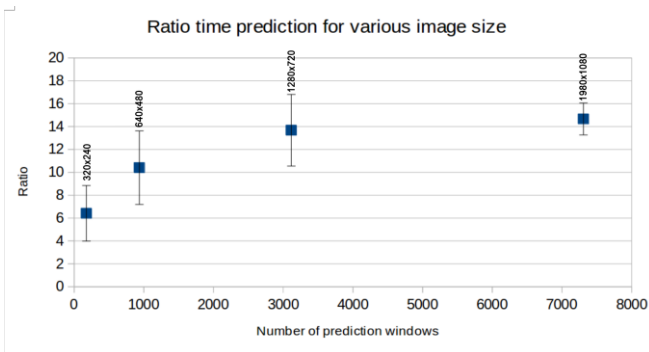


Figure. 9. Ratio du temps de prédiction pour différentes tailles d'image.

Ratio = (temps de prédiction pour l'image originale entière) / (temps de prédiction sur la dernière carte de fonction + temps pris pour réaliser la dernière carte de caractéristique). Fenêtres glissantes par pas de 16px

sur l'image originale et 4px sur la dernière carte de caractéristiques. Expérience réalisée sur 200 images d'une vidéo.

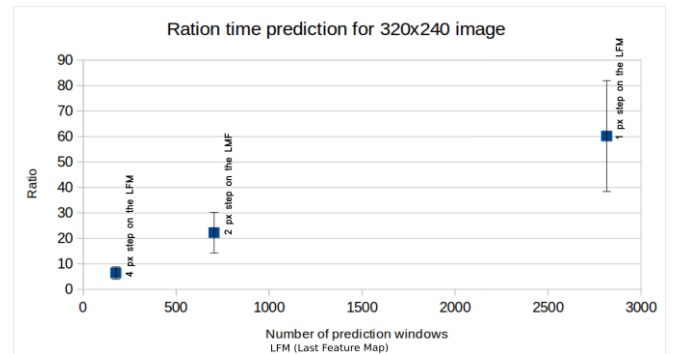


Figure. 10. Ratio du temps de prédiction pour une taille d'image 320x240 avec différent pas de glissement. Expérience réalisée sur 200 images d'une vidéo.

La classification et la localisation affichée à la Fig. 11 et 12, a été effectuée en faisant glisser une fenêtre de 12x12 sur la carte de caractéristique par pas de 2 pixels. Figure. 9.1 (c) et la Fig. 9.2. C) décrivent le masque de détection et de localisation. La couleur rouge représente l'incendie détecté et la couleur verte représente la fumée détectée. L'intensité des couleurs rouge et verte varie en fonction de la probabilité de détection de feu et / ou de fumée. La localisation du feu et de la fumée dans la carte des

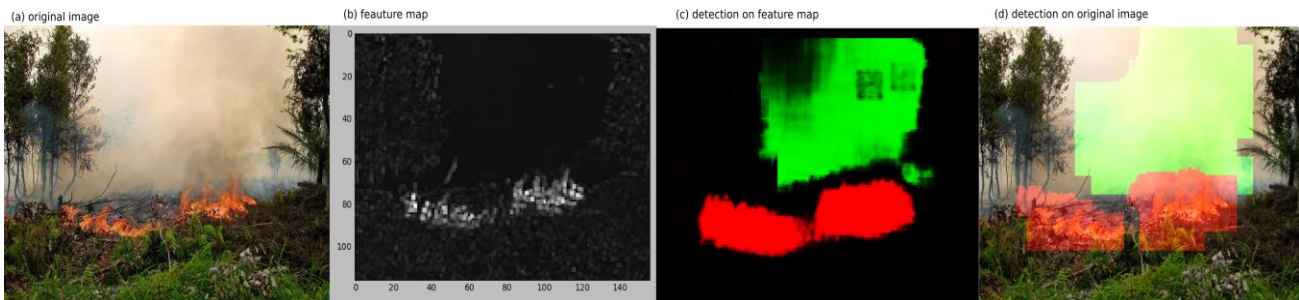


Fig. 11. Feu/fumée détection sur une image de forêt

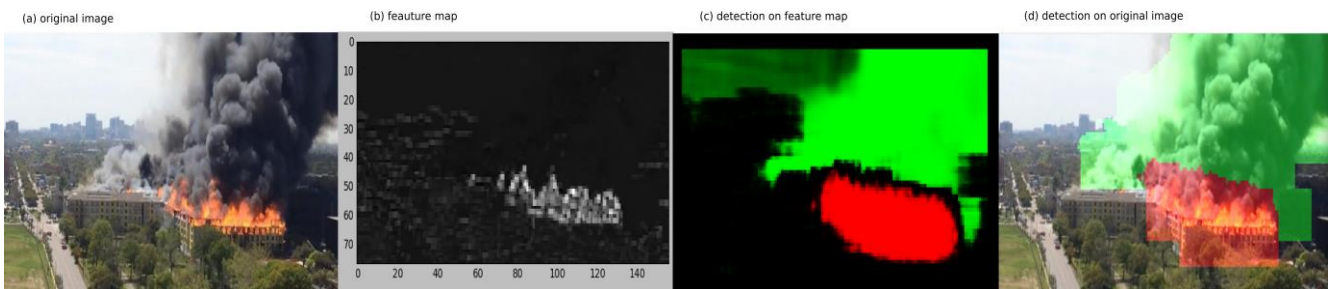


Fig. 12. Feu/fumée détection sur une image d'immeuble

caractéristiques permet de projeter ces positions sur l'image RVB originale (Rectangles rouge et vert).

## 6 Conclusion

Dans cet article, une méthode de détection de fumée et de feu dans le spectre du visible a été présentée. L'algorithme proposé utilise une approche d'apprentissage profond basée sur le réseau neuronal convolutif. La matrice de confusion et les courbes ROC indiquent une très bonne précision globale pour l'étape de détection. Nous avons montré que notre méthode de travail sur la dernière carte de caractéristique permet de diminuer le temps de prédiction d'un facteur allant de 6 à 60.

Dans les travaux futurs, nous souhaitons améliorer la méthode en utilisant un réseau de neurones convolutionnels 3D pour prendre en compte la dynamique temporelle dans notre modèle. En effet, le réseau convolutif est actuellement limité à gérer les entrées 2D qui nous amènent à traiter l'entrée vidéo seulement image par image. En revanche, le réseau convolutif 3D extrait des caractéristiques spatiales et temporelles en effectuant des convolutions 3D. Ainsi, l'information de mouvement du feu et de la fumée pourrait être encodée, ce qui permettrait de diminuer considérablement le temps de détection. De plus, pour optimiser la détection et la localisation de la fumée et du feu sur une vidéo, nous devons améliorer notre base d'image. La fumée est plus difficile à détecter et à localiser en raison de la nature de sa forme et de sa texture. Nous pourrions améliorer cette détection en augmentant la taille du réseau ou en ajoutant des images de l'objet « fumée » dans la base d'apprentissage. Notre modèle ne détecte que le feu rouge, pour détecter d'autres couleurs de feu, nous devons augmenter notre base d'image avec d'autres couleurs de feu telles que bleue, etc ... En outre, nous prévoyons de comparer notre algorithme aux méthodes conventionnelles sur une plus grande variété des images de feu vidéo: différents matériaux, sources et ventilations. Enfin, une approche intéressante serait de coupler notre réseau convolutionnel dans le spectre du visible avec une image infrarouge dans la bande LWIR afin de focaliser notre recherche dans une zone déterminée de l'image. Ceci permettrait également de différencier la fumée et les nuages.

## References

- [1] S. Verstockt, A. Vanoosthuyse, S. Van Hoecke, P. Lambert, and R. Van de Walle, Multi-sensor fire detection by fusing visual and non-visual flame features, In Proceedings of International Conference on Image and Signal Processing, June 2010, pp. 333–341.
- [2] B. U. Toreyin, Y. Dedeoglu, U. Gudukbay, A. E. Cetin, Computer vision based method for real-time fire and flame detection, Pattern recognition letters, 2006, 27,1, pp. 49-58.
- [3] T. Çelik, H. Özkaramanlı and H. Demirel, Fire and smoke detection without sensors: Image processing based approach, *Signal Processing Conference, 2007 15th European*, Poznan, 2007, pp. 1794-1798.
- [4] K. Borges, P. Vinicius, J. Mayer and E. Izquierdo, Efficient visual fire detection applied for video retrieval, *Signal Processing Conference, 2008 16th European*, IEEE, 2008.
- [5] K. Poobalan and S. Liew, Fire detection algorithm using image processing techniques, Proceedings of the 3<sup>rd</sup> International Conference on Artificial Intelligence and Computer Science (AICS2015), October 2015, pp. 160-168
- [6] P. Gomes, P. Santana and J. Barata, A vision-based approach to fire detection, *International Journal of Advanced Robotic Systems*, 09-2014.
- [7] E. Çetin et al, Video fire detection – Review, *Digital Signal Processing*, Volume 23, Issue 6, December 2013, pp. 1827-1843
- [8] K. Fukushima, Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position. *Biological Cybernetics*, 1980, 36(4), pp.193–202.
- [9] D. H. Hubel and T. N. Wiesel, Ferrier lecture: Functional architecture of macaque monkey visual cortex, *Proceedings of the Royal Society of London, Series B, Biological Sciences*, 1977, 198(1130):pp. 1–59.
- [10] Y. Lecun, L. Bottou, Y. Bengio and P. Haffner, Gradient-based learning applied to document recognition, in *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278-2324, Nov 1998.
- [11] D. Ciresan, U. Meier; J. Masci; L.M. Gambardella and J. Schmidhuber, *Flexible, High Performance Convolutional Neural Networks for Image Classification, Proceedings of the Twenty-Second international joint conference on Artificial Intelligence-Volume Volume Two 2: pp. 1237–1242, November 2013.*
- [12] X. Bing, N. Wang, T. Chen and M. Li, Empirical Evaluation of Rectified Activations in Convolutional Network, *CoRR abs/1505.00853 (2015): n. pag.*
- [13] A. L. Maas, A. Y. Hannun and A. Y. Ng, Rectify nonlinearities improve neural network acoustic model, *ICML 2013 Workshop on Deep Learning for Audio, Speech, and Language Processing*, June 2013, Atlanta
- [14] F. Bastien, P. Lamblin, R. Pascanu, J. Bergstra, I. Goodfellow, A. Bergeron, N. Bouchard, D. Warde-Farley and Y. Bengio. "Theano: new features and speed improvements". *NIPS 2012 deep learning workshop*.
- [15] J. Bergstra, O. Breuleux, F. Bastien, P. Lamblin, R. Pascanu, G. Desjardins, J. Turian, D. Warde-Farley and Y. Bengio. "Theano: A CPU and GPU Math Expression Compiler". *Proceedings of the Python for Scientific Computing Conference (SciPy) 2010. June 30 - July 3, Austin*
- [16] LASAGNE, Lightweight library to build and train neural networks in Theano, <https://github.com/Lasagne/Lasagne>, 13 August 2015.
- [17] S. Frizzi, R. Kaabi, M. Bouchouicha, J. M. Ginoux, E. Moreau and F. Fnaiech, "Convolutional neural network for video fire and smoke detection," *IECON 2016 - 42nd Annual Conference of the IEEE Industrial Electronics Society*, Florence, 2016, pp. 877-882.



# Contrôle par apprentissage constructiviste pour le trafic coopératif

M. Guériau<sup>1,2</sup>F. Armetta<sup>1</sup>S. Hassas<sup>1</sup>R. Billot<sup>3</sup>N-E. El Faouzi<sup>2</sup><sup>1</sup> Univ. Lyon, UMR CNRS 5205 LIRIS, F-69622 Villeurbanne, France<sup>2</sup> LICIT, Univ. Lyon – IFSTTAR, LICIT, F-69675 Bron – ENTPE, LICIT, F-69518 Vaulx En Velin, France<sup>3</sup> IMT Atlantique, Lab-STICC, Univ. Bretagne Loire, F-29238 Brest, France

maxime.gueriau@ifsttar.fr

## Résumé

Lorsqu'un système autonome évolue dans un environnement complexe, en partie inconnu ou dynamique, il n'est pas possible de fournir une représentation exhaustive a priori facilitant son processus de prise de décision. Pour illustrer ce problème, nous choisissons le cas du contrôle décentralisé du trafic coopératif, où une unité d'infrastructure est en charge de réguler localement le flux, en envoyant des consignes aux véhicules connectés. Ce contrôle est le fruit d'une stratégie construite par l'apprentissage d'une représentation précise (états perception-action) des différents états de trafic. Nous proposons un modèle capable, sans connaissances expertes, d'utiliser un ensemble de méthodes de classification représentées sous la forme d'une population d'agents et de les combiner dynamiquement pour construire une représentation précise de l'environnement. Notre approche s'inscrit dans une démarche d'apprentissage constructiviste où la population d'agents construit collectivement une représentation qui exploite, suivant l'usage, les discrétisations possibles de l'espace de perception proposées par les individus.

## Mots Clef

Apprentissage constructiviste, prise de décision, contrôle.

## Abstract

Decision making in autonomous systems is particularly challenging in unknown and changing complex environments, where providing a complete a priori representation is not possible. To illustrate the problem, we consider a decentralized control of road traffic, where a control device of the distributed infrastructure locally controls traffic by sending recommendation messages to connected vehicles. We propose an approach able to combine, without prior domain-knowledge, a set of existing traditional unsupervised learning methods that collaborate as a population of agents in order to build an efficient representation. Our approach follows a constructivist learning perspective, where the population of agents is able to collectively build a representation that dynamically combines discretizations from the individuals.

## Keywords

Constructivist learning, decision making, control.

## 1 Contexte

La pertinence de la décision prise par un système autonome repose directement sur sa capacité à discriminer ses différentes interactions sensorimotrices. Cette compétence dépend de la précision de la représentation faite de l'environnement par le système mais aussi du problème visé. Dans les approches classiques en Intelligence Artificielle (IA), cette représentation est donnée au système a priori sous la forme de connaissances expertes. Dans le cas d'un environnement inconnu a priori et/ou dynamique, il est plus difficile de fournir une telle représentation. Le problème est donc plutôt de permettre au système de construire dynamiquement une représentation assurant une prise de décision efficace.

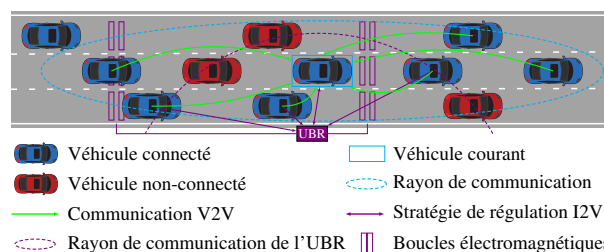


FIGURE 1 – Architecture générale des systèmes de transport intelligents coopératifs.

Comme exemple d'un environnement de prise de décision complexe, nous avons choisi le cas de la régulation du trafic routier dans le contexte des systèmes de transport intelligents coopératifs (C-ITS), illustré par la figure 1. Ces systèmes complexes s'appuient sur les avancées récentes en matière de communication et de technologies de l'information afin d'améliorer l'écoulement du flux de trafic. Dans un futur proche, l'infrastructure routière sera partagée par des véhicules connectés et non-connectés. Les véhicules connectés pourront, grâce à un protocole de communication sans fil, échanger des informations entre eux

(véhicule à véhicule – V2V) et avec l'infrastructure (infrastructure à véhicule – I2V et véhicule à infrastructure V2I). Dans les C-ITS, les unités de contrôle côté infrastructure, nommées Unités de Bord de Route (UBR), ont pour rôle de collecter des données sur leur section dédiée. Plus que de simples relais d'information, une UBR peut aussi jouer le rôle d'unité de régulation décentralisée du trafic en propageant des consignes aux véhicules connectés par communication I2V.

## 1.1 Contributions

Le premier objectif de ce travail est de proposer un modèle permettant à un système autonome de construire une représentation à partir de ses interactions bas-niveau avec son environnement (section 2). Le modèle suit une perspective constructiviste et couple deux processus inter-dépendants : la construction d'hypothèses d'"états perception-action" et l'élaboration d'une "stratégie de contrôle". Ces processus sont gérés par une population d'agents en compétition pour construire une représentation précise de l'environnement (ensemble d'états perception-action) et d'un processus de prise de décision permettant de choisir une action pertinente. Ce dernier mécanisme est modélisé comme une forme d'apprentissage par renforcement où le retour sur expérience est fourni par l'environnement et utilisé pour renforcer les hypothèses d'état perception-action proposées par les agents. L'originalité de notre contribution est d'utiliser un ensemble de représentations concurrentes modélisées sous la forme d'agents autonomes qui s'adaptent dynamiquement grâce à des mécanismes d'intelligence collective, dans le but de produire une représentation émergente d'un contexte dynamique.

Le second objectif de l'article est de proposer une application innovante de notre modèle conceptuel. Nous avons choisi l'environnement du véhicule connecté où une unité décentralisée de l'infrastructure connectée (UBR) est un agent de contrôle du trafic qui implémente notre modèle. La partie 3 montre des résultats intéressants de notre travail avec une implémentation simplifiée du modèle. Les résultats montrent également des voies d'amélioration possibles du modèle.

## 1.2 État de l'art

Les approches classiques en IA reposent sur une abstraction de l'environnement proposée par le concepteur du système. Cette représentation dépend très fortement des spécificités du problème visé et de l'environnement associé (capteurs/effecteurs). Un système exploitant une telle représentation aura donc une capacité d'adaptation à d'autres problèmes/environnements limitée [3]. Comme l'autonomie d'un système dépend de sa capacité d'adaptation, les approches constructivistes proposent des mécanismes de construction itérative de la représentation de l'environnement par un agent, à partir de ses interactions. Ces approches s'inspirent des travaux précurseurs en sciences cognitives de Piaget [9] sur le développement de l'enfant, et ont convergé vers la définition de caractéristiques per-

mettant de modéliser l'autonomie [12] : l'agent doit être incarné, situé, et exploiter un processus développemental épigénétique. La notion d'épigénétique a été proposée par Piaget, et décrit le processus de développement de l'individu à travers ses interactions avec son environnement. L'approche la plus prometteuse en IA développementale consiste à reproduire le processus d'acquisition de connaissance de l'être humain. Seuls quelques travaux novateurs [7] proposent de construire une représentation uniquement à partir des interactions sensorimotrices de l'agent. Ces approches se heurtent au problème de l'amorçage de la représentation, qui survient dès que l'on cherche à démanteler la construction de la représentation de l'environnement, notamment sans utiliser de connaissances expertes *a priori*. Pour contourner ce problème, quelques approches proposent d'aider le système en lui donnant quelques mécanismes de bas niveau, permettant d'amorcer le processus d'apprentissage [10]. Le modèle des auteurs utilise le gaz neuronal croissant (GNG) - une adaptation des cartes auto-organisées - afin de générer une première discrétisation de l'espace des perceptions du système. Cela permet de réduire le nombre de dimensions des données perçues (espace d'entrée) afin de simplifier l'identification des actions les plus pertinentes correspondant à cet espace. Cependant, pour fournir une telle discrétisation, il faut réaliser un partitionnement des données perçues en amont du processus d'apprentissage du système. Cela s'oppose en partie au cadre strict du constructivisme où cette discrétisation devrait être le résultat des interactions avec l'environnement. Le système doit donc être en mesure de faire évoluer cette représentation. Une façon d'offrir cette possibilité est de fournir différentes discrétisations au système, qu'il devra combiner afin d'en déterminer les plus utiles à son usage (actions). Dans cet article, nous proposons des mécanismes permettant au système d'utiliser plusieurs représentations concurrentes, qui sont mises à jour en parallèle. Ces abstractions (discrétisations) des données brutes perçues peuvent être générées en amont de la phase d'apprentissage du système, à partir de différents algorithmes d'apprentissage (K-moyennes, GNG). D'un point de vue de l'ingénierie du trafic, les approches classiques de régulation sont définies comme un ensemble de règles expertes qui s'appliquent dans un contexte discret et prédéfini. Ce type de stratégies peut aussi bien s'appliquer au niveau du système [8] qu'à une échelle plus locale [4]. Les SMA sont largement utilisés pour traiter des problèmes de gestion et de régulation des systèmes de transport [2], mais la plupart des contributions s'intéressent à la coordination d'intersections. Le déploiement des véhicules connectés a tout de même suscité un intérêt plus récent vers le contrôle sur autoroute. Ces approches proposent des stratégies dynamiques exploitant les technologies basées sur la communication, comme les limitations de vitesse variables, et permettent des stratégies de régulation plus décentralisées.

## 2 Construction d'une représentation par apprentissage concurrent

L'objectif à long terme de notre modèle est de concevoir un système de prise de décision dynamique capable de définir les frontières de chaque état perception-action, afin d'identifier précisément le périmètre de chaque action étant donné un ensemble de perceptions. En effet, un tel système serait capable de mieux comprendre les dynamiques de l'environnement et aurait une probabilité plus élevée de sélectionner l'action la plus pertinente dans le contexte actuel estimé. Plusieurs problèmes sont à traiter afin d'arriver à un tel résultat. D'abord, il est nécessaire de choisir quelles sont les connaissances de base à fournir au système sans pour autant incorporer trop de connaissances expertes dans sa représentation. Ensuite, le périmètre exact des actions peut ne pas être atteignable selon le problème. La difficulté est de trouver la meilleure façon de modéliser cette notion de périmètre sans pour autant limiter l'ensemble des états possibles. Enfin, le défi le plus intéressant se situe au niveau du renforcement de la représentation, qui doit mener à un contrôle efficace de l'environnement par le système. Nous proposons un modèle conceptuel générique dont les mécanismes généraux traitent les problématiques précédentes, et qui peut être facilement adapté dans son implémentation pour traiter des problèmes spécifiques.

### 2.1 Aperçu

Le modèle proposé s'appuie sur des mécanismes d'intelligence collective afin d'assister le système dans sa tâche itérative de construction de sa représentation de l'environnement. Cette tâche prend la forme de deux processus couplés, perception et décision, comme illustré dans la figure 2. Les données d'entrée, collectées par les capteurs,

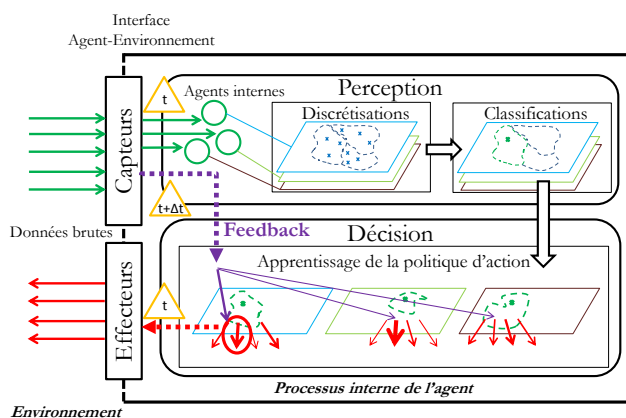


FIGURE 2 – Aperçu du modèle.

représentent des observations de l'environnement et constituent la perception de bas-niveau du système. Ces données sont continues et non-identifiées dans la mesure où aucune information experte n'est communiquée au système au sujet de la pertinence ou de l'importance de chaque variable perçue. Cette perception est transmise aux agents "discrét-

iseurs", dont le rôle est de fournir des abstractions à partir de ces informations. Ces abstractions correspondent à des façons alternatives d'interpréter le signal d'entrée. Chaque agent applique une stratégie de discrétisation interne. Les discrétisations générées peuvent évoluer en temps réel pour adapter le périmètre des (états) classes que l'agent propose. Plusieurs mécanismes peuvent améliorer la précision des discrétisations proposées par les agents. Le processus multi-agent produit un ensemble de partitions concurrentes pour le système correspondant à différentes représentations possibles. La sélection de l'action repose sur un apprentissage par renforcement qui fait intervenir toutes les représentations concurrentes. Le modèle, présenté dans

#### Algorithme 1 Processus général du modèle

```

Instancier l'ensemble  $A$  des actions possibles
Instancier l'ensemble  $D$  des agents discrétiseurs
Instancier l'ensemble  $S$  des états sélectionnés concurrents (vide)
répéter # boucle qui représente le processus de
  décision du système
  si  $S$  n'est pas vide alors # bloc non exécuté à la
    première itération
     $F = \text{feedback}(P)$ 
    pour tout  $s_i \in S$  faire # récupérer le lien  $L$  entre
      l'état  $s_i$  et l'action  $a^*$ 
       $L = \text{getPercepActionLink}(s_i, a^*)$ 
       $\text{incrementReward}(L, F)$ 
    fin pour
     $\text{reinforce}(D, a^*, F)$ 
  fin si
   $P = \text{getPerception}()$  # début du processus de
    sélection de l'action
   $S.\text{removeAll}()$ 
  pour tout  $d_i \in D$  faire
     $p_i = \text{discretizerPerception}(P)$  # récupération
      de la perception de l'agent
     $d_i.\text{discretize}(p_i)$ 
     $s_i = D_i.\text{classify}(p_i)$ 
     $S.\text{add}(s_i)$  # ajout de l'hypothèse d'état  $s_i$  à
      l'ensemble  $S$ 
  fin pour
   $a^* = \text{actionSelection}(S)$ 
   $\text{execute}(a^*)$ 
fin répéter # fin du processus de décision avant la
  prochaine perception (et feedback)

```

l'algorithme 1, vise à combiner dynamiquement les états perception-action précédemment appris afin de construire une représentation encore plus précise de l'environnement.

### 2.2 Perception de haut-niveau

Le système utilise des représentations individuelles proposées par ses agents discrétiseurs pour construire une représentation de plus haut-niveau. Chaque discrétiseur exploite son propre processus de classification exécuté dans

son propre espace de perception, qui peut varier d'un agent à l'autre. Cela offre la possibilité d'évaluer des représentations dédiées à des canaux spécifiques de la perception. Ainsi, les discrétiseurs peuvent percevoir des données à partir des mêmes capteurs ou utiliser des dimensions différentes à l'échelle de la perception du système. Nous proposons de décrire le modèle de façon générique, ce qui permet de repousser le choix de la méthode de classification à l'étape d'implémentation du modèle. Les algorithmes en-ligne ou hors-ligne peuvent être utilisés à la condition qu'ils soient compatibles avec le type des données d'entrées et qu'ils permettent de générer une représentation sous la forme d'un ensemble d'états discrets. L'ensemble des variables perçues par chacun des agents peut être défini de manière experte ou initialisé aléatoirement, en fonction du problème visé. Le processus de perception du système peut se résumer en deux étapes consécutives (étapes 1 et 2), telles que représentées dans la figure 3. Chaque fois que le système

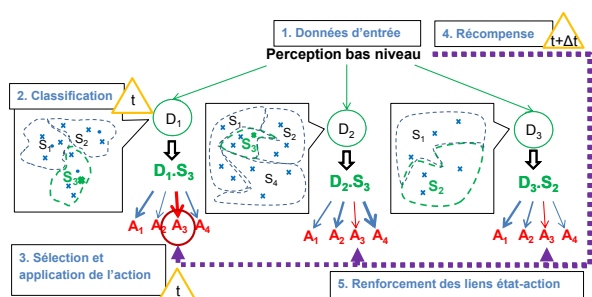


FIGURE 3 – Processus de discrétisation, de classification et de sélection d'action.

perçoit son environnement, tous les agents discrétiseurs  $D$  mettent à jour en parallèle leur propre espace d'hypothèses d'état, en fonction de leur méthode de classification choisie (K-means, GNG, etc.). Par exemple, la discrétisation de l'agent  $D_i$  est composée de l'ensemble des hypothèses d'état :  $\{S_1, S_2, \dots, S_j\}$ . Pour chaque pas de temps de perception du système, chaque agent  $D$  perçoit ses variables d'entrée et classe sa perception comme une hypothèse d'état discret. À partir de l'ensemble des hypothèses d'état (une par agent), le système tente de choisir l'action la plus pertinente dans le contexte défini par les différents agents.

### 2.3 Contrôle de l'environnement

Les discrétiseurs mémorisent la probabilité de sélection de chaque action lorsqu'ils classent leur perception en tant qu'une hypothèse d'état donnée. Ces probabilités sont modélisées sous la forme de liens état-action qui permettent de garder une trace au niveau du système du résultat des actions consécutives, évalué à travers le retour sur expérience (*feedback*) fourni par l'environnement. À l'issue de l'étape de perception, tous les discrétiseurs ont sélectionné une de leurs hypothèses d'état en réponse à la perception de bas-niveau propagée. Le système doit désormais choisir une action qui soit la plus pertinente possible en fonction

des différentes hypothèses proposées (une par agent). Le système doit alors trouver un compromis entre l'exploitation (appliquer l'action générant le meilleur *feedback*) et l'exploration (tester des actions non utilisées jusque là). Ce choix est illustré par les étapes 3 à 5 dans la figure 3. Le système fait face à un dilemme classique d'exploration-exploitation [1], puisque l'espace des actions est supposé fini et qu'il n'y a pas de corrélation évidente entre des hypothèses d'état distinctes. Cette situation fait écho au problème dit du bandit à  $n$  bras, relatif à l'élaboration d'une stratégie optimale de sélection et de test de différentes machines à sous dans un casino. Plusieurs stratégies, disponibles dans la littérature, peuvent être utilisées dans notre modèle. La plupart utilisent un paramètre objectif à minimiser basé sur la notion de regret, exprimé comme la différence entre la récompense courante et la récompense optimale (théoriquement inconnue). Notre modèle ne nécessite pas que la stratégie soit optimale dans la mesure où la sous-exploration de l'espace des actions risque de nuire aux actions prises par le système. Néanmoins, cette hypothèse dépend encore une fois du type de problème visé et pourra être étudiée en comparant différentes implémentations du modèle.

Le retour sur expérience fourni au système fait partie de sa perception de bas-niveau et doit être défini lors de la formalisation du problème (*i.e.* pour l'implémentation). Cette fonction de *feedback* permet une évaluation par le système du résultat de ses actions dans l'environnement. Cette valeur est utilisée comme une récompense pour l'algorithme d'exploration-exploitation et contribue au renforcement de la représentation construite. Le résultat de la sélection est un lien état-action dont chaque état est directement lié à un agent discrétiseur. Ainsi, le processus de renforcement contribue à l'apprentissage du système, en l'aidant à identifier les hypothèses d'état les plus précises parmi l'ensemble des discrétisations proposées par la population d'agents discrétiseurs.

## 3 Application au trafic coopératif

Le modèle générique présenté dans cet article a été implémenté dans le cadre d'une application innovante : la régulation du trafic coopératif. Le système prend place dans une unité de bord de route, chargée de réguler dynamiquement un flux de trafic. Pour cela, il est équipé de capteurs permettant de percevoir les véhicules connectés et les informations du flux (débits/vitesses). Un agent discrétiseur est associé à chacun de ces capteurs. Le cadre de simulation utilisé pour les expérimentations est présenté dans [5]. Nous donnons ici les détails de l'implémentation de notre modèle de contrôle : description des caractéristiques des agents discrétiseurs et actions/*feedback* utilisés. Ensuite, les résultats en simulation mettent en avant le comportement du modèle et confirment les bénéfices de notre approche qui, grâce à des représentations concurrentes, permet de réguler un flux partiellement composé de véhicules connectés.



### 3.1 Cas d'étude

Nous avons choisi le cas d'étude du contrôle du trafic coopératif pour implémenter notre modèle et étudier les bénéfices de notre approche. Nous visons une situation de trafic mixte où le flux est composé de véhicules connectés et non-connectés. Les véhicules connectés peuvent échanger des informations entre eux (communication V2V) et avec l'infrastructure (I2V/V2I). L'infrastructure est composée d'unités de bord de route (UBR), en charge de réguler le trafic sur leur section dédiée. Une UBR perçoit des informations lui permettant d'estimer l'état de trafic courant grâce à ses capteurs (par exemple, des boucles électromagnétiques qui permettent d'estimer le débit et les vitesses sur chaque voie). L'objectif d'une UBR est de propager des consignes dans des messages à destination des véhicules connectés. Une telle stratégie de contrôle décentralisée vise à réduire ou à retarder le phénomène de congestion.

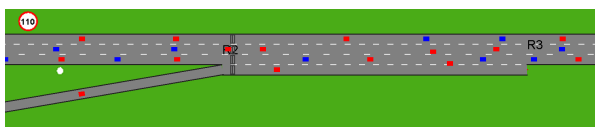


FIGURE 4 – Capture d'écran du scénario dans le simulateur.

Comme illustré par la figure 4, toutes les expérimentations sont menées dans un même scénario représentant une section autoroutière rectiligne à trois voies avec la présence d'un convergent (voie d'insertion) vers sa fin. L'UBR (représentée par le disque blanc) perçoit les débits et les vitesses grâce aux capteurs sur la route (rectangles gris) ainsi que les vitesses des véhicules connectés dans son rayon de communication (150 mètres). L'unité de bord de route implémente le modèle afin de choisir les consignes les plus pertinentes à envoyer aux véhicules connectés pour éviter, ou au moins réduire, la congestion sur sa section. Les consignes sont propagées sous la forme de messages I2V aux véhicules connectés. La voie d'insertion, à cause de la différence de vitesse avec la section principale, déclenche, suivant les conditions de trafic dense, l'apparition de congestion qui se propage depuis la voie de droite jusqu'à l'ensemble des voies. Nous étudions l'effet de la stratégie de régulation d'une UBR sur le flux en faisant l'hypothèse d'un pourcentage fixé à 30% de véhicules connectés, mélangés de façon homogène parmi les véhicules non connectés sur la section principale.

**Simulation.** Le scénario choisi pour la simulation est modélisé dans le simulateur MASCAT [5] (*Multi-Agent Simulator for Connected and Automated Traffic*), une extension multi-agent du simulateur de trafic microscopique MovSim [11]. Cette plateforme de simulation permet d'instancier différents modèles microscopiques (nommés "lois de poursuite") qui décrivent le comportement longitudinal des véhicules. Le réseau routier est représenté sous la forme d'un graphe qui permet de modéliser les sections

multi-voies, les limitations de vitesses et les connexions entre les sections. Notre extension permet d'ajouter la communication entre les véhicules et l'infrastructure. Un protocole basé sur l'envoi de message permet à toutes les entités, modélisées comme des agents autonomes, de partager des informations et d'adapter leur comportement.

**Scénarios.** Afin de générer différents états distincts pour l'unité de bord de route, nous avons déterminé les flux d'entrée (section principale et voie d'insertion) à partir de données réelles issues de boucles électromagnétiques. Le jeu de données est composé d'un mois (juin 2001) de données boucles, récoltées entre 8h et 9h sur l'autoroute A6. Nous n'avons conservé que les données de deux boucles, choisies pour la correspondance entre le réseau réel et celui de notre scénario (figure 4). Cela a permis de produire 30 simulations de 60 minutes chacune. La diversité des situations de trafic simulées permet d'étudier le comportement du modèle dans des contextes sensiblement variables (trafic fluide, régime critique ou congestionné).

### 3.2 Définition des comportements

Afin de reproduire les dynamiques du flux de trafic, des lois de poursuite modélisent le mouvement longitudinal des véhicules. Ce comportement de base est modifié par les messages de consignes reçus par les véhicules connectés.

**Unité de Bord de Route.** Le résultat de la stratégie de contrôle de l'UBR est le choix, à chaque pas de temps de décision (fixé à 120 secondes), d'une action parmi les suivantes :

- $A_1$  : pas d'envoi de consigne.
- $A_2-A_3$  : consigne de changement de voie (droite à gauche – gauche à droite).
- $A_4-A_5$  : consigne de modification d'interdistance (1,8 s–1,2 s).
- $A_6-A_7$  : consigne de limite de vitesse (110 km/h–50 km/h).

Toutes les actions mènent à la propagation d'un message de consigne (sauf  $A_1$ ) contenant le type de consigne et le paramètre associé aux véhicules connectés. Ces messages définissent aussi une zone de pertinence (qui couvre l'ensemble de la section jusqu'à la voie d'insertion) ainsi qu'une date d'expiration (fixée elle-aussi à 120 secondes). La consigne contenue dans un message est appliquée par le véhicule qui l'a reçue tant qu'il se trouve dans la zone de pertinence du message et que celui n'a pas expiré.

**Véhicules.** Le comportement longitudinal des véhicules connectés et non-connectés est régi par une loi de poursuite. Nous avons choisi *Intelligent Driver Model* (IDM), qui est un modèle sans collision capable de reproduire les instabilités du flux de trafic [11]. La décision de changement de voie est modélisée grâce à la stratégie opportuniste nommée MOBIL [6]. Seuls les véhicules connectés sont équipés de dispositifs de communication sans fil leur permettant de percevoir les messages de consigne dans un rayon de 150 mètres, ce qui correspond aux technologies

actuelles. Nous ne considérons par les messages partiellement reçus ou les pertes d'informations dans la mesure où le protocole réel est censé les traiter. Les messages de consigne reçus par les véhicules connectés sont retransmis aux autres véhicules (à portée) jusqu'au début de la section (*i.e.* tant qu'ils se trouvent dans la zone de pertinence du message).

Les véhicules connectés affichent dans leur interface embarquée la consigne contenue dans les messages, afin que le conducteur puisse l'intégrer à son comportement. Nous proposons donc une interprétation par le véhicule de chacune des consignes. Pour les actions  $A_2$  et  $A_3$ , la consigne de changement de voie modifie la stratégie opportuniste tant que le message reste pertinent (zone et durée). Le véhicule va donc essayer de rejoindre la voie la plus à gauche (respectivement à droite) en entreprenant un ou des changement(s) de voie jusqu'à atteindre cette voie cible. Le modèle de changement de voie intègre une estimation du créneau spatio-temporel disponible. Cette estimation permet au véhicule de décider si un changement de voie est suffisamment sûr (au sens de ces paramètres) pour être effectivement exécuté. Cela implique qu'une plus grande proportion de véhicules connectés changera de voie lorsque le trafic sera fluide (et inversement lorsque le trafic sera dense). Pour les actions  $A_4$  à  $A_7$ , la seule modification consiste à changer temporairement les paramètres correspondants dans le modèle longitudinal (respectivement, l'interdistance désirée et la vitesse désirée) pour les valeurs indiquées dans le message de consigne. Toutes ces modifications sont maintenues tant que le message reste pertinent et sont automatiquement annulées dès que le message expire ou que le véhicule quitte la zone de pertinence définie.

### 3.3 Implémentation du modèle

Le modèle présenté dans cet article est implémenté en tant que processus de décision d'une unité de bord de route. Cette UBR est capable de percevoir par l'intermédiaire d'un capteur de type électromagnétique le débit moyen, la concentration et les vitesses sur la voie d'insertion et sur les trois voies de la section principale. En plus, les véhicules connectés situés dans son rayon de communication partagent leurs vitesses respectives. L'ensemble de ces données constitue la perception bas-niveau du système. La perception et la décision du système sont ici discrétisées et exécutées toutes les 120 secondes. L'UBR perçoit des données à partir de ses capteurs, agrégées sur la période précédente de 120 secondes, et choisit et applique immédiatement une action (*i.e.* envoie une consigne). Après une période de 120 secondes supplémentaires, la consigne expire et l'UBR reçoit, au même moment, un *feedback* utilisé pour renforcer la représentation et une nouvelle perception de l'environnement.

**Agents discrétiseurs.** Afin de montrer les bénéfices du modèle, nous proposons une implémentation avec deux agents discrétiseurs. Chaque agent a pour rôle de générer une représentation (discrétisation) à partir des variables

qu'il perçoit. Nous avons décidé de lier chaque discrétiseur, nommés  $D_1$  et  $D_2$ , à un capteur spécifique. Les agents diffèrent donc principalement par leur perception. Ainsi,  $D_1$  perçoit les données des boucles électromagnétiques (débit et densité moyens) pour chacune des trois voies et la voie d'insertion, agrégées toutes les 120 secondes.  $D_2$  exploite les données provenant du capteur de communication de l'UBR et construit un modèle (moyenne et écart-type) de la distribution estimée des vitesses à partir des messages des véhicules connectés qui sont passés dans son rayon de communication durant son dernier pas de temps de décision. Les discrétisations des agents sont générées à partir d'une classification par l'algorithme des K-moyennes (stabilisé avec 10000 itérations à partir d'initialisation aléatoires) pour respectivement 4 et 3 classes ( $D_1$  et  $D_2$ ). Les résultats de la discrétisation par les deux agents sur les données issues des 15 premières simulations sont représentés sur la figure 5. L'apprentissage des états perception-action de chaque agent est réalisé sur les 15 mêmes simulations, en exploitant le *feedback* du système.

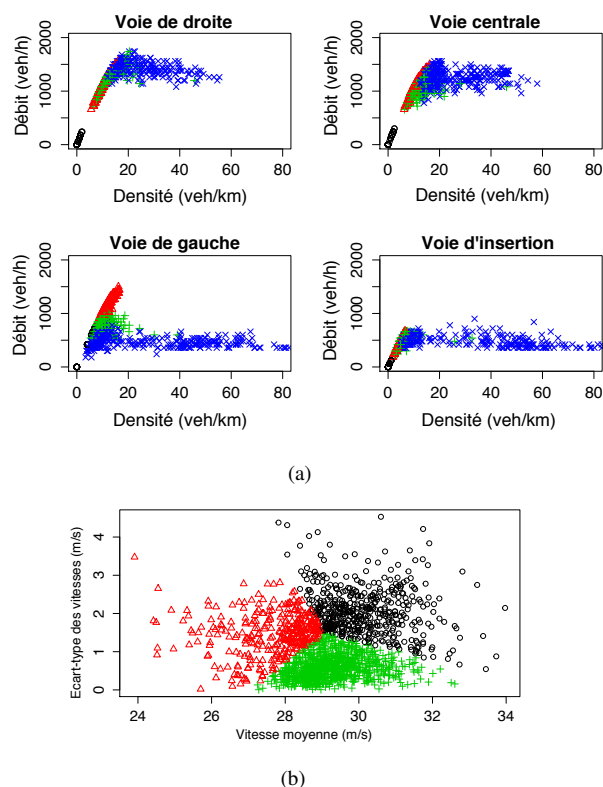


FIGURE 5 – Discretisations générées par (a)  $D_1$  [8 dimensions] et (b)  $D_2$  [2 dimensions] sur le jeu de données d'apprentissage.

**Renforcement de la représentation.** Les récompenses recueillies par les agents sous la forme de liens état-action sont calculées à partir d'une variable de l'environnement choisie. Dans le scénario proposé, nous avons retenu la vitesse moyenne la plus faible (sur les 4 boucles électroma-

gnétiques) en tant que *feedback* du système, puisque maximiser cette valeur correspond à l'objectif global d'amélioration de l'écoulement du flux. Ce retour sur expérience externe, est utilisé sous la forme de récompenses dans l'algorithme d'exploration-exploitation pour renforcer à la fois la stratégie de contrôle et les représentations proposées. Parmi les algorithmes de bandit à  $n$  bras de la littérature, notre choix s'est arrêté sur *Upper Confidence Bound* [1] (UCB). Son initialisation consiste à essayer chaque machine une seule fois. Ensuite, l'algorithme sélectionne systématiquement la machine  $j$  qui maximise  $\hat{x}_j + \sqrt{\frac{2 \ln n}{n_j}}$  où  $\hat{x}_j$  correspond à la récompense moyenne obtenue pour la machine  $j$ ,  $n_j$  est le nombre de fois que la machine  $j$  à été jouée jusqu'ici et  $n$  est le nombre total d'essais. La mise à jour des récompenses concerne tous les liens perception-action (notés  $L$  dans l'algorithme 1), et non pas seulement celui qui a été effectivement sélectionné par l'algorithme. De cette façon, le bénéfice d'une action est également propagé aux liens qui ont contribué à sa sélection. Cette récompense permet d'accélérer la phase d'exploration, en s'appliquant à autant de liens que d'agents discrétiseurs pour chaque itération.

### 3.4 Expérimentations

Les expérimentations sont menées en deux étapes principales. Dans un premier temps, un ensemble de données d'entrée (la perception des agents  $D$ ) est nécessaire afin d'appliquer les algorithmes de classification et obtenir des discrétisations stables (sans appliquer les actions). Puis, un jeu de simulations est utilisé pour l'apprentissage des liens perception-action en utilisant le processus complet d'exploration et d'application des actions (sur les 15 simulations). La seconde étape confronte le système à des scénarios alternatifs, pour vérifier que la représentation construite par le système l'aide à appliquer une stratégie de régulation efficace sur le flux. Cette étape est menée sur la deuxième partie des simulations. La situation témoin de chaque simulation est l'observation d'un ensemble d'indicateurs dans le même scénario mais sans utiliser l'infrastructure (aucune consigne n'est envoyée). Nous comparons trois implémentations du modèle : le cas avec un seul agent ( $D_1$  et  $D_2$  séparément), et une combinaison dynamique des représentations de ces deux agents grâce au modèle.

**Indicateurs.** Afin d'étudier les effets des différentes stratégies, nous avons sélectionné trois indicateurs : le temps total passé (TTS), la vitesse moyenne sur la section et le pourcentage de congestion. Le TTS est simplement la somme des temps de trajet de tous les véhicules sur la section ; des valeurs plus faibles indiquent un meilleur écoulement. La vitesse moyenne permet d'observer, au cours du temps, l'homogénéité du flux. Le pourcentage de congestion est la proportion spatiale de la section où les véhicules circulent à moins de 30 km/h. On peut s'attendre à ce que l'implémentation du modèle profite des représentations individuelles proposées par les agents et mène à l'amélioration

de l'écoulement du flux, observé au regard de ces indicateurs.

**Résultats.** Parmi les 15 simulations disponibles, nous en avons sélectionné 3 différentes qui présentent des situations de trafic différentes. La figure 6 présente le tracé des trois indicateurs sur ces simulations. Les indicateurs

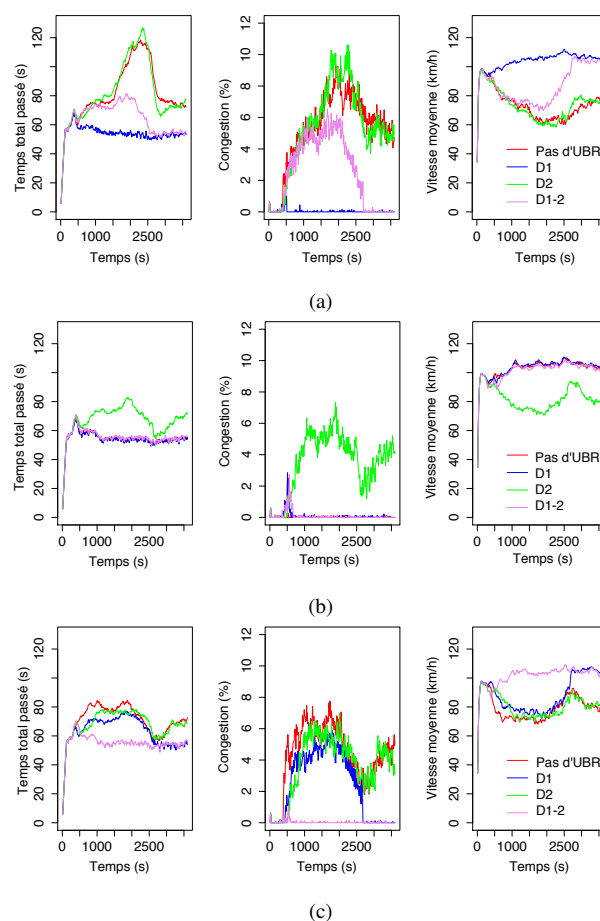


FIGURE 6 – Résultats des 3 simulations : (a) Lundi 18, (b) Jeudi 21 and (c) Mardi 28.

permettent d'évaluer les effets des différentes implémentations du modèle, en les comparant avec la simulation témoin (courbes rouges). Dans les simulations (a) et (c), sans régulation, les instabilités générées par les changements de voies depuis la voie d'insertion se propagent à l'ensemble du flux. Les trois simulations montrent des comportements différents du modèle. Dans la simulation (a), la stratégie de contrôle de l'agent  $D_1$  permet d'éviter l'apparition de congestion alors que celle de l'agent  $D_2$  donne de moins bon résultats que dans la simulation témoin. Ces résultats sont bien visibles en observant le tracé des TTS. Des valeurs faibles de temps total passé indiquent que les véhicules parcourent la section plus rapidement en moyenne. La combinaison dynamique de  $D_1$  et  $D_2$  produit une stratégie mixte qui donne une amélioration mitigée. La simulation (b) illustre un cas spécifique où l'un des deux discrétiseurs

seurs propose une mauvaise représentation ( $D_2$ ), puisque l'on n'observe pas de congestion dans la simulation témoin. Néanmoins, le modèle est capable de converger vers les meilleures représentations individuelles ( $D_1$ ) dans ce contexte, ce qui permet de ne pas empirer la situation observée dans la simulation témoin. Dans la figure 6 (c), l'état de trafic modélisé se rapproche du régime critique du flux, puisque l'ampleur de la congestion est limitée et la vitesse moyenne ne diminue que peu (comparé à la simulation (a)). Dans ce contexte, les deux stratégies apprises par chaque discrétiseur ont un impact positif sur le flux, et l'amélioration de la représentation par le modèle est visible. Le système est capable, grâce aux récompenses exploitant le *feedback*, de construire une représentation de plus haut niveau à partir de celles des agents, qui permet d'éviter la propagation de congestion sur la section. La figure 7 représente

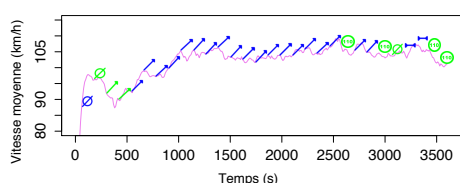


FIGURE 7 – Répartition entre les agents ( $D_1$  en bleu,  $D_2$  en vert) des consignes choisies par le modèle dans la simulation (c).

de façon simplifiée la contribution de chaque discrétiseur (en reprenant le code couleur de la figure 6) à la stratégie de régulation globale du système. Le schéma montre que des états perception-action des deux agents sont utilisés par le modèle. L'action  $A_2$ , la plus utilisée, semble logique puisqu'elle permet de laisser plus de place aux véhicules entrant pour s'insérer dans le flux principal. Ainsi, les résultats confirment la capacité du modèle à utiliser différentes représentations concurrentes et à les combiner en une représentation plus précise.

**Discussion.** Comme dans la plupart des problèmes d'apprentissage par renforcement, le choix du *feedback* est critique et complexe. De plus, comme le système exploite différentes discrétisations concurrentes, le choix des données d'entrée (perception) est important. Cela est dû à une problématique de fusion de capteurs : certains phénomènes sont perceptibles ou améliorables en utilisant seulement certaines dimensions. Cela signifie que le système doit percevoir des variables pertinentes pour son usage pour être capable d'adapter sa stratégie de contrôle aux variations de l'environnement. Dans les expérimentations, nous avons vu que le concepteur doit fournir un premier ensemble de discrétiseurs capables d'identifier des états pertinents, avant même de les faire combiner par le modèle. Ce problème pourrait être traité par l'ajout de mécanismes permettant au système de créer de manière autonome de nouveaux agents, voire de les faire évoluer dynamiquement, en exploitant son expérience d'interaction.

Comme les données d'entrée peuvent provenir de différentes dimensions (ou variables), une possibilité est de donner une perception complète au système. Dans ce cas, lors de l'implémentation du modèle, le temps d'exploration de l'ensemble des combinaisons de toutes les variables pour chaque discrétisation risque d'être très important. C'est pourquoi nous proposons d'utiliser différents agents qui ne travaillent que sur un sous-ensemble de la perception de bas-niveau. Les résultats obtenus en simulation laissent entrevoir le potentiel de convergence du système (en termes de profils de liens perception-action) vers les combinaisons de variables les plus pertinentes pour l'usage du système (*i.e.* au regard de la stratégie de contrôle produite). Cela pourra être particulièrement utile pour des problèmes où la combinaison des variables d'entrée n'est pas connue *a priori*.

## 4 Conclusions et perspectives

Nous avons proposé un modèle générique permettant à un système de construire une représentation de son environnement à partir de ses interactions. Par un apprentissage constructiviste visant à construire itérativement le processus de prise de décision d'un système de contrôle, le modèle utilise une population interne d'agents discrétiseurs pour amorcer la construction de la représentation. La représentation qui en est le fruit se compose d'états perception-action qui évoluent dans un apprentissage par renforcement exploitant un retour de l'environnement afin d'explorer l'espace de recherche. Après avoir présenté le modèle d'un point de vue théorique, nous avons proposé une application dans le cas d'étude de la régulation de trafic coopératif. L'objectif est d'utiliser l'infrastructure et les véhicules connectés pour proposer un système d'aide à la décision autonome. Ce cadre applicatif innovant met en valeur les effets de notre approche. Les résultats obtenus en simulation montrent que la combinaison dynamique des discrétisations individuelles permet au système d'adopter une stratégie de régulation plus efficace.

La prochaine étape de notre travail se concentrera sur les évolutions du modèle. La construction du modèle conceptuel est aussi itérative, afin de pouvoir tester et valider chacun des mécanismes par l'implémentation à chaque étape. Nous cherchons désormais à concevoir un nouveau type d'agent capable d'exploiter les hypothèses d'état proposées par les discrétiseurs en proposant des associations entre ces états. Dans un premier temps, deux types d'associations seront considérées : la spécialisation ou l'agrégation. La question du compromis entre le temps nécessaire à l'exploration et le nombre d'associations reste à étudier.

## Références

- [1] P. Auer, N. Cesa-Bianchi, and P. Fischer. Finite-time analysis of the multiarmed bandit problem. *Machine learning*, 47(2-3) :235–256, 2002.

- [2] A. L. Bazzan and F. Klügl. A review on agent-based technology for traffic and transportation. *The Knowledge Engineering Review*, 29(03) :375–403, 2014.
- [3] R. A. Brooks. Intelligence without representation. *Artificial intelligence*, 47(1) :139–159, 1991.
- [4] R. C. Carlson, I. Papamichail, and M. Papageorgiou. Local feedback-based mainstream traffic flow control on motorways using variable speed limits. *Intelligent Transportation Systems, IEEE Transactions on*, 12(4) :1261–1276, 2011.
- [5] M. Guériau, R. Billot, F. Armetta, S. Hassas, and N.-E. El Faouzi. Un simulateur multiagent de trafic coopératif. In *23es Journées Francophones sur les Systèmes Multi-Agents (JFSMA'15)*, pages 165–174. Cépaduès, 2015.
- [6] A. Kesting, M. Treiber, and D. Helbing. General lane-changing model mobil for car-following models. *Transportation Research Record : Journal of the Transportation Research Board*, 1999(1) :86–94, 2007.
- [7] J. Mugan and B. Kuipers. Autonomous representation learning in a developing agent. In *Computational and Robotic Models of the Hierarchical Organization of Behavior*, pages 63–80. Springer, 2013.
- [8] M. Papageorgiou, C. Diakaki, V. Dinopoulou, A. Kotsialos, and Y. Wang. Review of road traffic control strategies. *Proceedings of the IEEE*, 91(12) :2043–2067, 2003.
- [9] J. Piaget. The construction of reality in the child. *Journal of Consulting Psychology*, 19(1) :77, 1955.
- [10] J. Provost, B. J. Kuipers, and R. Miikkulainen. Self-organizing distinctive state abstraction using options. In *Proc. of the 7th International Conference on Epigenetic Robotics*, 2007.
- [11] M. Treiber and A. Kesting. *Traffic flow dynamics : data, models and simulation*. Springer, 2013.
- [12] J. Zlatev and C. Balkenius. Introduction : Why "epigenetic robotics"? In *Proceedings of the First Conference on Epigenetic Robotics.*, number 85, 2001.



# Can Synthetic Data Handle Unconstrained Gaze Estimation ?

A. Kacete<sup>1</sup>R. Ségurier<sup>2</sup>M. Collobert<sup>3</sup>J.Royan<sup>1</sup><sup>1</sup> Institute of Research and Technology b-com<sup>2</sup> Centralesupelec<sup>3</sup> Orange Labs

amine.kacete@b-com.com

## Résumé

Dans cet article, nous visons à résoudre le problème d'estimation du regard sans contrainte en utilisant une approche basée apparence. Contrairement aux méthodes précédentes qui fonctionnent dans un environnement relativement contraint, nous proposons une approche permettant un mouvement libre de la tête et des distances utilisateur-capteur significatives en utilisant un capteur RVB-P. Notre article présente les contributions suivantes : (i) une estimation directe du regard à partir des apparences RVB des deux yeux et de la profondeur du visage ; (ii) une stratégie de sélection de canal pendant l'apprentissage pour évaluer l'implication de chaque canal dans la prédiction finale ; (iii) adaptation d'un modèle 3D déformable du visage intégrant un modèle regard paramétrique pour générer un ensemble d'apprentissage RVB-P. Nous collectons également un ensemble d'échantillons annotés réels à partir de la Kinect permettant d'évaluer le potentiel de l'apprentissage synthétique face à des configurations réelles et d'établir une comparaison objective avec un apprentissage purement réel. Les résultats sur plusieurs utilisateurs démontrent le grand potentiel de notre approche.

## Mots Clef

Estimation du regard, Eye tracking, Champs aléatoires, Données synthétiques, Modèle déformable 3D.

## Abstract

In this article, we aim at solving unconstrained gaze estimation problem using appearance-based approach. Unlike previous methods working in relatively constrained environment, we propose an approach that allows free head motion and significant user-sensor distances using RGB-D sensor. Our paper presents the following contributions : (i) A direct estimation by inferring gaze information from RGB eyes and depth face appearances ; (ii) A channel selection strategy during the learning to evaluate the involvement of each channel in the final prediction ; (iii) Adapting a 3D face morphable model by integrating a parametric gaze model to render an important synthetic RGB-D training set. We also collect real labeled samples using Kinect

sensor that allows for evaluating the potential of synthetic learning in handling real configurations and establish an objective comparison with real learning. Results on several users demonstrate the great potential of our approach.

## Keywords

Gaze estimation, Eye tracking, Random Forest, Synthetic data, 3D morphable model.

## 1 Introduction

Gaze estimation plays a key role in several computer vision applications. In facial expression recognition fields, it allows access to important information such as the cognitive and expressive state of the person. In human behavior analysis, it allows the point of interest of the user, which represents the input of various devices such as user attention while driving and helping disabled people. Several industrial solutions are commercialized. They provide good accuracy on gaze estimation. Some of these solutions use complex hardware specifications (embedded camera on a head-mounted system) making them inappropriate for large scale public use. Other solutions use a range of infrared cameras to detect corneal reflection, but they remain very sensitive to illumination conditions. Recently, researches focus on using low cost devices such as monocular cameras, a comprehensive survey presented in [9], considers two main categories of gaze estimation, features-based methods and appearance-based methods.

### 1.1 Feature-based methods

These methods rely on the extraction of some features such as the pupil center, the eye corners, the iris contour or the corneal reflection, which are used to build a 3D eye model and determine the visual axis. [8] and [25] used the pupil center corneal reflection, from the IR lights which are used to illuminate the eye regions from different directions giving different image appearances, the corneal reflection is built by subtracting these images. [22] and [11] estimated the shape of the iris by fitting an ellipse to infer the gaze. [15] and [10] estimate the gaze direction from the 2D locations of the pupils and the corners in the

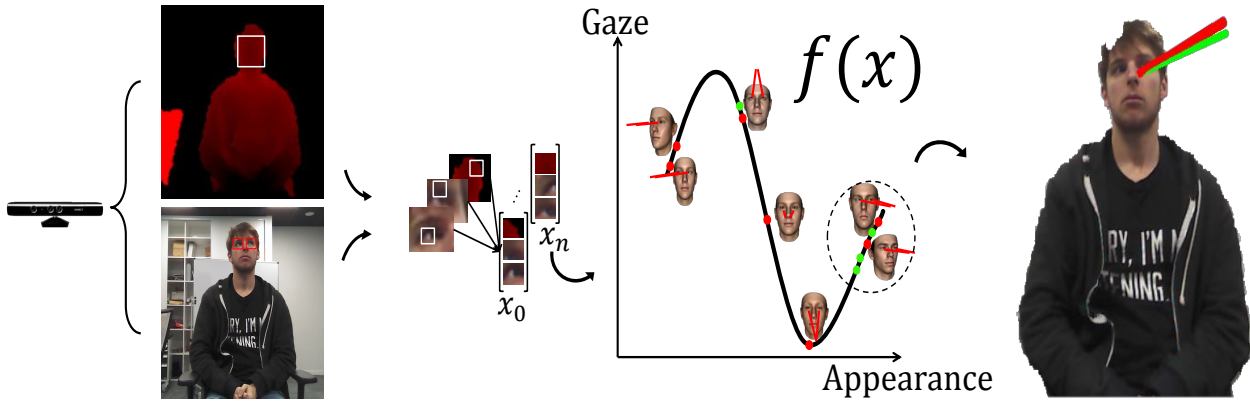


FIGURE 1 – Automatic gaze estimation based on our approach. We build a 3-channels global vector represented by the two RGB eye images and the face depth information using the depth sensor multimodal data, we extract a set of patches and project it through the forest represented here as the mapping function  $f(x)$  (the learned gaze sample clusters are defined as the red centroid points). Each single tree casts votes for each patch (defined as the green points). By performing a non-parametric clustering technique, a final estimation is calculated (represented as the green line, the red one defines the ground truth).

eye image. All the above methods simplify the anatomical structure of the eyeball and define the gaze direction as the optical axis. [4] proposed an extended 3D eye model based on the pupil and the corners to estimate the visual axis but still require a high image resolution to detect the corners accurately, in addition, they manually labeled the pupils centers. The main limitation of these methods lies in the direct link between their gaze estimation precision and the accuracy of the eye’s key-points localization (pupil, corners etc.) which requires a high image resolution and small head pose changes.

## 1.2 Appearance-based method

Appearance-based methods learn the mapping from the eye image appearances space to the gaze estimation space. Many algorithms have been proposed. [1] trained a neural network with 2k samples to learn the mapping function. [20] proposed a weighted linear interpolation to estimate an unknown gaze point from 252 sparse samples. [23] trained a semi-supervised Gaussian process on 80 samples relatively sparse. [26] proposed support vector regressors to achieve a high non-linear mapping. [19] proposed an incremental learning strategy using an on-line sample acquisition from a video stream updating the mapping function for a number of limited head pose configurations. [14] introduced the adaptive linear regression for the learning on a very sparse training set. The accuracy of these previous approaches is significantly affected in unconstrained environment with high head pose changes. Recently, some methods aim to manage such trouble by considering gaze estimation and head pose as two independent geometrical components, these approaches can be seen as semi appearance-based methods. [13] proposed to separate head pose component from the global gaze estimation system. By performing an initial estimation under frontal

configuration assumption and a geometric compensation with the head pose parameters, the final gaze estimation is inferred. Using the same paradigm, [16] projected the training gaze sample in frontal manifold using a frontalization step based on the head pose parameters calculated using a specific 3D model fitting. These last two methods solved the problem of head changes successfully but still working under low user-camera distances. To cover all the eye image appearance variability [24] recorded around 200k training samples and used a deeper strategy using a convolutional neuronal network to learn a very robust mapping function achieving a high gaze estimation accuracy but very constrained by an important computational time making this method not real-time.

In this paper, we consider the high non-linear problem of gaze estimation under head pose changes and large user-sensor distances as a regression task. To learn such mapping robustly, we propose a novel approach that considers a global gaze manifold instead of learning in frontal configurations and geometrically correct the final estimation using head pose parameters as usually done. We train an ensemble of regression trees able to capture robustly gaze information on an important 3-channels training samples ( $\text{channel}^{(0)}$ ,  $\text{channel}^{(1)}$  defines the gray scale images relative to right and left eye respectively,  $\text{channel}^{(2)}$  defines the depth image of the face) organized as a set of patches (where a patch defines a small group of nearby pixels). We apply a channel-selection during the training to evaluate the importance and involvement of each channel in the final estimation. We define the gaze vector  $g$  as the vector stretching the gravity center of the face and the gazed 3D point. To provide a significant set of training data for learning the trees, we render a very important amount of gaze samples using a 3D statistical morphable model with integrating dynamic gaze model. We also build an important



gaze database recorded with the Microsoft Kinect sensor. Rendered synthetic data are exclusively used for the learning and real data are used for both learning and testing.

Fig. 1 describes an overview of our automatic gaze estimation system. The rest of the paper is organized as follows : Sec. 2 describes our method in details. In Sec 3, we detail how training data are generated. Sec. 4 describes the experiments and evaluates the precision of our approach. Sec. 5 concludes the paper.

## 2 Our method

We use randomized regression trees to estimate the two angles  $(\theta, \gamma)$ , which represent the horizontal and vertical orientation of the gaze vector  $g$ , from the RGB and depth cues combined on 3-channels patches. In Sec. 2.1, we provide some background on regression trees. In Sec. 2.2 and Sec. 2.3 we detail the training and testing steps of our forests respectively. Sec.3 describes the generation of our training RGB-D gaze samples.

For the next sections, we define also head pose parameters as  $\mathcal{H}$  (with  $\mathcal{H} : [R|T]$ ).

### 2.1 Random regression forest

Recently, many applications in computer vision have used Random Forest to achieve the mapping from complex input spaces into discrete or continuous output spaces. Introduced by [2], randomized trees deal with different tasks such as classification in [7, 12] and regression in [18, 5].

Regression forest is an ensemble of trees predictors which splits the initial problem in two low complex problems in a recursive way. At each node, a simple binary test is performed, according to the result of the test, a data sample is directed towards the left or the right child. The tests are selected to achieve an optimal clustering. The terminal nodes of the tree called leaves, store the estimation models approximating the best the desired output. To achieve high generalization, the trees are trained in a decorrelated way (with introducing randomness in both the training data provided for each tree and the set of binary tests).

### 2.2 Training

We supervised the training of each tree  $T$  in the forest  $\mathcal{T} = \{T_i\}$  using a set of annotated patches  $\{\mathcal{P}_i = (\mathcal{I}_i^c, g_i)\}$  randomly selected from the training data where :

- $\mathcal{I}_i^c$  represents the extracted visual features vector from a given patch  $\mathcal{P}_i$ ,  $c$  defines the feature channel. We used 3 channels namely the two grayscale intensities extracted from the two eye images and the depth values extracted from the face.
- $g_i$  represents the output gaze vector represented with two components  $(\theta, \gamma)$ .

Starting from the root, at each non-leaf node, we define a simple binary test  $t_{x_1, y_1, x_2, y_2, c, \tau}$  :

$$\begin{cases} 1, & \text{if } \mathcal{I}_i^c(x_1, y_1) - \mathcal{I}_i^c(x_2, y_2) \leq \tau \\ 0, & \text{otherwise} \end{cases}$$

where  $(\mathcal{I}_i^c(x_1, y_1) - \mathcal{I}_i^c(x_2, y_2))$  represents the difference of intensity between two locations  $(x_1, y_1)$  and  $(x_2, y_2)$  in the channel  $c$ . Supervising the training consists in finding at each non-leaf node the optimal binary test  $t^*$  that maximizes the purity of the data clustering. Maximizing the clustering purity is achieved by maximizing the information gain defined as the differential entropy  $H$  of the set of patches at parent node  $\mathcal{P}$  minus the weighted sum of the differential entropies computed at the children  $\mathcal{P}_{\mathcal{L}}$  and  $\mathcal{P}_{\mathcal{R}}$  :

$$E = H(\mathcal{P}) - (\omega_{\mathcal{L}}H(\mathcal{P}_{\mathcal{L}}) + \omega_{\mathcal{R}}H(\mathcal{P}_{\mathcal{R}})) \quad (1)$$

The weights  $\omega_{j \in \{\mathcal{R}, \mathcal{L}\}}$  are defined as the ratio of patches reached to the parent and the right or left child respectively, *i.e.*,  $\frac{|\mathcal{P}_{j \in \{\mathcal{R}, \mathcal{L}\}}|}{|\mathcal{P}|}$ . Assuming that the gaze vector  $g$  at each node is a random variable with a multivariate Gaussian distribution such as  $p(g) = \mathcal{N}(g, \bar{g}, \Sigma)$ , allows us to rewrite Eq. 1 as follows :

$$E = \log |\Sigma(\mathcal{P})| - (\omega_{\mathcal{L}} \log |\Sigma(\mathcal{P}_{\mathcal{L}})| + \omega_{\mathcal{R}} \log |\Sigma(\mathcal{P}_{\mathcal{R}})|) \quad (2)$$

The learning process finishes when the data reach a predefined maximum value of the tree or when the number of patches let down a threshold value yielding the creation of the leaves. Each leaf  $l$  stores the mean of all the gaze vectors which reached it with the corresponding covariance.

### 2.3 Testing

Given an unseen instance, we extract a set of patches from the RGB eye regions and the face depth information after a face detection step. Each patch is passed through all the learned trees in the forest. Using the optimal stored binary test, each tree processes the patch until reaching a leaf. The gaze vector estimation according to a single tree is given by the reached leaf  $l$  in terms of the stored distribution  $p(g|l) = \mathcal{N}(g, \bar{g}, \Sigma)$ . The gaze vector estimation for a given patch  $\mathcal{P}_i$  over all the trees is calculated as follows :

$$p(g|\mathcal{P}_i) = \frac{1}{|\mathcal{T}|} \sum_t p(g|l_t(\mathcal{P}_i)) \quad (3)$$

All the estimations corresponding to the extracted patches are grouped in a set of votes. Before performing the clustering of these votes, we discard the estimations from the leaves with high variance considered as non-informative. To locate the centroid of the cluster of the votes, we perform 5 mean-shift iterations using a Gaussian kernel.

## 3 Data generation

To provide a representative training dataset, we use two types of data : synthetic and real data.

### 3.1 Synthetic data

In Computer vision community, machine learning techniques are considered as a very elegant way to tackle problems. They demonstrated a great potential in terms of

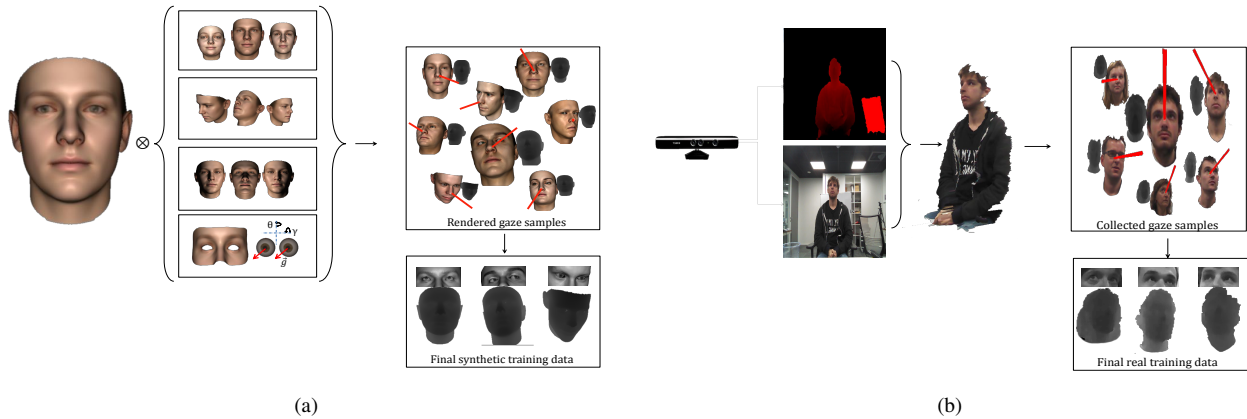


FIGURE 2 – Data generation. (a) represents synthetic data rendering using the 3D morphable model of [17]. By introducing some variabilities such as identity, head pose changes and lighting conditions, we integrated a dynamic gaze model (represented by two global textured spheres), we rendered the final RGB-D gaze training samples with the correspondent gaze annotation illustrated in red line. (b) We performed the same strategy using real data grabbed from the multimodal Kinect sensor by introducing the same previous variabilities. To obtain gaze annotation, a 2D moving point is gazed by the user (knowing the rigid transformation sensor-screen, the stretching vector from user head gravity and the projection of the moving point in the word coordinate can be calculated). These real data are principally used to evaluate accurately the performance of the synthetic data in handling gaze estimation.

efficiency and robustness. Nevertheless to achieve a high generalization across unseen scenarios, these methods often require a very representative training data set. Given that the building of high amount of labeled data is a very tedious process, synthetic data represent a promising solution. Indeed, the annotation is performed automatically instead of manual labeling. [3] developed an iterative model based on Gabor-filters applied on an empty image containing some seed points to render a fingerprint training samples. [27] rendered iris image samples obtained from a 2D polar projection of a cylindrical representation of continuous fibers. [21] improved face authentication by generating multiple virtual images using simple geometric transformations. [18] used a motion capture strategy to record RGB and depth cues of the body part movements, by varying body size and shape, scene position, camera position and mirroring the recorded data, they synthesize a highly varied training allowing a robust body part pose estimation. [6] tackled the head pose estimation problem with synthetic depth images by rendering an enormous amount of training data using a 3D statistical morphable model. In our method, we first generate our synthetic training gaze samples by rendering the 3D morphable model proposed by [17]. This model is built from around 200 scans of human faces. It contains a very high mesh density including the face, frontal neck and the ears. The shape and texture of the model is composed as a linear combination of 199 components. They can be deformed according to the following equation :

$$\mathcal{A} = \mathcal{A}_0 + \mathcal{M}_A \alpha \tag{4}$$

where  $\mathcal{A}$  can denote the generated texture or shape respec-

tively.  $\mathcal{A}_0$  denotes the mean,  $\mathcal{M}$  represents the basis components perturbed with parameters  $\alpha$ .

Fig.2a shows an overall of the generation process. To introduce face identity variation, we perturb the first 50 basis components of the shape and texture by  $\pm 1.5$  of the standard deviation of each mode. To render images in different head pose configurations, we apply random rigid transformations on the model : the rotations spans  $\pm 60$  for yaw and  $\pm 40$  for pitch and we translate the model along the  $z$  axis within 200 cm range for scale variability. Furthermore, to produce illumination variability, we generate light sources with different intensities and directions. Unfortunately, the basis components related to the shape and the texture of this model do not monitor the gaze direction.

So, to integrate a dynamic gaze system to the model able to generate different gaze direction instances, we decided to delete all the vertices related to the eye regions. Two spheres are placed as the eyeballs instead. We fix the diameters to human average eyeball namely 25 mm. We use different textures for the eyeballs to handle iris appearance variability. Moreover to control eyelids movements resulting from the gazing up and down, we introduce a linear translation for each vertex surrounding the eye regions as blendshapes. By defining the starting and the ending position in the global mesh, all the coefficients of the linear translations can be calculated. Thanks to the topology of the model, all these modifications keep the same behavior under identity variation. To generate gaze samples, we apply random rotations to the eyeballs, the gaze information angles can be easily computed knowing the location of the eyeballs.

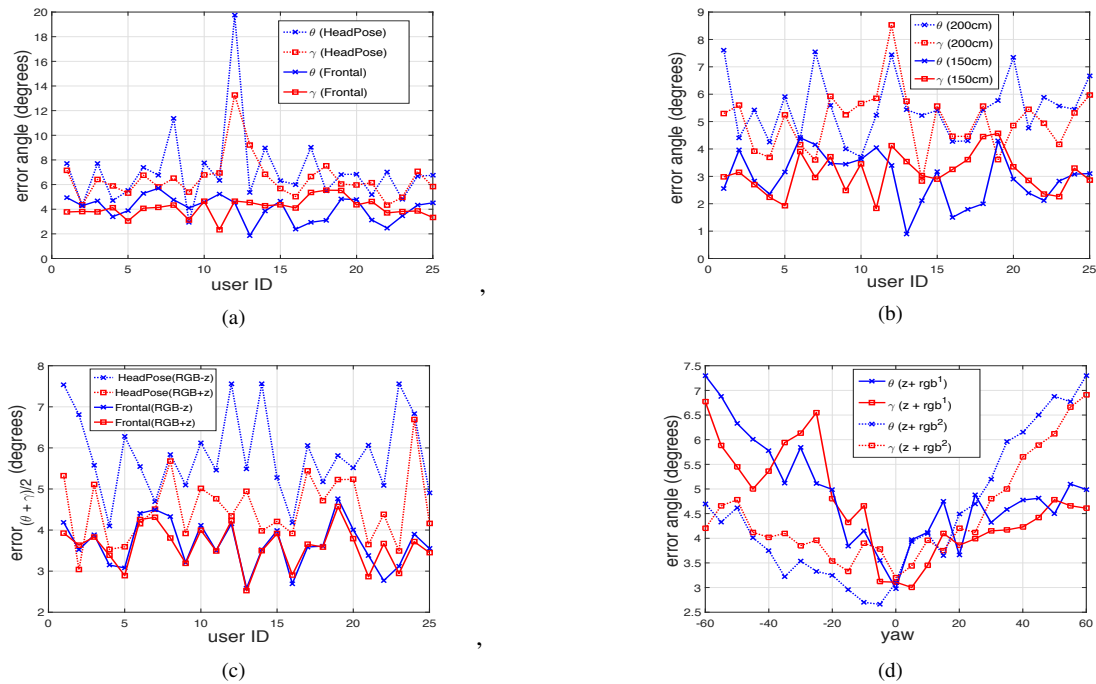


FIGURE 3 – (a) the mean error for the two gaze directions under frontal and head pose changes. (b) the mean error for the two gaze directions under two distances from the sensor. (c) mean error over the two directions with and without using channel<sup>(2)</sup> in frontal and head pose configurations respectively. (d) mean error of the two direction over head pose variation (yaw angle variation) with different channels combination

### 3.2 Real data

In the other hand, we recorded real gaze sample data using Microsoft Kinect sensor. The database contains 17k RGB-D images of 42 people (15 females and 27 males, 4 with glasses and 38 without glasses) gazing different targets displayed on a screen. The subject performed 4 scenarios, gazing with a fixed head at roughly  $d_0 = 150$  cm from the sensor, gazing at same distance  $d_0$  under head pose changes and the two others scenarios are performed at about  $d_1 = 200$  cm from the sensor. Knowing the Kinect intrinsic parameters and its rigid transformation to the screen, the displayed gaze points can be projected to the Kinect world space. The gaze vector is represented as the vector stretching the head gravity center (computed using face detection area) and the 3D gazed point. The acquisition was under SXGA and VGA resolution for RGB and depth respectively recorded at 15 fps. Fig.2b describes the acquisition process. First we grab RGB and depth information, by using the known calibration between the two sensors, a 3D textured mesh can be reconstructed. We show then, in analogy with synthetic data, some real training data used for both learning and testing.

## 4 Experimental results

In our experiments, we trained different forests either on real data or synthetic data. The nature of the experiment determines the training parameters.

We trained our forest  $\mathcal{T}_{g_s}$  using 400k RGB-D synthetic gaze samples under several head pose changes. We extracted 15 patches from each sample giving 6M training data. Face depth image size is fixed to  $(150 \times 150)$ , eye RGB images to  $(80 \times 70)$  and the size of each channel of the extracted patches for each channel is fixed to  $(16 \times 16)$ . Some training and testing parameters are fixed according to some empirical observations, *e.g.*, the maximum depth to 18 and at each node we randomly generate 400 splitting candidates with 50 thresholds giving a total number of 20k binary tests. At testing, we extracted a total of 30 patches from each gaze test sample with 20 regression trees. We tested our forest on 25 users from the real images database discussed previously.

### 4.1 Robustness to head pose and distance variations

We evaluate the gaze estimation accuracy using our trained forest  $\mathcal{T}_{g_s}$  under unconstrained environment. Fig. 3a represents the global error of gaze estimation over 25 users under frontal and head changes configurations. For each user, a mean error across different gaze samples performed under two distances is computed. In frontal case, the mean error over all the users is less than  $3^\circ$  for the two directions whereas the error is less than  $6.5^\circ$  for head pose changes case. This difference in accuracy between the two configurations is directly linked to the high eye image appearances

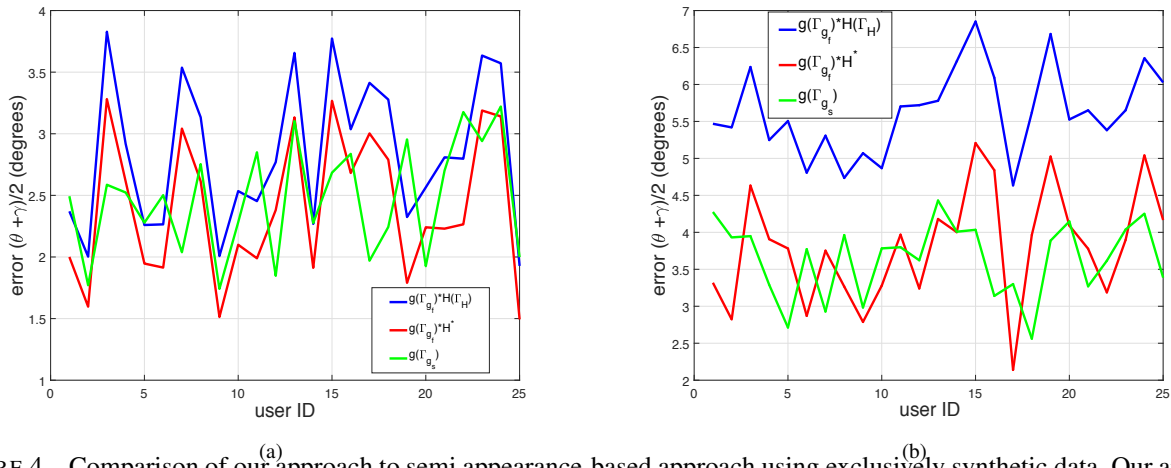


FIGURE 4 – Comparison of our approach to semi appearance-based approach using exclusively synthetic data. Our approach (in red) performs gaze estimation using RGB-D cues assuming a global gaze manifold (under head pose changes) in red. Semi appearance-based approach (in green) performs gaze estimation assuming frontal configuration with a geometrical correction using estimated head pose parameters. In blue : a semi appearance-based approach with ideal head pose parameters. (a) and (b) : mean gaze error across the two directions for frontal and head pose configuration respectively.

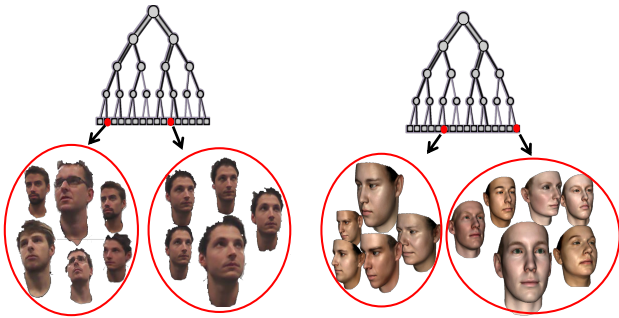


FIGURE 5 – Visualizing some clusters captured during the training step using only channel<sup>(2)</sup> with real and synthetic data respectively.

variability across head pose configuration making the trees prediction less accurate. In Fig. 3b we report the error as a function of distance from the sensor for a frontal configuration. The experiments show a mean error of  $2.9^\circ$  and  $3.1^\circ$  for  $\theta$  and  $\gamma$  respectively at 150 cm from the sensor. At 200cm, we notified a slightly higher errors,  $4.8^\circ$  and  $5.0^\circ$  for the two directions respectively. The difference in accuracy between the two distances is related to the RGB eye images and face depth appearances which are significantly variable depending on the distance from the sensor.

## 4.2 Channel selection importance

To evaluate the involvement of each channel (from the two eye RGB images and face depth information) at testing time, we realized experiments using the forest  $\mathcal{T}_{g_s}$  with and without depth channel and compared gaze estimation accuracy in both cases. Fig.3c illustrates the importance of depth information in our approach especially in head pose changes scenario. Gaze estimation errors are very

close with and without depth information in frontal scenario whereas the error gap is approximately  $1.5^\circ$  in head pose changes configuration proving the importance of this channel in such case. Depth information is more suitable to encode geometric similarities between data samples which represent the head pose information.

Fig.3d describes the influence of the two RGB channels (corresponding to right and left eye) on gaze estimation accuracy across different yaw angle values. These results are expected since eye appearance is very sensitive to head pose changes especially for *yaw* angle variation. For instance, positive values of yaw deform the left eye appearance until a complete disappearance giving high estimation errors for the two directions without the visible channel namely right eye (*i. e.*, dotted lines in Fig 3d) and reciprocally. Using our channel selection strategy introduced on the forest learning, we can quantify the involvement of each RGB channel in the final gaze estimation across head pose changes.

Fig.5 shows some clusters with low variances captured by the forest  $\mathcal{T}_{g_s}$  during training step. The process is achieved with real training data.

## 4.3 Semi appearance-based versus appearance-based approach

Fig.4 illustrates the robustness of learning gaze in a global manifold (under head pose estimation). Instead of separating gaze and head pose as usually done, we trained two supplementary forests  $\{\mathcal{T}_{g_f}, \mathcal{T}_H\}$  on exclusively synthetic data as follows :

- $\mathcal{T}_{g_f}$  : is the learned gaze estimation forest using only RGB (eye images) cues under frontal configuration exclusively. The forest is trained with the same parameters as  $\mathcal{T}_{g_s}$ .

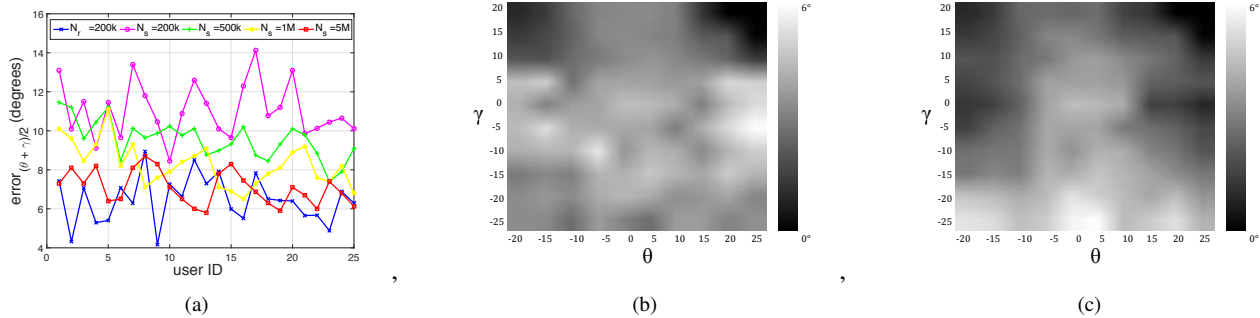


FIGURE 6 – (a) gaze estimation error (over  $\theta$  and  $\gamma$  and over 150cm and 200cm) under learning with real and synthetic data ( $N_r$  and  $N_s$  are the number of real and synthetic training data used respectively). (b) and (c) Gaze estimation error distribution based real and synthetic learning respectively.

—  $\mathcal{T}_{\mathcal{H}}$  : is the learned head pose estimation forest using RGB-D cues (face depth and face RGB images). The training parameters are fixed as done in [6] using  $100k$  training data.

Fig.4a illustrates the mean error of the gaze estimation across the two direction under frontal scenario using different approaches. In red, our approach using the forest ( $\mathcal{T}_{g_s}$ ), in green, frontal gaze estimation corrected with head pose parameters using  $\{(\mathcal{T}_{g_f}, \mathcal{T}_{\mathcal{H}})\}$  and in blue, frontal gaze estimation using  $\mathcal{T}_{g_f}$  corrected with an ideal head pose (driven from the OpenGL camera calibration). We noticed that errors are very close which is an expected result due to the fact that head pose parameters are not involved in the frontal scenario. Fig.4b describes the mean error in a head pose scenario. Our gaze estimation approach presents a lowest error compared to the frontal gaze estimation corrected with head pose parameters even if it is optimal. Correcting gaze estimation with head pose in a geometrical way makes the errors related to each component accumulated unlike our approach which performs a direct mapping producing an unique error for gaze estimation.

#### 4.4 Learning with real data versus learning with synthetic data

To evaluate the realism of our rendered synthetic data and their ability to handle unconstrained gaze estimation problem, we trained a forest  $\mathcal{T}_{g_r}$  on  $500k$  exclusively real training data under head pose changes extracted from the previous database (the other training parameters are kept fix as in  $\mathcal{T}_{g_s}$ ). To achieve a comparative analysis, we trained in the same way as  $\mathcal{T}_{g_s}$  different forest using different number of synthetic training data. The fact that learning forest is a very computational time task, to learn different forest in an acceptable time, we reduce the number of binary tests generated at internal nodes to  $1k$  instead of  $20k$  which affects considerably the estimation accuracy but yields sufficiently good results to compare different scenarios. Fig. 6a describes gaze estimation errors across different user under real and synthetic learning. A first observation can be driven by the figure, learning with the same number of train-

ing data in real and synthetic case does not perform the same accuracy which can be explained by the difference in the realism between real and synthetic data. During test, extracted patches appearance from a testing user is more closer to the gaze clusters appearances encoded using the forest  $\mathcal{T}_{g_r}$ . Increasing the number of synthetic training data make  $\mathcal{T}_{g_s}$  increasingly close to  $\mathcal{T}_{g_r}$ . This can be explained by enhancing the generalization ability across unseen scenarios with more training data (We evaluated the factor between real and synthetic training producing approximately same accuracy to  $1/9$ ).

In Fig. 6b and Fig. 6c we illustrate gaze estimation error distribution over all 5 best testing users using  $\mathcal{T}_{g_r}$  and  $\mathcal{T}_{g_s}$  respectively. We can notice the importance of the error under synthetic learning in Fig. 6c for  $\gamma$  less than  $-20^\circ$  values resulting from the eyes closure when gazing down.

## 5 conclusion

In this paper, we presented a robust approach to handle gaze estimation problem in unconstrained environment using an ensemble of regression trees grouped in a single forest with high ability of generalization. To ensure the robustness, we include both RGB and depth cues as input during learning assuming a global gaze samples manifold under head pose variation. To enhance the generalization, we render an important amount of training data using a 3D morphable model with an integrated dynamic gaze model. We also, build a database with real images to evaluate the accuracy of the gaze estimation in real scenario with accuracy. Different experiments scenarios demonstrate that our approach present a great potential regarding state-of-the-art methods.

## Références

[1] S. Baluja and D. Pomerleau. Non-intrusive gaze tracking using artificial neural networks. Technical report, DTIC Document, 1994.  
 [2] L. Breiman. Random forests. *Machine learning*, 45(1):5–32, 2001.

- [3] R. Cappelli, A. Erol, D. Maio, and D. Maltoni. Synthetic fingerprint-image generation. In *Pattern Recognition, 2000. Proceedings. 15th International Conference on*, volume 3, pages 471–474. IEEE, 2000.
- [4] J. Chen and Q. Ji. 3d gaze estimation with a single camera without ir illumination. In *Pattern Recognition, 2008. ICPR 2008. 19th International Conference on*, pages 1–4. IEEE, 2008.
- [5] A. Criminisi, J. Shotton, D. Robertson, and E. Konukoglu. Regression forests for efficient anatomy detection and localization in ct studies. In *Medical Computer Vision Workshop*. 2010.
- [6] G. Fanelli, J. Gall, and L. Van Gool. Real time head pose estimation with random regression forests. In *CVPR*, 2011.
- [7] J. Gall, A. Yao, N. Razavi, L. Van Gool, and V. Lempitsky. Hough forests for object detection, tracking, and action recognition. *TPAMI*, 2011.
- [8] E. D. Guestrin and M. Eizenman. General theory of remote gaze estimation using the pupil center and corneal reflections. *Biomedical Engineering, IEEE Transactions on*, 53(6) :1124–1133, 2006.
- [9] D. W. Hansen and Q. Ji. In the eye of the beholder : A survey of models for eyes and gaze. *TPAMI*, 2010.
- [10] T. Ishikawa, S. Baker, I. Matthews, and T. Kanade. Passive driver gaze tracking with active appearance models. In *Proceedings of the 11th World Congress on Intelligent Transportation Systems*, October 2004.
- [11] S. Kohlbecher, S. Bardinst, K. Bartl, E. Schneider, T. Poitschke, and M. Ablassmeier. Calibration-free eye tracking by reconstruction of the pupil ellipse in 3d space. In *Proceedings of the 2008 symposium on Eye tracking research applications*, pages 135–138. ACM, 2008.
- [12] V. Lepetit, P. Lagger, and P. Fua. Randomized trees for real-time keypoint recognition. In *CVPR*, 2005.
- [13] F. Lu, T. Okabe, Y. Sugano, and Y. Sato. A head pose-free approach for appearance-based gaze estimation. In *BMVC*, pages 1–11, 2011.
- [14] F. Lu, Y. Sugano, T. Okabe, and Y. Sato. Inferring human gaze from appearance via adaptive linear regression. In *Computer Vision (ICCV), 2011 IEEE International Conference on*, pages 153–160. IEEE, 2011.
- [15] Y. Matsumoto and A. Zelinsky. An algorithm for real-time stereo vision implementation of head pose and gaze direction measurement. In *Automatic Face and Gesture Recognition, 2000. Proceedings. Fourth IEEE International Conference on*, pages 499–504. IEEE, 2000.
- [16] K. A. F. Mora and J.-M. Odobez. Gaze estimation from multimodal kinect data. In *Computer Vision and Pattern Recognition Workshops (CVPRW), 2012 IEEE Computer Society Conference on*, pages 25–30. IEEE, 2012.
- [17] P. Paysan, R. Knothe, B. Amberg, S. Romdhani, and T. Vetter. A 3d face model for pose and illumination invariant face recognition. In *Advanced Video and Signal Based Surveillance*, 2009.
- [18] J. Shotton, R. Girshick, A. Fitzgibbon, T. Sharp, M. Cook, M. Finocchio, R. Moore, P. Kohli, A. Criminisi, A. Kipman, et al. Efficient human pose estimation from single depth images. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, 35(12) :2821–2840, 2013.
- [19] Y. Sugano, Y. Matsushita, Y. Sato, and H. Koike. An incremental learning method for unconstrained gaze estimation. In *Computer Vision–ECCV 2008*, pages 656–667. Springer, 2008.
- [20] K.-H. Tan, D. J. Kriegman, and N. Ahuja. Appearance-based eye gaze estimation. In *Applications of Computer Vision, 2002.(WACV 2002). Proceedings. Sixth IEEE Workshop on*, pages 191–195. IEEE, 2002.
- [21] N. P. H. Thian, S. Marcel, and S. Bengio. Improving face authentication using virtual samples. In *Acoustics, Speech, and Signal Processing, 2003. Proceedings.(ICASSP'03). 2003 IEEE International Conference on*, volume 3, pages III–233. IEEE, 2003.
- [22] J.-G. Wang and E. Sung. Study on eye gaze estimation. *Systems, Man, and Cybernetics, Part B : Cybernetics, IEEE Transactions on*, 32(3) :332–350, 2002.
- [23] O. Williams, A. Blake, and R. Cipolla. Sparse and semi-supervised visual mapping with the  $s^3$ gp. In *Computer Vision and Pattern Recognition, 2006 IEEE Computer Society Conference on*, volume 1, pages 230–237. IEEE, 2006.
- [24] X. Zhang, Y. Sugano, M. Fritz, and A. Bulling. Appearance-based gaze estimation in the wild. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pages 4511–4520, 2015.
- [25] Z. Zhu and Q. Ji. Novel eye gaze tracking techniques under natural head movement. *Biomedical Engineering, IEEE Transactions on*, 54(12) :2246–2260, 2007.
- [26] Z. Zhu, Q. Ji, and K. P. Bennett. Nonlinear eye gaze mapping function estimation via support vector regression. In *Pattern Recognition, 2006. ICPR 2006. 18th International Conference on*, volume 1, pages 1132–1135. IEEE, 2006.
- [27] J. Zuo, N. A. Schmid, and X. Chen. On generation and analysis of synthetic iris images. *Information Forensics and Security, IEEE Transactions on*, 2(1) :77–90, 2007.

# Représentations et modèles en extraction d'événements supervisée

D. Kodelja

R. Besançon

O. Ferret

CEA, LIST, Laboratoire Vision et Ingénierie des Contenus,  
Gif-sur-Yvette, F-91191, France.

{dorian.kodelja,romaric.besancon,olivier.ferret}@cea.fr

## Résumé

*Cet article de synthèse retrace l'histoire des techniques d'extraction d'événement à partir de textes. Cette tâche a d'abord été traitée par des règles lexico-syntaxiques puis des classifieurs supervisés utilisant tous deux des représentations complexes, fortement dépendantes du domaine et sujettes à la propagation d'erreurs. Récemment, plusieurs approches neuronales ont amélioré l'état de l'art en mettant en avant la réduction des prétraitements nécessaires et donc des possibilités d'erreur. Finalement, nous nuancions ce postulat en présentant de nouvelles méthodes réintroduisant ces informations linguistiques.*

## Mots Clef

Extraction d'information, apprentissage supervisé, réseaux de neurones, représentations distribuées.

## Abstract

*We present in this survey the successive approaches to supervised event extraction from texts. The first rule-based systems and the classical statistical methods use complex and domain-dependent representations that are prone to error propagation. In response to these problems, recent neural network systems using embeddings have linked their success to the absence of the preprocessing steps producing these errors. We nuance this viewpoint by presenting recent methods reintroducing these linguistic features.*

## Keywords

Information Extraction, supervised learning, neural networks, distributed representations.

## 1 Introduction

L'extraction d'information est un champ de recherche dont l'objectif consiste à extraire automatiquement des informations factuelles structurées dans un domaine donné à partir de données textuelles peu ou pas structurées. Les premiers systèmes d'extraction d'informations, développés manuellement pour un besoin précis dans un domaine spécifique, n'étaient cependant absolument pas réutilisables. Que ces

Ce travail a été partiellement financé par l'ANR dans le cadre du projet ASRAEL (ANR-15-CE23-0018).

documents émanent du domaine biomédical (articles), industriel (rapports trimestriels), de la Presse ou du Gouvernement, la tendance actuelle est à une explosion tant du volume d'information disponible que de sa variété. C'est pourquoi les approches successives de l'extraction d'information ont créé des systèmes de plus en plus modulaires et universels. Par ailleurs, en fonction de la disponibilité de documents annotés, l'extraction peut être supervisée ou non supervisée. À mi-chemin entre ces deux familles d'approches, des approches semi-supervisées peuvent utiliser un nombre réduit d'exemples annotés, un modèle incomplet ou des exemples privés de contextes pour amorcer un système supervisé. Nous nous placerons ici dans le cadre de l'extraction supervisée, présentée à la section 2. Différentes campagnes d'évaluation ont fortement guidé le développement de cette tâche et font l'objet de la section 3. L'extraction supervisée d'informations textuelles faisant appel à des ressources linguistiques et des connaissances fortement dépendantes du domaine, nous analyserons à la section 4 les différentes approches possibles en nous concentrant sur la transition vers les approches neuronales et les représentations distribuées et leur utilité pour le développement de systèmes génériques, donc plus robustes et facilement adaptables à de nouveaux domaines.

## 2 Présentation des tâches

Ainsi que l'illustre la Figure 1, l'extraction d'information supervisée au sein de sources bruitées et non structurées est une tâche complexe, décomposable en plusieurs sous-tâches généralement traitées séquentiellement. L'extraction d'événements est envisagée comme une tâche de remplissage de formulaire : le type de formulaire correspond à un type d'événement et impose le remplissage d'un nombre variable de champs identifiant les rôles associés à ce type d'événement. Le système extrait alors deux types d'informations : des mentions et des liens entre ces mentions, notamment leur rôle. Cette tâche est donc relativement proche de l'extraction de relations binaires, non présentée ici.

### 2.1 Extraction de mentions

**Reconnaissance d'entités nommées** La première étape d'un système d'extraction d'information consiste à identifier dans le texte l'ensemble des entités pouvant remplir un rôle vis-à-vis d'un événement. Une même entité pou-

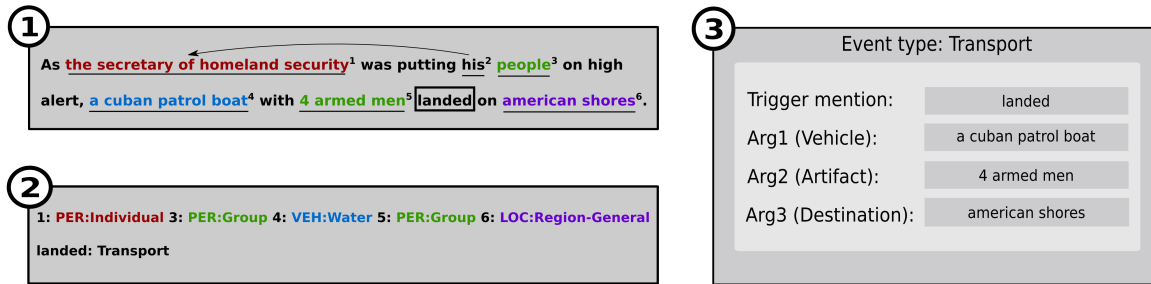


FIGURE 1 – *La reconnaissance d'entités nommées* identifie les différentes mentions d'entités (soulignées dans le cadre (1)) de la phrase et leur type d'entité (cadre (2)). La *résolution de coréférences* permet d'identifier (ici par une flèche) lorsqu'un pronom ou une mention font référence à une même entité. La *détection d'événement* identifie les triggers de la phrase et leur associe un type. Le type du trigger indique le type du formulaire du cadre (3). Les arguments de ce formulaire sont ensuite sélectionnés parmi les entités identifiées précédemment.

vant apparaître plusieurs fois dans un texte (notamment via des pronoms), il s'agit en fait d'extraire des *mentions d'entités* au sein du texte. Trois types d'entités généralement déclinés en sous-types sont généralement distingués : les entités réellement nommées, telles que les noms propres et acronymes de personnes, de lieux ou d'organisations (EN-AMEX), les références temporelles telles que les durées ou les dates (TIMEX) et les valeurs numériques telles que les prix ou les pourcentages (NUMEX).

**Détection d'événement** La majorité des systèmes font l'hypothèse simplificatrice qu'un événement est intégralement défini dans une seule phrase. Ce parti pris est critiquable [28] mais motivé par la plus grande richesse des informations exploitables à l'échelle phrastique. La détection d'événement revient alors à classer des phrases selon un type d'événement prédéfini. La classification d'événement est alors généralement assimilée à la détection de déclencheurs (ou *triggers*) au sein de la phrase. Cette modélisation, introduite par la campagne d'évaluation ACE, est prédominante parmi les approches récentes. Cette simplification n'est pas intrinsèquement nécessaire mais facilite le développement, et notamment l'identification d'événements multiples au sein d'une phrase.

## 2.2 Extraction de liens entre mentions

**Résolution de coréférence** La résolution de coréférence, introduite lors de la campagne MUC-6, vise à identifier au sein du texte des mentions faisant référence à une même entité. Cette étape s'appuie principalement sur l'identification de mentions similaires malgré les possibles variations et sur la résolution d'anaphores pronominales pour identifier les correspondances entre mentions d'entité et entre mentions d'entité et pronoms y faisant référence. Elle permet en particulier d'éviter les redondances et de désambiguïser les pronoms pour l'extraction d'événements et de relations. Bien que cette tâche ait initialement été définie pour les entités, la résolution de coréférences entre événements est également courante et permet notamment la consolidation des extractions.

**Extraction d'événement** Une fois la phrase associée à un type d'événement donné via l'extraction d'un *trigger*, il reste à identifier les entités jouant un rôle dans celle-ci. Cette tâche consiste à prédire, pour chaque entité nommée, son rôle dans l'événement considéré. Dans un cadre supervisé, il est possible de définir un schéma indiquant les types d'entités autorisés pour les différents rôles, ce qui permet de restreindre les possibilités.

## 2.3 Approches séquentielles et approches jointes

Nous avons présenté séparément les différentes tâches car elles sont généralement traitées de manière séquentielle. Cependant, des interdépendances existent entre ces différentes étapes et peuvent être exploitées. Les approches jointes concernent généralement la prédiction conjointe de triggers et d'arguments. Les phrases suivantes explicitent l'interdépendance de ces tâches :

1. A cameraman died when an American tank **fired** on the Palestine Hotel.
2. He has **fired** his air defense chief.

Ici, le mot "fired" est ambigu mais la présence du mot "tank" correspond probablement au rôle *instrument*, ce qui permet de déduire qu'il s'agit d'un événement de type *Attack*. Dans la deuxième phrase, puisque "Air Defense chief" est un intitulé de poste, l'interprétation de "fired" en tant que trigger de type *End-Position* est évidente.

## 3 Campagnes d'évaluation

Nous présentons ici plusieurs campagnes d'évaluations qui ont motivé l'apparition et le développement des tâches présentées précédemment.

### 3.1 Message Understanding Conferences

Les campagnes d'évaluations MUC<sup>1</sup> sont une série de conférences organisées par le DARPA (*Defense Advanced Research Projects Agency*) afin de stimuler la recherche

1. [http://itl.nist.gov/iaui/894.02/related\\_projects/muc/proceedings/proceedings\\_index.html](http://itl.nist.gov/iaui/894.02/related_projects/muc/proceedings/proceedings_index.html)



en extraction d'information, initialement pour l'analyse de documents militaires. La première campagne d'évaluation MUC (1987) était relativement peu contrainte, sans critères d'évaluation formels. Elle portait sur des documents traitant de repérages et d'interventions en mer. À partir de la deuxième édition (1989), la tâche fut précisément définie comme le remplissage d'un formulaire. Cette tâche s'est complexifiée au fil des éditions pour atteindre 47 champs différents au sein de 11 formulaires (types d'événements) différents pour la cinquième édition. Le principal tournant dans ce programme fut l'introduction de la précision et du rappel comme métriques d'évaluations lors de MUC-3.

Les résultats obtenus lors de la cinquième édition (1993) furent plutôt satisfaisants mais les systèmes étaient fortement spécialisés. Cette spécialisation implique un lourd travail d'adaptation pour chaque nouveau domaine, loin de l'objectif de conception d'un système générique et universel. Pour répondre à ces attentes, les deux dernières éditions de MUC ont identifié des sous-tâches considérées comme fondamentales et pouvant faire l'objet d'évaluations spécifiques : la reconnaissance d'entités nommées et la résolution de coréférences.

### 3.2 Automatic Content Extraction

Dans la continuité de MUC, les campagnes ACE<sup>2</sup> présentent une tâche de détection et de suivi d'entités incluant la reconnaissance d'entités nommées et la résolution de coréférences. L'édition 2004 introduit la tâche de détection et de caractérisation de relations, qui consiste à extraire des relations étiquetées selon 24 types. ACE présente également une tâche de détection et de caractérisation d'événements couvrant 6 types d'événements et 33 sous-types.

Les documents annotés proviennent de différentes sources : des dépêches d'agence de presse, des bulletins et débats télévisés, des blogs et groupes de discussion en ligne et enfin des échanges téléphoniques. Cette pluralité de sources permet à ACE 2005 de s'imposer comme un cadre de référence pour l'extraction d'événements, et notamment l'étude de l'adaptation au domaine dans ce contexte. ACE 2008 propose deux versions pour les tâches d'extraction d'entités et de relations. La première, à l'échelle locale, correspond aux tâches d'ACE 2004. La seconde version, à l'échelle du corpus, impose donc la résolution de coréférence multi-document, aussi bien pour les entités que pour les relations.

### 3.3 Autres campagnes

Il existe plusieurs autres campagnes d'évaluation en extraction d'événements ou sur des tâches similaires. La campagne TAC<sup>3</sup> (*Text Analysis Conference*) propose une tâche de peuplement de base de connaissances comprenant plusieurs sous-tâches d'extraction d'événement, d'*entity linking* (identifier et faire correspondre les entités d'un texte à celles d'une base de connaissances existante et ajouter les entités manquantes à la base), *slot-filling* (compléter cer-

tains champs des formulaires de la base à partir des textes). Il existe également plusieurs campagnes d'évaluation en domaine spécialisé, et plus particulièrement dans le domaine biomédical pour permettre d'exploiter les milliers d'articles publiés chaque jour. Les différentes campagnes BioCreatives<sup>4</sup> et I2B2<sup>5</sup> fournissent ainsi des jeux de données annotés sur l'identification de relations entre traitements et maladies ou entre protéines et gènes par exemple.

## 4 Méthodes

Bien que les différentes approches d'extraction d'événement présentent des variations importantes, il est possible de les comparer selon plusieurs axes :

- les systèmes ou les différents modules d'un système peuvent être plus ou moins dépendants du domaine concerné ;
- les modèles peuvent s'appuyer sur un degré variable de connaissances linguistiques ;
- la conception des modèles peut nécessiter plus ou moins de données annotées.

Les premiers systèmes [11] d'extraction d'information à partir d'un texte se voulaient universels. Ils ne faisaient pas l'hypothèse de l'existence d'un domaine particulier ou d'un type d'information spécifique à extraire. Ces systèmes visaient à réaliser une analyse complète du document (syntaxique, sémantique et pragmatique) afin de comprendre le texte dans son ensemble. Bien que cette approche soit théoriquement pertinente pour "résoudre" l'extraction d'information indépendamment du domaine ou de la tâche cible, ces systèmes étaient trop complexes à développer et nécessitaient trop de ressources (les développeurs de TACITUS rapportent 36h de calculs pour les 100 messages du jeu de test MUC-3 [12]) et de connaissances à modéliser. De manière générale, la compréhension de texte suppose une analyse en profondeur de l'intégralité des documents. L'implémentation et l'application de ces traitements s'avèrent impossibles, même à l'heure actuelle. En se fixant une tâche moins ambitieuse par l'introduction de la dimension de domaine et d'information structurée et spécifique (entités, attributs, relations et événements), l'extraction d'information se différencie de la compréhension de texte par la profondeur et la couverture de l'analyse linguistique nécessaire au fonctionnement d'un système.

### 4.1 Approches à bases de connaissances

**Utilisation de motifs lexico-syntaxiques** Ce changement de paradigme de conception permet l'apparition, dès la fin des années 1980, des premiers systèmes d'extraction d'information. Ce changement annonce aussi la prédominance des approches intra-phrastiques. En effet, puisque l'objectif n'est plus la compréhension globale du texte mais l'extraction ponctuelle d'informations, de nombreux systèmes ne travaillent dans un premier temps qu'au niveau local. Une phase de consolidation est généralement réali-

2. <http://www.itl.nist.gov/iad/mig/tests/ace/>  
3. <https://tac.nist.gov/>

4. <http://www.biocreative.org/>  
5. <https://www.i2b2.org/NLP/>

sée à la fin pour fusionner les formulaires construits localement. Ces premiers systèmes (ATRANS [18], SCISOR [26]) tirent leur efficacité de règles et d'expressions régulières définies manuellement. De ce fait, ils nécessitent toujours une conception complexe réalisée spécifiquement par un expert et propre au domaine cible. Ils se caractérisaient par un aspect rigide et monolithique et étaient difficilement adaptables à d'autres langues ou objectifs.

Le système FASTUS [12] popularise sur la campagne MUC-3 l'utilisation d'automates à états finis en cascades et plus généralement l'approche séquentielle. Ce système réalise ainsi séquentiellement 5 étapes de reconnaissance de motifs et de *chunking*, la sortie de chaque module étant l'entrée du module suivant. L'intérêt de cette décomposition est l'apparition de la modularité au sein du système. Celle-ci permet une modification plus aisée qui facilite le développement et l'adaptation du système. De plus, les trois premières étapes opèrent au niveau linguistique et sont très peu dépendantes du domaine. De ce fait, l'adaptation au domaine ne concerne que les 2 dernières étapes.

**Extraction de motifs lexico-syntaxiques** Si les systèmes à base d'automates en cascade marquent un tournant au regard de la lourdeur de la tâche de conception d'un système, cette conception est toujours manuelle et nécessite l'intervention de connaissances expertes à la fois sur le système et sur le domaine. Un nouveau changement de paradigme intervient avec l'utilisation de méthodes d'extraction de motifs. Le travail nécessaire pour l'adaptation à un nouveau domaine est ainsi passé de la conception des règles à l'annotation d'un corpus.

Ces méthodes (RAPIER [21], Autoslog [27]) utilisent différentes représentations des exemples telles que des sacs de mots, l'étiquetage en parties du discours ou des arbres syntaxiques. [10] propose une approche séquentielle d'identification des mentions d'événements puis des arguments pour finir par la classification du type d'événements. Ce système s'appuie sur l'utilisation de structures syntaxiques et de classifieurs séquentiels et constitue donc aussi un pré-curseur des approches à base de classifieurs.

## 4.2 Apprentissage de classifieurs

À la différence des systèmes précédents, les systèmes utilisant des classifieurs considèrent la tâche d'extraction d'événement comme une tâche de classification de séquence. Un texte est un ensemble de phrases traitées comme des séquences de *tokens*. La détection d'événement consiste alors à appliquer à chaque élément de la séquence un classifieur entraîné à détecter les *triggers* et leur type, séquentiellement ou de manière jointe. Il en va de même pour les arguments, généralement prédits parmi les entités nommées détectées en amont.

On dénote au sein de cette famille d'approches deux tendances. La majorité des études utilisent des approches séquentielles en traitant d'abord l'identification puis la classification des *triggers* puis des arguments ([1, 6, 10]). Mais certaines études utilisent également des approches jointes

[4, 16]. Ces méthodes tentent de réduire le problème de propagation des erreurs symptomatique des approches séquentielles. De plus, elles peuvent ainsi tenir compte de l'interdépendance entre arguments et *triggers* ou entre détection et caractérisation des *triggers* ou des arguments. Néanmoins, ces approches se rejoignent sur les types de classifieurs et de représentations choisis. Ces approches ont évolué parallèlement à celles concernant l'extraction de relations. C'est pourquoi nous citerons ici indifféremment des études portant sur les deux tâches. Les classifieurs sont le plus souvent des machines à vecteurs de support [33, 17, 13] ou des classifieurs de type maximum d'entropie [29, 23]. L'efficacité de ces approches étant particulièrement dépendante de la qualité des représentations choisies, la création de représentations adaptées est essentielle. Les approches à bases de classifieurs n'ont ainsi supprimé l'effort d'élaboration de règles que pour le remplacer par un effort d'ingénierie des représentations aussi conséquent.

Il apparaît cependant qu'une fois des représentations efficaces obtenues, celles-ci s'avèrent assez génériques pour être transposées dans des domaines proches. [33] introduit pour l'extraction de relations la plupart des traits (*features*) utilisés dans l'état de l'art. Ces représentations sont produites à plusieurs niveaux : au niveau lexical (sac de mots et tête de mention pour chaque mention, premiers et deuxièmes mots des contextes gauche, milieu, et droit), syntaxique (chemins dans l'arbre syntaxique complet entre les deux mentions, *chunking* puis extraction des têtes des groupes nominaux) et sémantique (utilisation des types d'entités ACE et de WordNet [20]). [29] reprend ces représentations et complète la représentation lexicale par l'utilisation de bigrammes des mots du contexte central. D'autres informations sémantiques sont proposées, telles que l'utilisation de synonymes de WordNet [16] ou d'hyponymes de Framenet [15]. Il est à noter que le niveau de granularité maximum de ces représentations est généralement le mot bien que des approches descendent au niveau des morphèmes pour l'extraction d'information en chinois [6]. La représentation des mots est généralement de type local ou *one-hot*, c'est-à-dire par un vecteur binaire de taille  $N$  où  $N$  est la taille du vocabulaire et dont seule la dimension correspondant au mot est active. Cette représentation symbolique pose deux problèmes [30] : d'une part elle ne permet pas de capturer convenablement la sémantique du mot (ce à quoi tentent de pallier les approches par morphèmes); d'autre part, elle est particulièrement parcimonieuse et sujette au fléau de la dimension. Enfin, si les vocabulaires cible et source sont différents, le système n'aura aucune information sur les mots nouveaux.

## 4.3 Plongement lexical

Le plongement lexical (ou *word embeddings*) est une représentation distribuée des mots permettant de répondre aux deux problèmes soulevés à la section précédente. Dans une représentation locale (ou symbolique), un élément est associé à une représentation unique (un indice). Au contraire,

dans une représentation distribuée, un élément est décrit par plusieurs indices et un indice est utilisé pour décrire plusieurs éléments. Si la représentation locale est plus facile à comprendre et à produire par un humain, elle ne permet pas de capturer la proximité sémantique entre les éléments ou d'isoler différentes propriétés sous-jacentes.

Pour le plongement lexical, l'hypothèse distributionnelle ("You shall know a word by the company it keeps!" [9]) postule que des mots apparaissant dans des contextes similaires ont des sens similaires. On peut apprendre ces représentations sur de grands corpus de textes non annotés. Ceci confère un autre avantage à ces représentations : elles permettent d'assurer une meilleure robustesse aux systèmes d'extraction d'information, même pour des mots non présents dans le jeu d'apprentissage. Plusieurs approches successives ont proposé des représentations distribuées, notamment l'analyse sémantique latente [8] et le clustering de brown [3]. Il apparaît ensuite une série de représentations [2, 7, 19] extraites de réseaux de neurones entraînés à prédire un mot à partir de son contexte ou inversement. Leur utilisation en extraction de relations commence avec [29] qui exploite des *clusters* de Brown pour augmenter les descripteurs des têtes de mentions.

#### 4.4 Architectures neuronales

Peu après l'introduction de ces représentations distribuées et en réponse aux propagations d'erreurs et aux difficultés d'adaptation inhérentes aux systèmes séquentiels, de nombreuses études ont commencé à appliquer des approches neuronales à différentes tâches de l'extraction d'information. Elles mettent en avant la suppression des prétraitements à l'origine de ces problèmes. Ces approches neuronales utilisent différentes architectures de réseaux de neurones popularisées dans les communautés de reconnaissance de la parole et de vision par ordinateur. Plusieurs systèmes d'extraction de relations [14, 32] proposent des systèmes extrayant une représentation de la phrase à l'aide d'un réseau convolutif (CNN). Ces systèmes ne fournissent en entrée que les *word embeddings* des *tokens* de la phrase, généralement ceux de [19]. Ces *word embeddings* sont modifiés durant l'apprentissage, ce qui permet de les adapter au domaine cible. L'opération de *pooling* du réseau convolutif ne conservant que l'information prédominante d'une phrase, [5] propose une variante du CNN utilisant le *dynamic multipooling* (ou *piece-wise CNN* [31]), c'est-à-dire l'extraction de représentations pour plusieurs parties de la phrase, ce qui permet notamment de mieux gérer les phrases contenant de multiples événements. Par ailleurs, l'application de convolutions est surtout adaptée à la détection de motifs locaux et consécutifs. Des systèmes utilisent au contraire des réseaux récurrents pour construire des représentations pouvant tirer profit de motifs non consécutifs à l'échelle de la phrase. De manière générale, un des intérêts de ces méthodes neuronales est d'obtenir une abstraction de la représentation d'entrée, de telle sorte que la représentation en sortie soit relativement invariante à de

TABLE 1 – Résultats de différents systèmes de détection d'événements (F1-mesure pour développement et test)

Méthodes	dév.	test
Meilleure méthode non neuronale : prédiction structurée [16] avec descripteurs locaux et globaux	67,9	67,5
CNN avec <i>embeddings</i> (mots) [24]	14,0	–
CNN avec <i>embeddings</i> (mots + positions) [24]	68,5	67,6
CNN avec <i>embeddings</i> (mots + positions + entités) [24]	70,7	69,0

faibles changements locaux, augmentant ainsi la robustesse de la représentation finale.

#### 4.5 Généralisation des représentations distribuées

Les différents systèmes ayant redéfini l'état de l'art à l'aide des approches neuronales ont tous mis en avant l'objectif de réduction des prétraitements nécessaires, notamment syntaxiques et sémantiques, pour se concentrer sur les représentations lexicales distribuées. On peut cependant rapidement voir la réintroduction d'informations syntaxiques puis sémantiques. Dans un premier temps, pour l'extraction de relations, la représentation en entrée du réseau est augmentée d'un vecteur codant la distance aux deux entités concernées (*position feature*) et qui permet, à titre indicatif, de gagner dans [32] plus de 9 points en F1-mesure pour l'extraction de relations et jusqu'à 54 points en extraction d'événements dans [24] (cf. Table 1). La majorité des systèmes augmente également la représentation de chaque *token* par un *embedding* de son type d'entité (*entity feature*, +2 points) [24], ou d'informations syntaxiques extraites d'arbres syntaxiques ou d'un *chunker* [22, 25]. Ces systèmes montrent bien que ces différentes informations linguistiques initialement supprimées des approches neuronales sont importantes pour le développement de systèmes performants. L'utilisation de représentations distribuées pour représenter ces informations confère plusieurs avantages déjà identifiés pour les représentations distribuées de mots. D'une part, ces représentations sont plus à même d'isoler les facteurs sous-jacents à ces différentes informations. D'autre part, ces représentations dans un espace dense sont plus expressives et permettent au réseau de tenir compte de la similarité entre différentes modalités. Enfin, puisque ces représentations sont apprises et transférables, elles permettent une meilleure adaptation à de nouveaux domaines et offrent la possibilité d'un apprentissage semi-supervisé. Ainsi, de la même manière que le pré-entraînement de *word embeddings* sur une base annexe offre une meilleure généralisation, cet apprentissage semi-supervisé est possible pour les autres types de représentations distribuées.

#### Conclusion et perspectives

Les premiers systèmes d'extraction d'information nécessitent l'élaboration manuelle de motifs lexico-syntaxiques finement adaptés à une tâche et un domaine donnés. Ceci

les rendaient coûteux et complexes à développer et particulièrement sensibles à un changement de domaine ou un décalage entre les jeux d'apprentissage et les jeux de test. Les approches suivantes utilisant des documents annotés nécessitaient toujours un temps de développement long, tout en étant plus génériques et modulables. L'introduction des représentations distribuées et des réseaux de neurones a ensuite grandement amélioré l'état de l'art. La synergie de ces différentes représentations permet de combler les faiblesses des précédents algorithmes sur plusieurs points. D'une part, elles permettent de réaliser un apprentissage semi-supervisé grâce à des bases annexes, notamment pour les *word embeddings* et d'autre part, elles offrent une meilleure expressivité et une meilleure interprétabilité par les classificateurs que les représentations symboliques. Cependant, l'abandon de ces dernières se fait au détriment de l'interprétabilité humaine du système, constat déjà réalisé pour la reconnaissance d'objets en vision. À l'avenir, ces représentations distribuées pourraient notamment être appliquées à l'extraction d'information ouverte en transférant des représentations apprises de manière semi-supervisée à cette tâche d'extraction peu ou pas supervisée.

## Références

- [1] David Ahn. The stages of event extraction. In *Workshop on Annotating and Reasoning about Time and Events*, 2006.
- [2] Yoshua Bengio, Réjean Ducharme, Pascal Vincent, and Christian Jauvin. A neural probabilistic language model. 3, 2003.
- [3] Peter F. Brown, Peter V. Desouza, Robert L. Mercer, Vincent J. Della Pietra, and Jenifer C. Lai. Class-based n-gram models of natural language. 18(4), 1992.
- [4] Chen Chen and Vincent Ng. Joint modeling for chinese event extraction with rich linguistic features. In *COLING*, 2012.
- [5] Yubo Chen, Liheng Xu, Kang Liu, Daojian Zeng, and Jun Zhao. Event extraction via dynamic multi-pooling convolutional neural networks. In *ACL-IJCNLP*, 2015.
- [6] Zheng Chen and Heng Ji. Language specific issue and feature exploration in chinese event extraction. In *NAACL-HLT*, 2009.
- [7] Ronan Collobert and Jason Weston. A unified architecture for natural language processing : Deep neural networks with multitask learning. In *ICML*, 2008.
- [8] S. T. Dumais, G. W. Furnas, T. K. Landauer, S. Deerwester, and R. Harshman. Using latent semantic analysis to improve access to textual information. In *SIGCHI*, 1988.
- [9] John R Firth. A synopsis of linguistic theory, 1930-1955. 1957.
- [10] Ralph Grishman, David Westbrook, and Adam Meyers. NYU's english ACE 2005 system description. In *ACE*, 2005.
- [11] Jerry R. Hobbs. Overview of the tacitus project. 1986.
- [12] Jerry R. Hobbs, Douglas E. Appelt, John Bear, David Israel, Megumi Kameyama, Marc Stickel, and Mabry Tyson. FASTUS : A cascaded finite-state transducer for extracting information from natural-language text. 1997.
- [13] Yu Hong, Jianfeng Zhang, Bin Ma, Jianmin Yao, Guodong Zhou, and Qiaoming Zhu. Using cross-entity inference to improve event extraction. In *ACL-HLT*, 2011.
- [14] Yoon Kim. Convolutional neural networks for sentence classification. In *EMNLP*, 2014.
- [15] Qi Li, Heng Ji, Yu Hong, and Sujian Li. Constructing Information Networks Using One Single Model. In *EMNLP*, 2014.
- [16] Qi Li, Heng Ji, and Liang Huang. Joint event extraction via structured prediction with global features. In *ACL*, 2013.
- [17] Shasha Liao and Ralph Grishman. Using document level cross-event inference to improve event extraction. In *ACL*, 2010.
- [18] Steven L. Lytinen and Anatole Gershman. ATRANS automatic processing of money transfer messages. In *AAAI*, 1986.
- [19] Tomas Mikolov, Ilya Sutskever, Kai Chen, Greg S. Corrado, and Jeffrey Dean. Distributed representations of words and phrases and their compositionality. In *Advances in Neural Information Processing Systems*, 2013.
- [20] George A. Miller. WordNet : A lexical database for English. 38(11), 1995.
- [21] Raymond J. Mooney and Mary E. Califf. Relational learning of pattern-match rules for information extraction. In *AAAI*, 1999.
- [22] Thien H. Nguyen, Kyunghyun Cho, and Ralph Grishman. Joint event extraction via recurrent neural networks. In *NAACL-HLT*, 2016.
- [23] Thien H. Nguyen and Ralph Grishman. Employing word representations and regularization for domain adaptation of relation extraction. In *ACL*, 2014.
- [24] Thien H. Nguyen and Ralph Grishman. Event detection and domain adaptation with convolutional neural networks. In *ACL-IJCNLP*, 2015.
- [25] Thien Huu Nguyen and Ralph Grishman. Combining neural networks and log-linear models to improve relation extraction. In *IJCAI Workshop on Deep Learning for Artificial Intelligence*, 2016.
- [26] Lisa F. Rau. Conceptual information extraction from financial news. In *HICSS*, 1988.

- [27] Ellen Riloff. Automatically constructing a dictionary for information extraction tasks. In *AAAI*, 1993.
- [28] Mark Stevenson. Fact distribution in Information Extraction. 40, 2006.
- [29] Ang Sun, Ralph Grishman, and Satoshi Sekine. Semi-supervised relation extraction with large-scale word clustering. In *ACL-HLT*, 2011.
- [30] Joseph Turian, Lev Ratinov, and Yoshua Bengio. Word representations : A simple and general method for semi-supervised learning. In *ACL*, 2010.
- [31] Daojian Zeng, Kang Liu, Yubo Chen, and Jun Zhao. Distant supervision for relation extraction via piecewise convolutional neural networks. In *EMNLP*, 2015.
- [32] Daojian Zeng, Kang Liu, Siwei Lai, Guangyou Zhou, and Jun Zhao. Relation classification via convolutional deep neural network. In *COLING*, 2014.
- [33] GuoDong Zhou, Jian Su, Jie Zhang, and Min Zhang. Exploring various knowledge in relation extraction. In *ACL*, 2005.



# Planification probabiliste : une heuristique pour garantir des solutions sûres

Rémi Lacaze-Labadie<sup>1</sup>Domitile Lourdeaux<sup>1</sup>Mohamed Sallak<sup>1</sup>

<sup>1</sup> Sorbonne universités, Université de technologie de Compiègne, CNRS, Heudiasyc UMR 7253, 57 avenue de Landshut – 60203 COMPIEGNE Cedex, France

{remi.lacaze-labadie - domitile.lourdeaux - mohamed.sallak}@hds.utc.fr

## Résumé

*Dans ce papier, nous proposons une nouvelle approche pour résoudre les problèmes de planification probabiliste en prenant en compte le risque que l'utilisateur est prêt à prendre vis à vis de la probabilité d'atteindre les buts. Notre approche garantit le fait que la probabilité d'atteindre un état satisfaisant les buts du problème sera toujours supérieure à un seuil de probabilité donné. Toutefois, un tel seuil ne permet pas toujours de tous les atteindre. C'est pourquoi, notre heuristique estime la probabilité d'atteignabilité de chaque but via une version relaxée du problème, puis sélectionne le sous-ensemble de buts atteignables avec une probabilité supérieure à ce seuil. Nous illustrons nos travaux par un problème d'exploration planétaire couramment utilisé dans ce contexte. Nos résultats de simulation montrent que l'heuristique proposée sélectionne les meilleurs buts et que le seuil de probabilité de les atteindre est toujours respecté.*

## Mots Clef

Planification dans l'incertain, heuristique, théorie de la décision.

## Abstract

*In this work, we propose a novel approach to solve probabilistic planning problems taking into account the risk that the decision maker is ready to accept regarding the probabilities of reaching the goals. Our approach guarantees that the probability to reach a state satisfying selected goals is above a certain limit threshold. Indeed, we relax the constraints that all goals must be satisfied, and select the most valuable set of goals whose reachability probability is above the threshold. To this end, we propose a goal selection heuristic based on the reachability probability and the cost between goals that are estimated using an abstracted version of the problem. Finally, a planetary exploration problem will be used for illustrating the effectiveness of the proposed approach. Our results show that the obtained selections cover the most valuable possible goals and respect the reachability probability threshold.*

## Keywords

Planning under uncertainty, heuristic search, decision theory.

## 1 Introduction

La planification est une branche de l'intelligence artificielle qui consiste à sélectionner et à ordonnancer des actions afin d'atteindre un objectif [6], c'est à dire un état satisfaisant tous les buts du problème. En particulier, nous nous intéresserons à la planification probabiliste, dont les effets des actions peuvent être non déterministes. Dans ce papier, nous définirons un objectif comme l'ensemble des buts d'un problème, c'est à dire l'ensemble des prédicats devant être satisfait par la solution. Nous formaliserons ces problèmes avec des processus de décision markovien (ou Markov Decision Process en anglais, MDP). Les MDPs peuvent être efficacement résolus à l'aide de la solution unique de l'équation de Bellman sous deux hypothèses : il existe au moins une politique qui atteint tous les buts avec une probabilité de 1 et tous les coûts des actions sont strictement positifs de sorte que toutes les politiques contenant une boucle infinie accumulent un coût infini. Les MDPs formulés sous ces hypothèses sont plus connus sous le nom de problèmes de Plus Court Chemin Stochastique (SSP en anglais). Ces hypothèses garantissent d'avoir toujours une solution à l'équation de Bellman. Cependant, elles sont contraignantes et ne permettent pas de modéliser tous les types de problèmes. C'est par exemple le cas pour les problèmes qui n'ont pas de politique permettant d'atteindre avec une probabilité de 1 tous les buts. Dans ce cas, nous allons chercher à maximiser cette probabilité. Des travaux ont déjà été menés sur le sujet, comme par exemple Teichteil-Königsburg [18] qui propose une méthode efficace pour trouver le chemin le plus probable pour atteindre les buts. Cette méthode est connue dans la littérature sous le nom de Stochastic Safest and Shortest Path ( $S^3P$ ) et résout les problèmes en recherchant d'abord tous les chemins optimisant la probabilité d'atteindre les buts, puis en sélectionnant celui minimisant le coût. Cependant cette méthode ne garantit pas une probabilité minimale d'atteindre les buts. L'objectif de ce papier est de proposer une heuristique sélectionnant un sous-ensemble de buts atteignable

avec un seuil de probabilité minimal choisit par le décideur. Nous positionnons nos travaux dans la catégorie des problèmes sur-contraints (*over-constrained problems* en anglais), dans lesquels les contraintes sont le manque de ressource (essence, énergie,...) empêchant tous les buts d'être atteints. En effet, nous considérons le risque, soit la probabilité de ne pas atteindre un but, comme une ressource limitant le nombre de buts atteignables. Dans cet article, nous ferons l'hypothèse que le décideur, en cas de situations à risque, préfère satisfaire moins de buts, avec cependant de meilleures chances qu'ils soient atteints. Pour ce faire, nous proposons de définir un seuil de probabilité d'atteindre les buts, que nous définirons dans le reste de l'article comme le *seuil de sûreté*. Nous proposons alors une méthode qui sélectionne ce que nous appellerons une *solution sûre* du point de vue du décideur, c'est-à-dire toutes solutions satisfaisant ce *seuil de sûreté*.

Les problèmes sur-contraints ont été étudiés dans le domaine de la planification sous le nom anglais de *Partial Satisfaction Planning problems (PSP)* [19] et *Over-Subscription Planning problems (OSP)* [14]. Le challenge de ce type de problèmes est la sélection du sous-ensemble de buts satisfaisant les contraintes. En effet, une recherche complète dans l'ensemble de l'espace d'état (pour trouver les meilleurs ensembles de buts) est impossible en pratique. De bonnes heuristiques ont alors été proposées pour sélectionner les buts [2, 10], cependant elles sont toutes basées sur des contraintes relatives aux limites de ressources. À notre connaissance, aucune heuristique ne propose une sélection des buts garantissant une probabilité de les atteindre au-dessus d'un seuil donné. L'idée de base de notre algorithme est de construire une version relaxée du problème afin d'estimer le coût entre les buts ainsi que la probabilité d'atteindre un but depuis un autre but. Puis, à partir de cette version relaxée du problème, un graphe est créé dont les nœuds représentent les buts et les arcs les coûts et probabilités. Enfin, la sélection des meilleurs buts se fait par un algorithme de recherche dans ce graphe.

Ce papier est organisé de la manière suivante. Dans la section 2, nous présentons nos motivations ainsi qu'un état de l'art sur les problèmes sur-contraints en planification. Puis nous formalisons notre problème et proposons notre heuristique dans les sections 3 et 4. La section 5 décrit nos résultats empiriques sur un cas d'étude, avant de conclure par nos perceptions à la section 6.

## 2 Motivation et contexte

Nos travaux s'inscrivent dans une catégorie de problèmes dans laquelle le décideur préfère satisfaire moins de buts si cela augmente la probabilité qu'ils se produisent. Dans cette section, nous donnons nos motivations à travers un exemple, discutons des problèmes de satisfaction de contraintes en planification puis concluons avec nos contributions dans ce domaine.

### 2.1 Motivation

Nous illustrons notre méthode tout au long de ce papier par un problème d'exploration planétaire (plus connu sous le nom de *mars rover*), dans lequel un robot doit naviguer sur mars, visiter différents points d'intérêts et effectuer des opérations telles que la collecte d'échantillons. Bien que ce problème ait été étudié dans le cas de ressources limitées (e.g. énergie) [14, 11], nous constatons que dans le cadre de missions qui peuvent s'avérer coûteuses, il est également important de considérer la notion de risque. En effet, le robot pourrait par exemple rester bloqué par un cratère, ce qui causerait sa perte. Pour mettre en évidence ces aspects, nous modifions légèrement ce problème en restreignant les buts à l'exploration de différents points d'intérêts, et en ajoutant la notion de risque sur les déplacements du robot. C'est à dire qu'à chaque déplacement, en fonction de la présence de dangers sur le sol, le robot a une probabilité de rester bloqué (ce qui provoque un échec de la mission). L'objectif de notre méthode sur ce problème est de trouver les meilleurs points d'intérêts à visiter, tout en garantissant que le robot ne prendra pas trop de risques et ne restera pas bloqué.

### 2.2 Problème de satisfaction de contraintes en planification

L'objectif de la planification classique est d'atteindre un état satisfaisant tous les buts du problème. Cependant, il n'est pas toujours possible de tous les atteindre, soit pour des raisons de conflits logiques entre les buts, soit par manque de ressources. On dit alors que le problème est sur-contraint. Dans ce cas, une solution peut seulement satisfaire un sous-ensemble de buts. On se retrouve alors dans un problème de planification sous contraintes dans lequel on cherche le meilleur sous-ensemble de buts satisfaisant les contraintes. Ces problèmes sont connus sous les noms de OSP, premièrement introduit par *Smith* [14] et de PSP introduit plus tard par *Briel et al.* [19]. Bien que ces deux types de problème ont les mêmes définitions et contraintes (l'ensemble des buts n'est pas atteignable), ils ont des perspectives différentes.

**Partial satisfaction planning (PSP).** *Le PSP Net Benefit* [13] est un problème dans lequel des scores (préférences) sont attribués aux différents buts, et des coûts aux actions. La solution recherchée est un plan avec le meilleur bénéfice net, c'est-à-dire le cumul des scores moins le coût pour atteindre les buts. L'inconvénient de cette approche est de supposer que les scores et les coûts sont comparables et donc du même ordre. Cependant, dans notre approche la probabilité d'atteindre un but n'est pas comparable avec un score, c'est pourquoi nous nous sommes tournés vers les OSPs.

**Over-subscription planning (OSP).** Dans les problèmes OSPs, l'objectif est de maximiser le score des buts atteints tout en respectant la limite imposée sur les ressources/coûts disponibles. Les méthodes utilisées pour ce



type de problèmes s'appuient souvent sur une approche en deux étapes [14, 4]. La première est de sélectionner, via une heuristique, un sous-ensemble de buts satisfaisant les contraintes du problème. La seconde étape est d'utiliser cette sélection pour guider le planificateur. Les buts sont donnés un par un au planificateur, qui planifie chaque but et s'arrête quand ils sont tous atteints ou que les ressources sont épuisées. La difficulté de cette approche est la sélection des meilleurs buts, car il n'est pas possible de faire une recherche complète dans tout l'espace d'état. Pour simplifier ce processus, nous réalisons une estimation des ressources utilisées pour atteindre les buts à l'aide d'une version relaxée du problème. Par exemple, *Smith* [14] propose une abstraction du problème sous la forme d'un Orienteering Problem (OP), qui est un graphe orienté dans lequel chaque nœud est un but (associé d'un score), et chaque arc l'estimation du coût entre les deux buts qu'il relie. L'OP est ensuite résolu par la recherche d'un chemin maximisant les scores sans dépasser la limite imposée sur les coûts. Similairement, *García et al.* [4] construisent une matrice de distance depuis une version relaxée du problème, et applique ensuite un algorithme de recherche en faisceau. Ces méthodes ont été appliquées sur des domaines déterministes, mais ne sont pas adaptées aux domaines probabilistes.

Dans les travaux de *Meuleau et al.* [11], le cas probabiliste a été étudié sous le nom de *Stochastic Over-Subscription Planning* (SOSP). La première étape de leur méthode est la résolution d'un MDP pour chaque but du problème (également appelé sous-tâche). Ces MDPs sont beaucoup plus petits et donc plus simple à résoudre que le problème initial. La seconde étape consiste à planifier ces sous-tâches une par une pour former la solution complète du problème. Bien que cette approche est adaptée aux domaines probabilistes, elle adopte une attitude neutre (*risk-neutral attitude*). A l'inverse, mais sur le même principe que *Geibel* [5], qui définit le risque d'une solution comme la probabilité d'entrer dans un état fatal ou non désiré, nous définissons le degré de sûreté d'une solution par sa probabilité d'entrer dans un état satisfaisant les buts sélectionnés par l'heuristique. Dans l'objectif de garantir une *solution sûre*, notre méthode de sélection des buts repose sur la construction d'un graphe dont chaque arc représente la probabilité d'atteindre un but depuis un autre. A l'aide de ce graphe, les buts sont sélectionnés un par un en respectant que la multiplication des probabilités reste au-dessus du *seuil de sûreté*.

### 2.3 Contribution

Comme vu dans la section précédente, ce sont les OSPs qui sont les plus proches de notre problématique, et c'est sur ces modèles que nous nous appuyerons. Nous proposons une approche pour limiter les risques liés aux problèmes de planification probabiliste, en réduisant le nombre de buts à atteindre afin de garantir une *solution sûre* du point de vue du décideur. Voici les points essentiels de notre contribution :

1. Une méthode pour estimer la probabilité de but

(*RP*), soit la probabilité d'atteindre un but depuis un autre but.

2. Un algorithme garantissant une *solution sûre*, définie avec un *seuil de sûreté* configurable.

## 3 Heuristique pour la sélection des buts

Chaque but a un score, qui représente son intérêt par rapport aux autres buts. L'heuristique pour la sélection des buts est en charge de choisir les meilleurs buts (avec la meilleure somme des scores) atteignables avec une probabilité supérieure au seuil de sûreté. Cette sélection est ensuite utilisée pour guider le planificateur. Notre heuristique se décompose en trois étapes :

1. Création d'une version relaxée du problème, réduisant l'espace d'état et permettant l'identification d'*états but* (état satisfaisant au moins un but du problème).
2. Une estimation au travers du problème relaxé, de la probabilité de but et du coût entre tous les états but identifiés.
3. La recherche dans un graphe (construit à l'aide des états but et estimations précédemment calculés) des meilleurs buts à atteindre.

Avant de détailler ces trois étapes, nous donnons la forme du problème relaxé utilisée pour formaliser l'heuristique.

**Formalisation.** Le problème relaxé est formalisé sous la forme d'un MDP [12]. L'ensemble des buts de ce MDP, que nous notons  $S_G$ , correspond aux états but identifiés lors de la phase de relaxation. Nous faisons le choix d'utiliser les MDPs car ils formalisent très bien les problèmes probabilistes. Le formalisme utilisé est donc un MDP sans facteur d'actualisation (*undiscounted MDP*) sous la forme  $\langle S, A, S_G, p, c, u, \alpha, s_0 \rangle$  tel que :

- $S$  est un ensemble fini d'états ( $s_0 \in S$  étant l'état initial).
- $A$  est un ensemble fini d'actions. Nous notons  $A(s)$  le sous-ensemble des actions applicables dans un état  $s$ .
- $S_G \subseteq S$  est une liste d'états but ( $s_g \in S_G$  désignant un état but).
- $u(s_g) > 0$  est le score de l'état but  $s_g$ .
- $p(s, a, s') \in [0, 1]$  est une fonction de transition qui donne la probabilité d'aller d'un état  $s$  à un état  $s'$  en appliquant l'action  $a$ .
- $c(s, a) > 0$  est le coût de l'action  $a$  dans l'état  $s$ .
- $\alpha \in [0, 1]$  est le *seuil de sûreté* et  $\beta = 1 - \alpha$  est le *seuil de risque*.

### 3.1 Relaxation du problème

Chercher la meilleure sélection des buts, via une recherche complète dans l'espace d'état, est en pratique impossible. En effet, le nombre d'états but peut vite devenir très grand. C'est pourquoi des méthodes d'abstraction de l'espace

d'état ont été utilisées pour résoudre ce type de problème [9, 8]. La plupart de ces méthodes appliquées aux OSPs créent une version relaxée du problème sous forme de graphe, dont les nœuds représentent les états but étiquetés de leurs scores et les arcs les ressources nécessaires pour aller d'un état but à un autre [14, 4]. La sélection des buts se fait alors à l'aide d'algorithmes de recherche dans un graphe, dont le chemin recherché doit maximiser la somme des nœuds tout en respectant une limite de capacité sur les arcs. La méthode de relaxation que nous utilisons dans notre heuristique est très proche de celles proposées par *Smith* ou *Benton et al.* [14, 2]. Pour être efficace, cette méthode suppose que les différents buts sont suffisamment indépendants entre eux, hypothèse que nous pouvons faire dans le cadre de notre cas d'application où les buts sont des sous-tâches indépendantes que le robot doit réaliser (visiter différents points d'intérêts). Nous ne rentrons pas dans les détails de la relaxation car ce n'est pas l'objectif de cet article, nous donnons cependant l'idée générale. Cela consiste à identifier les interactions et variables partagées entre les buts, ce qui permet ensuite de former un espace d'état restreint et donc l'identification des états but. Chaque état but identifié sur le problème relaxé correspond à un but du problème initial. Le plan relaxé ainsi formé, est une version simplifiée du problème, contenant suffisamment d'informations pour estimer les coûts et probabilités entre les buts.

### 3.2 Fonction probabilité de but et coût de but

S'inspirant de Teichteil-Königsbuch [17], et pour des raisons d'uniformisation, nous appellerons la fonction estimant la probabilité d'atteindre un but, la fonction probabilité de but (noté  $RP$ ). Quant à la fonction coût de but (noté  $V$ ), elle estime les coûts entre les états but. Ces deux fonctions sont des *fonctions de valeur*, qui associent une valeur pour chaque état. Nous les approximations via une méthode de *programmation dynamique* [16, 3] avec un algorithme *Value Iteration* appliqué au problème relaxé. L'algorithme *Value Iteration* s'appuie sur la résolution de l'équation de Bellman [1], qui approche la valeur exacte par raffinement itératif. Cette équation est définie comme suit :

$$V(s) = \min_{a \in A(s)} Q(s, a) \quad (1)$$

$$Q(s, a) = c(a, s) + \sum_{s'} p(s, a, s') V(s')$$

En général, une fonction de valeur est une estimation de l'intérêt d'être dans un état  $s$ , par rapport à l'ensemble des buts du problème. Dans le cas sur-contraint, nous estimons les fonctions de valeur, par rapport à chaque état but du problème relaxé. C'est donc un vecteur de fonctions de valeur, de la taille du nombre d'états but.

**Approximation des fonctions de valeur.** Tandis que l'approximation de la fonction coût de but est une application directe de l'équation de Bellman, la fonction probabilité de but en est une variante où les valeurs initiales sont

égales à 1 quand on entre dans un état but et 0 sinon. Sur le même principe que la formulation proposée par *Steinmetz et al.* [15], nous définissons  $RP(s, s_g)$ , la probabilité d'entrer dans un état but  $s_g$  depuis un état  $s$  et  $\sigma(s, a, s_g)$  la probabilité d'entrer dans un état but  $s_g$  depuis un état  $s$  en appliquant l'action  $a$  :

$$RP(s, s_g) = \max_{a \in A(s)} \sigma(s, a, s_g) \quad (2)$$

$$\sigma(s, a, s_g) = \sum_{s'} p(s, a, s') RP(s', s_g) \quad (3)$$

La fonction coût de but  $V(s, s_g)$  est défini sous la forme de l'équation de Bellman (cf. Eq.1), non pas en sélectionnant l'action qui minimise le coût mais l'action qui maximise la probabilité de but. De la sorte, les deux fonctions de valeur sont mise à jour avec la même action (i.e. qui minimise  $\sigma$ ) et donc les mêmes décisions de l'agent. La fonction coût de but est alors une estimation des futurs coûts quand l'agent agit en maximisant sa probabilité de but.

$$V(s, s_g) = Q(s, a, s_g) \text{ s.t. } a = \operatorname{argmax}_{a \in A(s)} \sigma(s, a, s_g) \quad (4)$$

$$Q(s, a, s_g) = c(a, s) + \sum_{s'} p(s, a, s') V(s', s_g) \quad (5)$$

Dans ce papier, nous faisons l'hypothèse qu'il existe au moins une politique avec une probabilité de succès non nulle et que tous les coûts sont strictement positifs. Ainsi toutes les politiques non-propres (*improper policies* en anglais) accumulent un coût infini, ce qui permet à l'équation de Bellman de toujours avoir une solution. Une fois que nous avons estimé la probabilité de but et le coût, on peut en déduire comme un cas particulier, la probabilité et le coût pour aller d'un état but  $s_{g1}$  à un autre état but  $s_{g2}$ , soit  $RP(s_{g1}, s_{g2})$  et  $V(s_{g1}, s_{g2})$ .

### 3.3 Sélection des buts

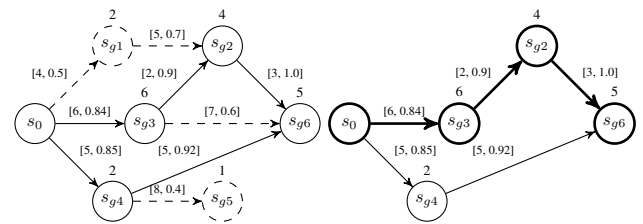


FIGURE 1 – Exemple d'un *SGGraph* avec  $\alpha = 0.75$  et des arcs étiquetés avec  $[V, RP]$ . Le meilleur chemin est  $(s_0, s_{g3}, s_{g2}, s_{g6})$  et les meilleurs buts  $\{s_{g3}, s_{g2}, s_{g6}\}$ .

Le processus de sélection des buts est chargé de choisir et d'ordonner le meilleur sous-ensemble de buts pouvant être satisfait par une solution sûre. Dans la plupart des travaux similaires [14, 11, 4], la sélection se fait par la résolution d'un *orientating problem* (OP). Un OP [7, 20] est un problème de graphe dans lequel chaque nœud a un score, et dont l'objectif est de trouver un chemin Hamiltonien maximisant la somme des scores sans dépasser une

certaine limite. Pour les problèmes OSPs, un graphe est construit depuis la version relaxée, ce qui permet ensuite de trouver les meilleurs buts via un algorithme de résolution d'OP. Dans notre cas, nous appelons *SGGraph*, le graphe orienté dont les nœuds sont l'état initial et les états but identifiés lors de la phase de relaxation. Les arcs du graphe représentent les dépendances entre états but (probabilité de but et coût de but). Le *SGGraph* est ensuite résolu comme un OP, en cherchant un chemin dont le produit des probabilités de but est supérieur au seuil de sûreté  $\alpha$ . Un exemple de *SGGraph* est donné dans la Figure 1. Plus formellement, nous définissons  $SGGraph = (S, E)$  avec  $S$  l'ensemble des nœuds tel que  $S = S_g \cup \{s_0\}$  et  $E$  les arcs avec pour étiquette la probabilité de but et le coût pour aller d'un nœud  $s_1$  à un autre nœud  $s_2$  sous la forme :  $[V(s_1, s_2), RP(s_1, s_2)]$ . Nous définissons également  $RP(path)$ , la probabilité d'atteindre tous les états but d'un chemin. C'est le produit des  $RP$  pour aller de l'état initial au premier état but du chemin, puis d'état but en état but dans l'ordre donné par le chemin. A noter que tout chemin valide doit commencer par l'état initial.

$$path = (s_0, s_{g0}, s_{g1}, \dots, s_{gn}) \text{ s.t. } s_{gi} \in S_g$$

$$RP(path) = RP(s_0, s_{g0}) \times \prod_{i=0}^{n-1} RP(s_{gi}, s_{gi+1}) \quad (6)$$

Un chemin est sûr, si sa probabilité de but  $RP(path)$  est au-dessus du seuil de sûreté  $\alpha$ .

$$RP(path) \geq \alpha \quad (7)$$

Nous définissons aussi l'utilité d'un chemin  $U(path)$  comme étant la somme des utilités des états but du chemin et le coût d'un chemin  $C(path)$  comme le coût total pour atteindre tous les états but du chemin.

$$U(path) = \sum_{i=0}^n u(s_{gi}) \quad (8)$$

$$C(path) = V(s_0, s_{g0}) + \sum_{i=0}^{n-1} V(s_{gi}, s_{gi+1}) \quad (9)$$

Le chemin optimal recherché, est un chemin sûr, maximisant  $U$  en premier critère et minimisant  $C$  en second. Ce second critère est utilisé lorsque plusieurs chemins sont sûrs et d'utilités équivalentes. Ce qui est fréquemment le cas quand plusieurs chemins sont composés des mêmes états but dans un ordre différent. Dans cette situation, le critère de coût permet de choisir l'ordre le moins coûteux. Enfin, nous appelons la sélection des buts  $list\_g$ , qui est le chemin optimal moins l'état initial. L'algorithme de recherche du chemin optimal est donnée dans la partie 4.

$$list\_g = (s_{g0}, s_{g1}, \dots, s_{gn}) \text{ s.t. } s_{gi} \in S_g$$

$$list\_g = SGGraph(s_0) \quad (10)$$

### 3.4 Un petit exemple

Nous illustrons l'utilisation de notre heuristique de sélection des meilleurs buts via la version modifiée du problème d'exploration, que nous avons présenté à la section 2.1.

**Relaxation du problème.** L'espace d'état est réduit à la position du robot et les actions à ses déplacements possibles dans les 4 directions cardinales (i.e. problème de type *grid world*).

**Probabilité de but et coût de but.** Dans la Figure 1,  $s_0$  est la position initiale du robot,  $s_{gi} (i \in 1, 2, \dots, 6)$  sont les points d'intérêts à visiter (buts) étiquetés de leurs scores. Les étiquettes des arcs correspondent respectivement, à la distance pour aller d'un point à un autre (coût) et à la probabilité d'atteindre un point depuis un autre.

**Sélection des buts.** Dans la Figure 1, il y a 2 chemins possibles :  $path\_a = (s_{g3}, s_{g2}, s_{g6})$  et  $path\_b = (s_{g4}, s_{g6})$ . En utilisant l'Eq.6 et l'Eq.8, nous obtenons :

$$\begin{aligned} - RP(path\_a) &= 0.84 \times 0.9 \times 1 = 0.756 & U(path\_a) &= 6 + 4 + 5 = 15. \\ - RP(path\_b) &= 0.85 \times 0.92 = 0.782 & U(path\_b) &= 2 + 5 = 7. \end{aligned}$$

En fixant  $\alpha = 0.75$ , les deux chemins  $path\_a$  et  $path\_b$  sont sûrs ( $RP(path\_a) \geq \alpha$  et  $RP(path\_b) \geq \alpha$ ). Cependant, le meilleur chemin est  $path\_a$  car de qualité supérieure ( $U(path\_a) > U(path\_b)$ ). Notons que le coût n'intervient pas dans cet exemple pour des raisons de simplicité, mais aurait été utilisé si les deux chemins étaient de même qualité. Enfin, nous extrayons depuis le  $path\_a$ , la sélection des buts, soit  $list\_g = (s_{g3}, s_{g2}, s_{g6})$ .

## 4 Algorithmes

Dans cette partie, nous commençons par présenter l'algorithme *Value Iteration* (VI) utilisé pour approximer nos deux fonctions de valeur (probabilité de but et coût de but). Nous détaillons ensuite notre méthode pour résoudre le *SGGraph* et finissons par présenter la solution globale du problème.

### 4.1 Algorithme VI

---

#### Algorithm 1 Relaxed Plan Value Iteration

---

```

1: function RPVALUEITERATION
2:   InitializeStates() ▷ see Eq.11
3:   repeat
4:      $\Delta_{RP} = 0, \Delta_V = 0$ 
5:     for each  $s \in \mathcal{S}$  do
6:       for each  $s_g \in \mathcal{S}_G$  do ▷ for all goal states
7:          $safe\_a = \operatorname{argmax}_{a \in A(s)} \sigma(s, a, s_g)$ 
8:          $V(s, s_g) = Q(s, safe\_a, s_g)$ 
9:          $RP(s, s_g) = \sigma(s, safe\_a, s_g)$ 
10:         $\Delta_V = \max(\Delta_V, R_V(s, s_g))$ 
11:         $\Delta_{RP} = \max(\Delta_{RP}, R_{RP}(s, s_g))$ 
12:   until  $\Delta_V < \epsilon$  &  $\Delta_{RP} < \delta$  ▷ see Eq.12

```

---

**Initialisation.** La fonction coût de but  $V$  est initialisée pour chaque état arbitrairement avec une valeur positive. La fonction probabilité de but  $RP$  est initialisée à 1 pour chaque état but et 0 sinon. Plus formellement :

$$\begin{aligned} \forall s_i, s_j \in S \text{ and } \forall s_g \in S_G, \\ V(s_i, s_j) \geq 0, RP(s_i, s_g) = 1 \text{ else } RP(s_i, s_j) = 0 \end{aligned} \quad (11)$$

**Condition d'arrêt.** L'algorithme VI s'arrête quand le résidu sur l'équation de Bellman entre deux itérations est suffisamment petit pour tous les états. Soit, quand le résidu sur Eq.2 et Eq.4 est suffisamment petit pour tous  $s$  et  $s_g$ .

$$\begin{aligned} R_V(s, s_g) &= |V(s, s_g) - V'(s, s_g)| \\ R_{RP}(s, s_g) &= |RP(s, s_g) - RP'(s, s_g)| \\ \forall s \in S, \forall s_g \in S_G, R_V(s, s_g) &< \epsilon, R_{RP}(s, s_g) < \delta \end{aligned} \quad (12)$$

**Calcul des fonctions de valeur.** Dans l'algorithme 1, les fonctions de valeur sont approximées en calculant itérativement Eq.2 et Eq.4 jusqu'à la condition d'arrêt définie plus haut. Pour trouver des solutions sûres, ces deux fonctions sont calculées en choisissant à chaque fois l'action la plus sûre, soit l'action maximisant la probabilité de but (ligne 7). Lignes 8 et 9,  $V$  et  $RP$  sont calculées avec cette action, puis le résidu maximal est stocké lignes 10 et 11. L'algorithme s'arrête quand  $\Delta_V$  et  $\Delta_{RP}$  sont suffisamment petits (nous discutons dans nos tests des valeurs de  $\epsilon$  et  $\delta$ ). La notation  $\sigma$  et  $Q$  est utilisée par souci d'espace, mais  $\sigma$  est à remplacer par la partie droite de Eq.3 et  $Q$  par celle de Eq.5.

## 4.2 Résoudre SGGraph

Sur le graphe  $SGGraph$ , le processus de sélection des buts consiste à trouver un chemin sans boucle et optimal (voir partie 3.3 pour la définition d'un chemin optimal). Dans un premier temps, nous simplifions le graphe en retirant tout élément ne pouvant pas faire partie d'une solution valide. C'est à dire tous les arcs dont la probabilité de but est inférieure au seuil  $\alpha$ , ainsi que tous les nœuds non reliés à l'état initial  $s_0$  (voir éléments en pointillés sur la Figure 1). Une fois le graphe réduit, l'étape suivante est la recherche du chemin optimal. La complexité en temps de la recherche d'un chemin dans un graphe orienté avec cycle est exponentielle. Par conséquent, des méthodes de recherche heuristique doivent être utilisées. Nous faisons le choix d'une recherche en faisceau, qui permet d'obtenir une bonne solution en un temps raisonnable (c'est cette même méthode qui a été utilisée dans d'autres travaux similaires [14, 4]). Ce type de recherche n'explore qu'un ensemble limité de fils à chaque niveau, réduisant ainsi la mémoire et le temps d'exécution nécessaire. Le nombre de fils exploré à chaque étape est paramétrable et correspond à la largeur du faisceau, que nous dénotons  $k$ . L'utilisation d'une heuristique permet de choisir quels sont potentiellement les meilleurs

fils à explorer. Dans notre cas, nous avons choisi une heuristique sélectionnant les fils ayant la meilleure probabilité de but direct, c'est à dire la probabilité de but entre l'état courant et le fil en question. Bien que la solution optimale ne soit pas garantie, car la recherche n'est pas complète, une bonne heuristique permet de trouver rapidement une bonne solution. Nous appliquons cet algorithme sur une recherche en profondeur, et évitons les cycles en marquant chaque nœud quand on rentre dans la récursion et en retirant la marque quand on en sort. Pour chaque chemin sûr rencontré, nous comparons l'utilité  $U$  (Eq.8) et si besoin le coût  $C$  (Eq.9) avec l'actuel meilleur chemin, puis conservons le meilleur des deux.

## 4.3 Planifier les buts sélectionnés

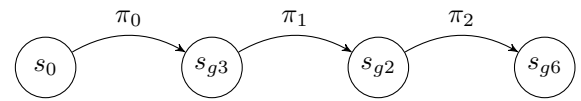


FIGURE 2 – La politique finale  $\pi = (\pi_0, \pi_1, \pi_2)$  est l'agrégation des politiques trouvées pour résoudre chaque but de la sélection.

Étant dans un contexte de planification probabiliste, nous ne pouvons pas simplement atteindre un par un les buts de la sélection, nous devons donc définir une politique pour chaque état but. L'idée est de planifier tous les états but  $s_{gi}$  de la sélection, comme des sous-problèmes indépendants, puis d'agréger les sous-solutions pour former la solution finale (cf. Figure 2). La définition de chaque sous-problème est la même que le problème initial en terme d'actions, états, fonction de transitions et coûts des actions. Elle diffère par contre pour l'état initial ainsi que les buts. L'état initial d'un sous-problème et l'état final du sous-problème précédent. Les buts  $G_i$  du sous-problème  $i$  sont l'ensemble des prédicats de l'état but  $s_{gi}$  correspondant à ce sous-problème. Nous résolvons chaque sous-problème avec un algorithme VI, qui estime  $V_i(s)$  et  $RP_i(s)$ , respectivement le coût et la probabilité d'atteindre un état satisfaisant  $G_i$ , avec la même approche que celle présenté précédemment sur le problème relaxé. Nous déterminons ensuite la politique  $\pi_i$  choisissant les actions qui maximisent d'abord la probabilité de but, puis le coût en cas d'égalité. Pour cela, nous définissons  $SA(s)$  l'ensemble des actions maximisant la probabilité d'atteindre  $G_i$  et  $\pi_i$  tel que :

$$\begin{aligned} SA(s) &= \operatorname{argmax}_{a \in A(s)} \left\{ \sum_{s'} p(s, a, s') RP_i(s') \right\} \\ \pi_i(s) &= \operatorname{argmin}_{a \in SA(s)} \left\{ c(a, s) + \sum_{s'} p(s, a, s') V_i(s') \right\} \end{aligned} \quad (13)$$

La solution du problème s'exprime sous la forme  $\pi(s) = (\pi_0, \pi_1, \dots, \pi_n)$ . Chaque politique est utilisée l'une après l'autre pour atteindre chaque but dans l'ordre de la sélection. Techniquement, les buts déjà atteints sont stockés en mémoire pour connaître à chaque moment la prochaine politique à utiliser.

## 5 Résultats numériques

### 5.1 Cas d'étude

Dans cette partie, nous présentons les résultats numériques que nous avons obtenus en appliquant notre heuristique sur la version simplifiée du problème d'exploration planétaire présentée précédemment. Ce problème, proche d'un problème de type *grid world*, permet d'illustrer nos travaux. Il est composé d'une grille pour modéliser la planète, dont chaque case est une position possible du robot. Les buts du problème sont des cases représentant les sites d'intérêts que le robot doit atteindre. Comme présenté dans la section 2.1, nous ajoutons la notion de risque à ce problème en distinguant deux types de case : les cases sûres et les cases risquées. L'idée est qu'à chaque fois que le robot part d'une case risquée, il a une certaine probabilité de rester bloqué. Le robot peut effectuer quatre actions : aller vers le nord, l'est, le sud ou l'ouest. Sur une case sûre, une action  $a \in A$  déplacera le robot sur la case désirée avec une probabilité  $p$ , ou sur l'une des trois autres directions avec une probabilité  $(1-p)/3$  si l'action échoue (définition classique d'un problème de type *grid world* en environnement probabiliste). Sur une case risquée, le robot a une probabilité  $q$  de rester bloqué, et donc  $1-q$  de se déplacer. L'objectif pour le robot est de visiter les sites avec les scores les plus élevés, tout en garantissant une probabilité de succès supérieur au seuil de sûreté. Une simulation est réussie si le robot atteint tous les sites sélectionnés sans rester bloqué. Dans un premier temps, nous évaluons pour différents seuils, la qualité des sélections des buts obtenue par l'heuristique, puis nous effectuons quelques tests de performance.

**Sélection des paramètres de condition d'arrêt.** Nous rappelons que  $\epsilon$  et  $\delta$ , sont les paramètres utilisés pour contrôler le résidu sur l'équation de Bellman, et donc la précision des estimations sur les fonctions de valeur. Pour choisir les valeurs les plus appropriées, nous avons mesuré la sensibilité de nos résultats pour des valeurs allant de 0.1 à 0.0001. En effectuant des tests sur des cas nominaux et extrêmes, nos résultats convergeaient vers les mêmes estimations à partir de 0.01 pour les deux paramètres, c'est donc la valeur que nous avons choisi.

### 5.2 Résultats

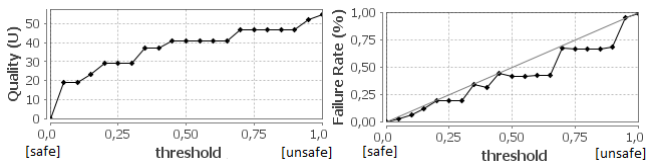


FIGURE 3 – Test sur la sélection des buts. La qualité ( $U$ ) est la somme des scores des buts. Le taux d'échec, et le pourcentage de simulations terminant sans avoir atteint les buts prévus par la sélection.

Dans nos tests présentés dans la Figure 3, nous avons utilisé une grille de  $25 \times 25$ , et différents seuils de risque  $\beta$

entre 0 et 1. Nous utilisons  $\beta$  au lieu de  $\alpha$  pour des raisons de simplicité de raisonnement, ce qui représente ainsi une limite de risque à ne pas dépasser. Nous avons défini dix buts générés de manière aléatoire, et 50% de cases risquées avec une probabilité  $q = 0.1$  de rester bloqué. Nous avons d'abord testé la sélection des buts pour différents seuils puis nous avons simulé chaque sélection pour vérifier le bon respect du seuil. Sur le graphe de gauche de la Figure 3, l'axe  $x$  représente le seuil de risque et l'axe  $y$  la qualité  $U$  obtenue en sommant les scores des buts de la sélection. On peut voir que la qualité croît de façon monotone quand le risque croît. Dans certains cas (par exemple pour  $\beta \in [0.45, 0.65]$ ), la qualité reste constante, signifiant qu'aucune autre meilleure sélection des buts est possible jusqu'à un certain seuil. Dans le deuxième graphe, nous testons via des simulations que chaque sélection des buts générée respecte bien le seuil de risque imposé. Pour ce faire, nous simulons chaque planification de la sélection 1000 fois (nous avons obtenu la même précision en testant avec 2000 simulations), et calculons le taux moyen d'échec, soit le taux de simulation n'ayant pu atteindre tous les buts prévus par la sélection. Comme prévu, le taux d'échec n'excède jamais le seuil  $\beta$  (la ligne diagonale). Toutefois, on repère parfois des intervalles relativement importants entre cette ligne et le taux d'échec. Ceci est dû au fait que le planificateur détecte qu'il ne peut ni augmenter la qualité, ni baisser le coût, avec le risque supplémentaire qui lui est alloué. Il maintient donc la probabilité de but aussi haute que possible.

### 5.3 Performances

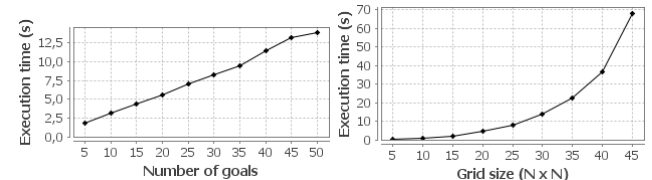


FIGURE 4 – Test de performances : sur la gauche, nous faisons varier le nombre de buts sur un espace d'état constant (grille :  $25 \times 25$ ). À l'inverse, sur la droite, c'est l'espace d'état qui varie et le nombre de buts qui est fixe (8).

Dans la Figure 4, nous traçons le temps d'exécution en seconde du processus de sélection des buts en fonction du nombre de buts total. Nous avons remarqué que ce temps d'exécution est en réalité très proche du temps d'exécution de l'algorithme 1 qui estime les fonctions de valeur. En effet, construire le graphe *SGGraph* et rechercher le meilleur chemin ne prend pas plus de 2% du temps total d'exécution de l'heuristique. Ceci à condition d'avoir une recherche en faisceau efficace. Dans notre cas nous avons testé plusieurs largeurs de faisceau et conclu qu'en moyenne, l'heuristique converge vers une solution unique avec une largeur de faisceau de  $0.25 \times$  le nombre total de buts. Avec cette largeur de faisceau, nous avons obtenu une recherche toujours inférieure à 250 ms pour une grille al-

lant jusque  $45 \times 45$  états et 50 buts.

## 6 Conclusion

Beaucoup de problèmes de planification probabilistes en condition réelle produisent des solutions non sûres, ce qui peut être critiques dans certains domaines tels que le domaine spatial (un robot explorateur peut coûter très cher, etc.). Dans ce papier, nous avons décrit une heuristique pour trouver des solutions satisfaisant les meilleurs buts possibles, tout en limitant la prise de risque. Notre méthode sélectionne les meilleurs buts qui pourront être atteints avec une probabilité de succès supérieure à un seuil donné. Cette sélection s'appuie sur un algorithme de type *Value Itération* qui estime la probabilité de but ainsi que le coût entre chaque but à l'aide d'une version relaxée du problème.

Nous travaux ont soulevé de nouveaux problèmes : une prochaine étape serait de travailler sur un choix autonome du seuil par le planificateur. Une solution envisagée serait d'avoir un seuil qui repose sur le critère de Hurwitz, permettant ainsi d'avoir un compromis variable entre une attitude sûre et une attitude risquée. De plus, l'algorithme VI, qui explore tout l'espace d'état, prend beaucoup de temps. Nous prévoyons donc d'appliquer notre méthode de sélection sur des algorithmes dit *Anytime* tels que RTDP ou L-RTDP qui peuvent converger beaucoup plus rapidement vers des solutions presque optimales. Et enfin, nous prévoyons d'appliquer notre heuristique sur des domaines plus complexes et larges, tels que la planification de situations critiques en environnement virtuel pour la formation de médecins.

## Références

- [1] R. Bellman. *Dynamic Programming*. Princeton University Press, 1957.
- [2] J. Benton, Minh Do, and S. Kambhampati. Anytime heuristic search for partial satisfaction planning. *Artificial Intelligence*, 173 :562–592, 2009.
- [3] DP. Bertsekas. *Dynamic Programming and Optimal Control 3rd Edition, Volume II Chapter 6 Approximate Dynamic Programming 6 Approximate Dynamic Programming*. 2011.
- [4] A. García-Olaya, T. De La Rosa, and D. Borrajo. Using the relaxed plan heuristic to select goals in oversubscription planning problems. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 7023 LNAI, pages 183–192, 2011.
- [5] P. Geibel. Reinforcement learning with bounded risk. *Proceedings of the Eighteenth International Conference on Machine Learning*, 9(D) :162–169, 2001.
- [6] M. Ghallab, DS. Nau, and P. Traverso. *Automated planning and acting*. Cambridge University Press, 2016.
- [7] BL. Golden, L. Levy, and R. Vohra. The Orienteering Problem. 1987.
- [8] M. Helmert, P. Haslum, J. Hoffmann, and R. Nissim. Merge-and-Shrink Abstraction : A Method for Generating Lower Bounds in Factored State Spaces. *Journal of ACM*, 61(3) :63, 2014.
- [9] M. Katz and D. Carmel. Implicit abstraction heuristics. *Journal of Artificial Intelligence Research*, 39 :51–126, 2010.
- [10] LA. Kramer and SF. Smith. Maximizing Flexibility : A Retraction Heuristic for Oversubscribed Scheduling Problems. *International Conference on Automated Planning and Scheduling*, 2005.
- [11] N. Meuleau, R. Brafman, and E. Benazera. Stochastic Over-Subscription Planning Using Hierarchies of MDPs. *International Conference on Automated Planning and Scheduling*, pages 121–130, 2006.
- [12] Groupe PDMIA. Processus décisionnels de Markov en intelligence artificielle. *Édité par Olivier Bu et Olivier Sigaud*, 1, 2008.
- [13] R. Sanchez-Nigenda and S. Kambhampati. Planning Graph Heuristics for Selecting Objectives in Over-subscription Planning Problems. In *Proceedings of the 15th International Conference on Automated Planning and Scheduling*, pages 192–201, 2005.
- [14] D. Smith. Choosing Objectives in Over-Subscription Planning. *International Conference on Automated Planning and Scheduling*, pages 393–401, 2004.
- [15] M. Steinmetz, J. Hoffmann, and O. Buffet. Revisiting Goal Probability Analysis in Probabilistic Planning. *International Conference on Automated Planning and Scheduling*, (0) :299–307, 2016.
- [16] RS. Sutton and AG Barto. Reinforcement learning. *Learning*, 3(9) :322, 2012.
- [17] F. Teichteil-Königsbuch. Problèmes de Plus Sûr et Plus Court Chemin Stochastique.
- [18] F. Teichteil-Königsbuch. Stochastic Safest and Shortest Path Problems. *Association for the Advancement of Artificial Intelligence*, pages 1825–1831, 2012.
- [19] M. van den Briel, R. Sanchez, and S. Kambhampati. Over-subscription in Planning : A Partial Satisfaction Problem. *ICAPS 2004 Workshop on Integrating Planning into Scheduling*, 2004.
- [20] P. Vansteenwegen, W. Souffriau, and DV. Oudheusden. The orienteering problem : A survey. *European Journal of Operational Research*, 209 :1–10, 2011.

# Empirical evaluation of a Q-Learning Algorithm for Model-free Autonomous Soaring

Erwan Lecarpentier<sup>1</sup>Sebastian Rapp<sup>2</sup>

Marc Melo

Emmanuel Rachelson<sup>3</sup><sup>1</sup> ONERA – DTIS (Traitement de l’Information et Systèmes)

erwan.lecarpentier@isae.fr

<sup>2</sup> TU Delft – Department of Aerodynamics, Wind Energy & Propulsion

s.rapp@tudelft.nl

<sup>3</sup> ISAE Supaero – DISC (Département d’Ingénierie des Systèmes Complexes)

emmanuel.rachelson@isae.fr

## Abstract

*Autonomous unpowered flight is a challenge for control and guidance systems : all the energy the aircraft might use during flight has to be harvested directly from the atmosphere. We investigate the design of an algorithm that optimizes the closed-loop control of a glider’s bank and sideslip angles, while flying in the lower convective layer of the atmosphere in order to increase its mission endurance. Using a Reinforcement Learning approach, we demonstrate the possibility for real-time adaptation of the glider’s behaviour to the time-varying and noisy conditions associated with thermal soaring flight. Our approach is online, data-based and model-free, hence avoids the pitfalls of aerological and aircraft modelling and allow us to deal with uncertainties and non-stationarity. Additionally, we put a particular emphasis on keeping low computational requirements in order to make on-board execution feasible. This article presents the stochastic, time-dependent aerological model used for simulation, together with a standard aircraft model. Then we introduce an adaptation of a Q-learning algorithm and demonstrate its ability to control the aircraft and improve its endurance by exploiting updrafts in non-stationary scenarios.*

## Keywords

Reinforcement Learning control, Adaptive control applications, Adaptation and learning in physical agents, UAVs.

## 1 Introduction

The number of both civil and military applications of small unmanned aerial vehicles (UAVs) has augmented during the past few years. However, as the complexity of their tasks is increasing, extending the range and flight duration of UAVs becomes a key issue. Since the size, and thus the energy storage capacity, is a crucial limiting factor, other means to increase the flight duration have to be examined.

A promising alternative is the use of atmospheric energy in the form of gusts and updrafts. This could significantly augment the mission duration while simultaneously save fuel or electrical energy. For this reason, there is a great interest in the development of algorithms that optimize the trajectories of soaring UAVs by harvesting the energy of the atmosphere. Since the atmospheric conditions are changing over time, it is crucial to develop an algorithm able to find an optimal compromise between exploring and exploiting convective thermal regions, while constantly adapting itself to the changing environment.

In this work we adapt a  $Q$ -learning (Watkins & Dayan, 1992) algorithm for this task. Our method is model-free, therefore suitable for a large range of environments and aircraft. Additionally, it does not need pre-optimization or pre-training, works in real-time, and can be applied online. Although the gap towards a fully autonomous physical demonstrator has not been bridged yet, our main contribution in this work is the *proof of concept* that a model-free reinforcement learning approach can efficiently enhance a glider’s endurance. We start by reviewing the state of the art in UAV static soaring and thermal modelling in Section 2 and position our contributions within previous related work. Then, in Section 3, we present the specific atmospheric model we used and its improvements over previous contributions, along with the thermals scenario used in later experiments. Section 4 details the aircraft dynamics model. We introduce our implementation of the  $Q$ -learning algorithm in Section 5 and discuss its strengths, weaknesses and specific features. Simulation results are presented in Section 6. We finally discuss the limitations of our approach and conclude in Section 7.

## 2 Related work

During the last decade, several possibilities to efficiently utilize atmospheric energy for soaring aircraft have been proposed. For a general introduction to static and dynamic soaring, refer to Chen & McMasters (1981) for instance.

For a more specific review on thermal centring and soaring in practice, see Reichmann (1993).

Most approaches to thermal soaring rely on the identification of some model of the wind field surrounding the aircraft. This estimated wind field is then used to track an optimized trajectory inside the thermal or between thermals, using various methods for identification and path planning (Allen, 2005; Allen & Lin, 2007; Lawrance & Sukkarieh, 2011; Lawrance, 2011; Bencatel *et al.*, 2013; Chen & Clarke, 2011; Chakrabarty & Langelaan, 2010). Such approaches demonstrated important energy savings (up to 90% in simulation (Chakrabarty & Langelaan, 2010)) compared to conventional flight. An alternative robust control algorithm (Kahveci & Mirmirani, 2008), based again on a pre-identification of a thermal model showed good results also.

In this paper, we reconsider the possibility to use a *Reinforcement Learning* (RL, Sutton & Barto, 1998) approach to optimize the trajectory. Using RL to exploit thermals has already been examined by Wharington (1998). In this work, a neural-based thermal centre locator for the optimal autonomous exploitation of the thermals is developed. After each completed circle, the algorithm memorizes the heading where the lift was the strongest and moves the circling trajectory towards the lift. However, this thermal locator is too time consuming for real-time on-board applications.

We introduce a *Q-learning* algorithm using a *linear function approximation*, which is simple to implement, demands less computational resources and does not rely on the identification of a thermal model. We empirically evaluate this online learning algorithm (Section 5) by interfacing it with a simulation model that couples the aircraft dynamics (Section 4) with an improved local aerological model (Section 3). We use the model to test our algorithm in several scenarios and show that it yields a significant endurance improvement. Our algorithm's main feature lies in its complete independence of the characteristics of the aerological environment, which makes it robust against model inaccuracy and estimation noise. Moreover, not explicitly estimating the thermal centre position and updraft magnitude saves valuable computational time.

### 3 Atmospheric model

Our updraft model expands on that of Allen (2006). His model possesses three desirable features : dependence of the updraft distribution in the vertical direction, explicit modelling of downdrafts at the thermal's border and at every altitude, and finally the use of an environmental sink rate to ensure conservation of mass. Although a complete literature review on modelling the convective boundary layer is beyond the scope of this paper, it should be noted that Allen (2006) is the first reference that includes these three modelling aspects.

We describe a thermal updraft as a symmetrical, bell-

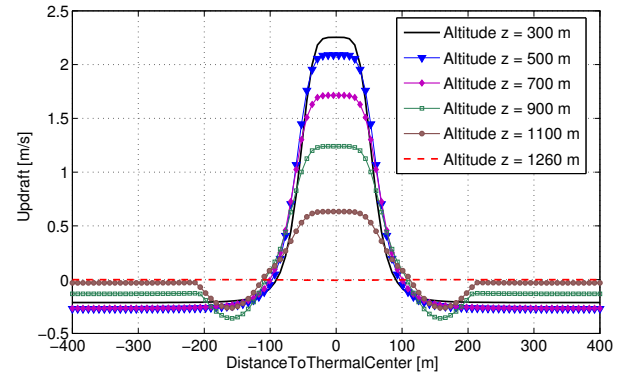


FIGURE 1 – Updraft distribution with altitude

shaped distribution as illustrated in Figure 1. This distribution is characterized by two radii  $r_1$  and  $r_2$ . At a given altitude  $z$ , if  $r$  denotes the distance to the thermal center, for  $r < r_1$  the updraft has a quasi-constant value of  $w_{peak}$ , then for  $r_1 < r < r_2$  this value drops smoothly to zero, and between  $r_2$  and  $2r_2$  appears a downdraft. The thermal has no influence further than  $2r_2$ .

The maximum updraft velocity  $w_{peak}$  evolves altitude-wise proportionally to  $w^* \left( \frac{z}{z_i} \right)^{\frac{1}{3}} \left( 1 - 1.1 \frac{z}{z_i} \right)$ , where  $w^*$  is an average updraft velocity and  $z_i$  is a scaling factor indicating the convective boundary layer thickness. Above  $0.9z_i$  all velocities are assumed to be zero.

Finally, based on the conservation of mass principle, an altitude-dependent global environmental sink rate is calculated and applied everywhere outside the thermals. For specific equations, we refer the reader to Allen (2006). We introduce three additional features that bring our simulation model closer to a real-life description, namely thermal drift, life-cycle and noise. First, in order to account for local winds, we let the thermals drift in the horizontal plane with a velocity  $(\bar{v}_x, \bar{v}_z)$ . Usually, the root point of a thermal is a fixed location and the thermal leans with the wind, so introducing a thermal drift is a poor description of this phenomenon. Nevertheless, for our simulations, it approximates the practical phenomenon of drift given that the aircraft model is reduced to a single point-mass. Thermals also have a finite life. We decompose a thermal's life in a latency phase of duration  $t_{off}$  and a growth, maturity and fade-off phase of duration  $t_{life}$ . After  $t_{off} + t_{life}$  the thermal dies. The life-cycle of a thermal is described by the updraft coefficient  $c_\xi(t)$  shown in Figure 2, using a shape parameter  $\xi$ . This  $c_\xi(t)$  coefficient is used as a multiplier on the total updraft. Finally, it is well-known among cross-country pilots that thermals are rarely round and present a great variety of shapes and much noise. In order to account for this fact and to model real-life uncertainties we added a Gaussian distributed noise  $n$  to the wind velocity.

We maintain a constant number  $N$  of thermals in the



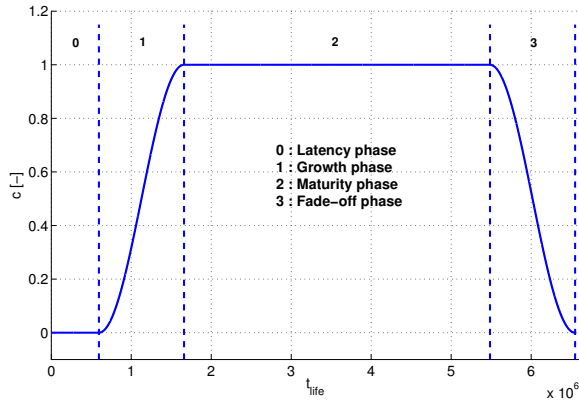


FIGURE 2 – Evolution of the updraft coefficient  $c_{\xi}(t)$

flight area, although some might be in their latency phase. Consequently, whenever a thermal dies, a new thermal is generated with randomly drawn parameters  $\{x_{th}, y_{th}, w^*, z_i, \bar{v}_x, \bar{v}_y, t_{off}, t_{life}, \xi\}$ .

## 4 Aircraft model

To model the dynamical behaviour of our aircraft, we used the equations derived by Beeler *et al.* (2003), which consider the aircraft as a point-mass, 6 degrees of freedom system, and take into account the three dimensional wind velocity vector of the atmosphere as well as a parametric model for the aircraft's aerodynamics. Let  $m$  be the glider's mass and  $g$  the gravity acceleration. The used variables are :

- $x, y, z$  the coordinates in the earth frame ;
- $V$  the magnitude of the aircraft's velocity in the earth frame ;
- $\gamma$  the angle of climb ;
- $\chi$  the course angle ;
- $\alpha$  the angle of attack ;
- $\beta$  the sideslip angle ;
- $\mu$  the bank angle ;
- $L, D$  and  $C$  the lift, drag and lateral force.

The corresponding equations are described below :

$$\begin{aligned} \dot{x} &= V \cos(\chi) \cos(\gamma) & \dot{z} &= V \sin(\gamma) \\ \dot{y} &= V \sin(\chi) \cos(\gamma) & \dot{V} &= -D/m - g \sin(\gamma) \\ \dot{\gamma} &= \frac{1}{mV} \left( L \cos(\mu) + C \sin(\mu) - \frac{g}{V} \cos(\gamma) \right) \\ \dot{\chi} &= \frac{1}{mV \cos(\gamma)} (L \sin(\mu) - C \cos(\mu)) \end{aligned}$$

The first three equations describe the kinematics and position rates in the earth frame. The last three equations define the dynamics of the glider aircraft. For a detailed presentation of the aerodynamic parameters and forces, we refer the reader to Beeler *et al.* (2003). Adopting this modelling directly implies taking the three angles  $\alpha$ ,  $\beta$  and

$\mu$  as control variables. Indeed the lift force depends on the bank angle, while the drag and lateral force depend on the three angles. For simplicity of notations we omitted to write these dependencies in the model's equations. The choice of the state and action spaces considered by the controller is discussed in Section 5.2.

## 5 Adaptive controller

### 5.1 Q-learning

RL (Sutton & Barto, 1998) is a branch of Discrete-time Stochastic Optimal Control that aims at designing optimal controllers for non-linear, noisy systems, using only interaction data and no *a priori* model. The only hypothesis underlying RL algorithms is that the system to control can be modelled as a Markov Decision Process (MDP, Puterman, 2005), even if this model is not available. An MDP is defined by a state space  $\mathcal{S}$ , an action space  $\mathcal{A}$ , a discrete-time transition function  $p(s'|s, a)$  denoting the probability of reaching state  $s'$  given that action  $a$  was undertaken in state  $s$ , and a reward model  $r(s, a, s')$  indicating how valuable the  $(s, a, s')$  transition was with respect to some predefined criteria.

The overall goal of an RL algorithm is to derive an optimal control policy  $\pi^*(s) = a$  that maximizes the expected cumulative sum of rewards  $\mathbb{E}(\sum_{t=0}^{\infty} \eta^t r_t)$  from any starting state  $s$  ( $\eta \in [0; 1[$  being a discount factor over future rewards). We focus on model-free RL algorithms that do not commit to the knowledge of the transition and reward models of the underlying MDP but use *samples* of the form  $(s, a, r, s')$  to learn an optimal policy. In our case, that means that an RL algorithm controlling the glider with an overall goal of gaining energy will use sensor data to build  $\pi^*$  online, without relying on a model of the atmosphere, or the aircraft's flight dynamics.

Q-learning, introduced by Watkins & Dayan (1992), is one of the most simple and popular online RL algorithms. It aims at estimating the optimal action-value function  $Q^*(s, a)$  in order to asymptotically act optimally. This function denotes the expected gain of applying action  $a$  from state  $s$ , and then applying an optimal control policy  $\pi^*$  :

$$Q^*(s, a) = \mathbb{E} \left( \sum_{t=0}^{\infty} \eta^t r_t \mid s_0 = s, a_0 = a, a_t = \pi^*(s_t) \right)$$

The key idea behind Q-learning is that the optimal action in state  $s$  is the one that maximizes  $Q^*(s, a)$ . Thus the optimal policy is greedy with respect to  $Q^*$  in every state. Estimating  $Q^*$  from  $(s, a, r, s')$  samples is a stochastic approximation problem which can be solved with a procedure known as *temporal differences*. The Q-learning algorithm is summarized in Algorithm 1.

Notice that Q-learning is an *off-policy* method, that is, it estimates  $Q^*$  assuming that a greedy policy w.r.t.  $Q$  is followed while this is untrue with probability  $\epsilon_t$ . As the latter tends towards zero, if the learnt Q-function has

**Algorithm 1 :  $Q$ -learning**


---

Initialize  $Q(s, a)$  for all  $(s, a) \in S \times A$ ,  
 $s_t \leftarrow s_0$ .  
**repeat**  
  Apply  $a_t = \arg \max_{a \in A} Q(s_t, a)$  with probability  
   $1 - \epsilon_t$ , otherwise apply a random action  $a_t$   
  Observe  $s_{t+1}$  and  $r_t$   
   $\delta_t = r_t + \eta \max_{a' \in A} (Q(s_{t+1}, a')) - Q(s_t, a_t)$   
  Update  $Q(s_t, a_t) \leftarrow Q(s_t, a_t) + \alpha_t \delta_t$   
   $s_t \leftarrow s_{t+1}$   
**until** simulation end

---

converged to  $Q^*$ , the agent tends to act optimally. As long as all state-action pairs are visited infinitely often when  $t \rightarrow \infty$ ,  $Q$  is guaranteed to converge to  $Q^*$  if the sequence of learning rates  $\alpha_t$  satisfies the conditions of Robbins & Monro (1951) :  $\sum_{t=0}^{\infty} \alpha_t = \infty$ ,  $\sum_{t=0}^{\infty} \alpha_t^2 < \infty$ . In the remainder of this section, we discuss how our problem differs from the vanilla MDP and  $Q$ -learning frameworks, and the design choices we made to accommodate these differences.

## 5.2 State and action spaces

Recall that the state of the aircraft, (Section 4), or the state of the atmospheric model (Section 3) are not fully observable. So it would be unrealistic to define the state space  $S$  as the observations of these values. Instead, we suppose that a state only defined by  $(\dot{z}, \dot{\gamma}, \mu, \beta)$  is accessible and that its dynamics still defines an MDP. Such a state is easily measurable with sensors such as pressure sensors, accelerometers or gyrometers. This key assumption is crucial to the success of our method since it reduces the size of the state space, easing the approximation of  $Q^*(s, a)$ . We shall see later that this choice of state variables has other advantages.

The considered actions consist in directly controlling the aircraft's bank and sideslip angles increments so that the action space is  $A = \{-\delta\mu, 0, \delta\mu\} \times \{-\delta\beta, 0, \delta\beta\}$ , resulting in  $|A| = 9$  different possible actions. We chose the values of  $\delta\mu$  and  $\delta\beta$  so that, given a certain control frequency, the cumulated effect of a constant action does not exceed the admissible dynamics of the aircraft. This results in a steady state change, representative of the actual behaviour of the actuators.

## 5.3 Reward model

The goal of our learning algorithm is to maximize the glider's endurance. This boils down to maximizing the expected total energy gain, so we wish that  $Q(s, a) = \mathbb{E}\{\text{total energy at } t = \infty\}$ . To achieve this, we choose :

$$r_t = \dot{E}_{aircraft} = \frac{d}{dt} \left( z + \frac{V^2}{2g} \right) \quad (1)$$

Thus we assume that this reward signal  $r_t$  is provided to the learning algorithm at each time step, representing the

(possibly noisy) total energy rate of the aircraft. Note that the variables  $\dot{z}$ ,  $V$  and  $\dot{V}$  can be measured online with classical sensors such as a GPS and an accelerometer.

## 5.4 Convergence in unsteady environments

The previous requirements on  $\epsilon_t$  and  $\alpha_t$  for convergence of  $Q$  to  $Q^*$  hold if the environment can indeed be modeled as an MDP. However, in the studied case, the environment is non-stationary since the thermals have a time-varying magnitude (thermal coefficient) and location (drift). Moreover, given the choice of state variables, since the agent is blind to its localization, the distribution  $p(s'|s, a)$  is not stationary and changes from a time step to the other. Consequently, our learning agent evolves in a constantly changing environment which is *not* a stationary MDP and we actually need to rely on its ability to learn and adapt quickly to changing conditions if we wish to approximate these conditions as quasi-stationary. In order to allow this quick adaptation, we need to force a permanent exploration of the state-action space and to constantly question the reliability of  $Q$ . This corresponds to making use of constant  $\alpha_t$  and  $\epsilon_t$  values, which need to be well-chosen in order to retain a close-to-optimal behaviour while quickly adapting to the changes in the environment. The choice of a simplified low-dimensional state space makes the adaptation to a non-stationary environment feasible. In fact, with our specific choice of state variables, in the short term, the learning agent observes a quasi-constant state  $(\dot{z}, \dot{\gamma}, \mu, \beta)$  and the optimal action in this state is almost constant as well. Indeed, the chosen variables evolve slowly through the time, making the evolution of the optimal action value slow as well. This allows to make maximal use of the collected samples since only a local approximation around the current state is required to compute the current optimal action. The success of the method is therefore due to the capacity of the  $Q$ -learning algorithm to track the optimal action quickly enough in comparison to the environment's dynamics.

## 5.5 Linear $Q$ -function approximation

In order to avoid the discretisation of the state space in the description of  $Q$ , we adopt a linear function approximation of  $Q(s, a)$ . We introduce sigmoid-normalized versions of the state space variables and define our basis functions  $\phi$  as the monomials of these normalized variables of order zero to two (15 basis functions). Then, by writing  $Q(s, a) = \theta^T \phi(s, a)$ , the update equation of  $Q$ -learning becomes  $\theta_{t+1} = \theta_t + \alpha_t \delta_t \phi(s_t, a_t)$ . There is abundant literature on choice of feature functions in RL, we refer the reader to Parr *et al.* (2008), Hachiya & Sugiyama (2010), or Nguyen *et al.* (2013) for more details.

To summarize, our glider is controlled by a  $Q$ -learning algorithm with fixed learning and exploration rates ( $\alpha$  and  $\epsilon$ ) to account for the unsteadiness of the environment. The optimal action-value function  $Q^*$  is approximated with a linear architecture of quadratic features defined over a set of observation variables  $(\dot{z}, \dot{\gamma}, \mu, \beta)$ . Finally, at each time

step, the chosen action is picked among a set of 9 possible increments on the  $(\mu, \beta)$  current values.

## 6 Simulation results

We identify three scenarios designed to empirically evaluate the convergence rate of the algorithm and the overall behaviour of the glider. These scenarios take place within a 1100m wide circular flight arena. Whenever the glider exits the arena, an autopilot steers it back in. The aircraft is initialized at  $z = 300\text{m}$  and  $V = 15\text{m/s}$ . According to Allen (2006), we set  $w^* = 2.56\text{m/s}$  and  $z_i = 1401\text{m}$ . The algorithm parameters were  $\epsilon = 0.01$ ;  $\alpha = 0.001$ ;  $\eta = 0.99$ ;  $\delta\beta = 0.003\text{ deg}$ ;  $\delta\mu = 0.003\text{ deg}$ ;  $\beta_{max} = 45\text{ deg}$ ;  $\mu_{max} = 25\text{ deg}$  and the observation frequency is  $1\text{kHz}$ .

The three scenarios are the following : flight in still air without thermal but a noisy downdraft; birth of a thermal along the trajectory; death of a thermal into which the UAV was flying. Qualitatively, the optimal policy in each case is respectively to adopt a straight flight configuration; to circle up within the thermal; and to switch from the circular trajectory to a straight one as in the first case. In each scenario, we refer to the optimal action-value function parameters as  $\theta_{opt}$ . In order to analyse the convergence rate of the algorithm, we built an empirical estimate  $\hat{\theta}_{opt}$  of those parameters with the value they take after the convergence of the algorithm and then compute the quantity  $\|\theta_t - \hat{\theta}_{opt}\|_2$  along 50 roll-outs of the system. The convergence results are reported in Figure 3 where the error bars indicate the standard deviation. One can see that the time required to adjust the parameters to each situation ranges between 30 and 40 seconds, which is compatible with the change rates of the glider’s environment. Note in particular that the glider’s behaviour might be optimal long before  $\theta$  converges to  $\theta_{opt}$  since from a certain state  $s$  the optimal action might be selected even if the parameters did not converge. Indeed, what matters is the ranking of the  $Q$ -values of the different actions rather than the  $Q$ -values themselves. Practically, the configurations vary between the three studied cases and the exploratory feature of the  $\epsilon$ -greedy policy allows to permanently adapt the  $Q$ -function to the situation.

The performance reached by the control algorithm can be measured via the total energy of the aircraft, capturing the reached altitude and the velocity. In the three aforementioned scenarios, the expected results are not the same. Indeed, in a steady atmosphere, the optimal policy only allows to minimize the loss of altitude by setting  $\beta = \mu = 0$ . Such a configuration is optimal since no thermals can be found and the glider can only maximize its long term energy by flying straight and avoiding sharp manoeuvres. Then, when a thermal is reached, the algorithm’s exploratory behaviour allows to capture the information that it is worth changing  $\beta$  and  $\mu$ , and adapts the trajectory to maximize the long-term return. In the third situation, when the glider flies inside

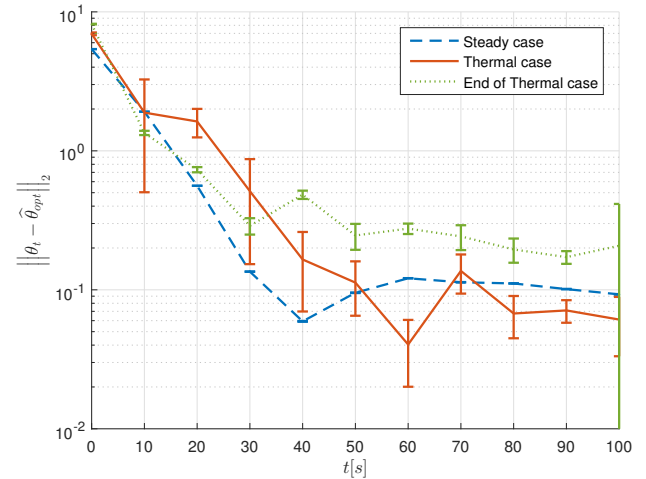


FIGURE 3 – Convergence of the action-value function

a dying thermal, the algorithm brings back the parameters to a steady atmosphere configuration and again minimizes the expected loss of energy.

Figure 4 shows the evolution of altitude and instantaneous rewards through time in a typical long-term scenario with multiple thermal crossings. Each altitude pike shows the entry of the aircraft into a thermal. First the trajectory is bent in order to maximize the altitude gain and when the thermal dies, the glider goes back to the steady flight configuration. Clearly, each gain-of-altitude phase corresponds to a positive reward and, conversely, a loss-of-altitude phase to a negative one. A 3D display of the trajectory inside a thermal is presented in Figure 5.

The  $Q$ -learning controller yields an overall behaviour close to the one of a human pilot while being unaware of its own location and of local wind field models. When flying in still air, the glider remains in “flat” flight attitude, thus maximizing its flight time expectancy. Whenever an updraft is spotted, it engages in a spiral, as shown in Figure 4. If the updraft dies, the aircraft comes back to the first configuration. This results in an overall trajectory composed with straight lines and circles as displayed in Figure 6.

Figure 4 also illustrates the reaction times of the glider and the overall command behaviour. It appears that the glider starts to circle up the thermals long before the value function has converged. Similarly, the convergence to a steady air optimal behaviour is faster than the  $Q$ -function convergence illustrated on Figure 3. When the glider reaches the thermal’s top, the updraft naturally decreases. Consequently one can notice the reduction of the bank angle (enlargement of the turning radius) computed by the algorithm in order to stay in the thermal while reaching a zero vertical velocity.

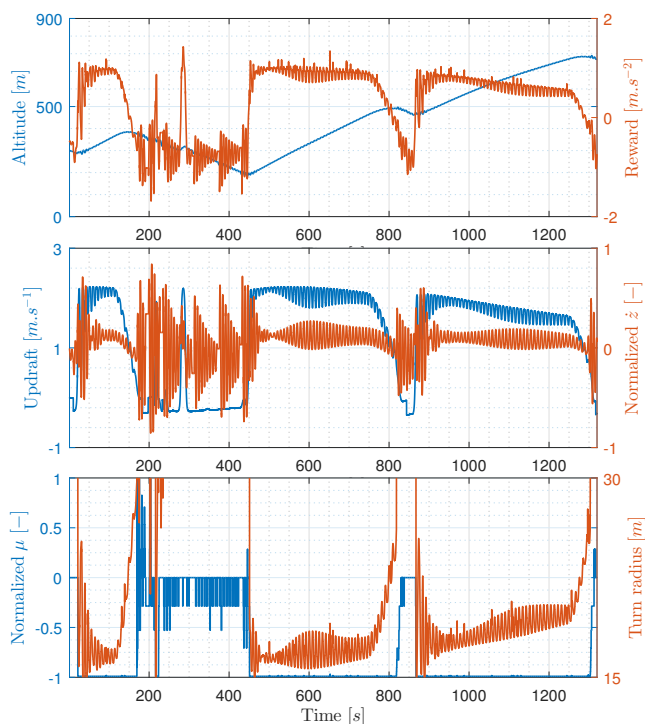


FIGURE 4 – Evolution of the aircraft variables with time

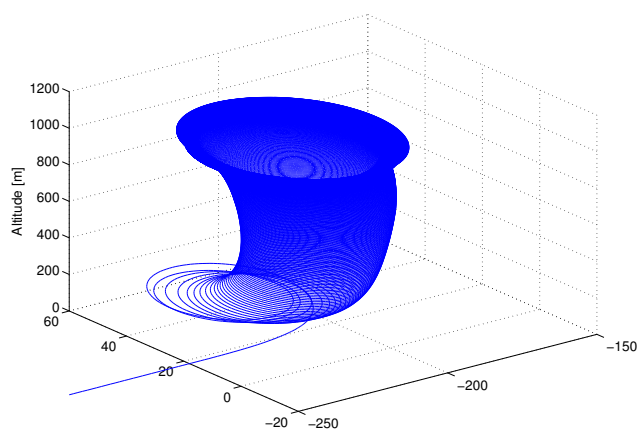


FIGURE 5 – Trajectory of the aircraft inside a thermal

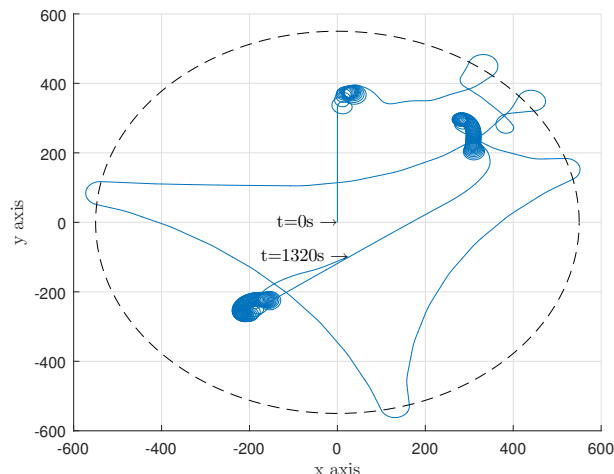


FIGURE 6 – An example of trajectory

## 7 Discussion and conclusion

In this paragraph, we discuss the limitations of our contribution, highlight directions for improvements and underline how our results make a difference compared to related work in the literature presented in Section 2. To summarize, we implemented a proof of concept that a computationally light algorithm like  $Q$ -learning could be adapted to take into account the time-varying conditions of thermal soaring flight and could make efficient online changes to the control behaviour of an autonomous glider. We take a critical look at this contribution.

First of all, we did not introduce a new RL algorithm *per se*, even though we shortly discuss the question of learning in unsteady environments. The choice of  $Q$ -learning is justified by its low computational footprint, despite the existence of a vast literature of efficient algorithms in online RL. Our contributions on the RL side are application-specific : first we justify the need for constant  $\alpha$  and  $\epsilon$  parameters to account for permanent exploration and adaptation in our unsteady environments. Secondly, we make a particular choice of state and action variables, such that, under an optimal policy, the system remains in a quasi-constant state (it would not be the case if the coordinates  $x, y, z$  were part of the state space for instance), thus limiting the need for exploration and making the learning process faster. Finally, we introduced a reward model based explicitly on the maximization of the long term energy of the aircraft, thus linking energetic considerations with the definition of the  $Q$ -function.

From a low-level control point of view, the hypothesis of a control frequency of 1kHz is somehow questionable and it should be decreased in further developments. We argue however that this frequency is representative of a measurement frequency and should thus still be used to update the  $Q$ -function. Exploratory actions artificially

account for the information collected due to the noise in wind conditions felt by the aircraft.

The 6 degrees of freedom aircraft model used in the simulation is a classical flight dynamics model that does not take into account the wind gradient in the wingspan direction. This gradient however is known to be a crucial information for human pilots, since it disambiguates whether a thermal centre is on the left or right hand side of the glider. Exploiting such information could bring more efficiency to the glider's control and avoid missing some thermals because the turn was initiated in the wrong direction.

Lastly, in this proof of concept, we based the action space on the aerodynamic angles  $\mu$  and  $\beta$  as it was done by Beeler *et al.* (2003). Since the  $Q$ -learning algorithm aims at maximizing the average energy gain in the long term, it does not improve the short-term stabilization of the longitudinal modes of the aircraft, leading to the oscillations shown in Figure 4. Even though this does not affect the overall long-term energy gains, a desirable improvement would consist in implementing a low-level stabilization loop (with a PID controller for instance), thus allowing to define the action space using aircraft attitude set points, rather than aerodynamic angles.

Overall, our contribution is three-fold. First we report on how to efficiently adapt a  $Q$ -learning algorithm to the non-steady, partially observable, control problem of thermal soaring. Then we empirically evaluate the performance of this algorithm in a rich simulation environment, illustrating how it can be used to improve the energy autonomy of soaring planes. Finally we discuss the strengths and limitations of this approach, thus opening research perspectives on this topic and providing first insights on these perspectives.

## Références

- ALLEN M. J. (2005). *Autonomous Soaring for Improved Endurance of a Small Uninhabited Air Vehicle*. Rapport interne, NASA Dryden Research Center.
- ALLEN M. J. (2006). *Updraft Model for development of Autonomous Soaring Uninhabited Air Vehicles*. Rapport interne, NASA Dryden Flight Research Center.
- ALLEN M. J. & LIN V. (2007). *Guidance and Control of an Autonomous Soaring UAV*. Rapport interne, NASA Dryden Flight Research Center.
- BEELER S., MOERDER D. & COX D. (2003). *A Flight Dynamics Model for a Small Glider in Ambient Winds*. Rapport interne, NASA.
- BENCATEL R., DE SOUSA J. T. & GIRARD A. (2013). Atmospheric flow field models applicable for aircraft endurance extension. *Prog. in Aerospace Sciences*, **61**.
- CHAKRABARTY A. & LANGELAAN J. (2010). Flight path planning for UAV atmospheric energy harvesting using heuristic search. In *AIAA Guidance, Navigation, and Control Conference*.
- CHEN M. & MCMASTERS J. (1981). From paleoaeronautics to altostratus - a technical history of soaring. In *AIAA Aircraft Systems and Technology Conference*.
- CHEN W. & CLARKE J. H. A. (2011). Trajectory generation for autonomous soaring UAS. In *17th International Conference on Automation and Computing*.
- HACHIYA H. & SUGIYAMA M. (2010). Feature selection for reinforcement learning : Evaluating implicit state-reward dependency via conditional mutual information. In *European Conference on Machine Learning and Knowledge Discovery in Databases*, p. 474–489.
- KAHVECI N. & MIRMIRANI M. (2008). Adaptive LQ control with anti-windup augmentation to optimize UAV performance in autonomous soaring application. In *IEEE Transactions on Control System Technology*.
- LAWRANCE N. (2011). *Autonomous Soaring Flight for Unmanned Aerial Vehicle*. PhD thesis, The University Of Sydney.
- LAWRANCE N. & SUKKARIEH S. (2011). Path planning for autonomous soaring flight in dynamic wind. In *IEEE International Conference on Robotics and Automation*.
- NGUYEN T., LI Z., SILANDER T. & LEONG T. Y. (2013). Online feature selection for model-based reinforcement learning. In *Int. Conf. on Machine Learning*.
- PARR R., LI L., TAYLOR G., PAINTER-WAKEFIELD C. & LITTMAN M. L. (2008). An analysis of linear models, linear value-function approximation, and feature selection for reinforcement learning. In *International Conference on Machine Learning*.
- PUTERMAN M. L. (2005). *Markov Decision Processes : Discrete Stochastic Dynamic Programming*. John Wiley & Sons, Inc.
- REICHMANN H. (1993). *Cross-Country Soaring*. Soaring Society of America.
- ROBBINS H. & MONRO S. (1951). A stochastic approximation method. *Ann. Math. Statist.*, **22**(3), 400–407.
- SUTTON R. S. & BARTO A. G. (1998). *Reinforcement Learning : An Introduction*. MIT Press.
- WATKINS C. J. C. & DAYAN P. (1992).  $Q$ -learning. *Machine Learning*, **8**, 279–292.
- WHARINGTON J. (1998). *Autonomous Control of Soaring Aircraft by Reinforcement Learning*. PhD thesis, Royal Melbourne Institute of Technology.



# HUMANS : suite logicielle pour la scénarisation d'environnements virtuels pour la formation à des situations socio-techniques complexes

D. Lourdeaux<sup>1</sup>A. Benabbou<sup>1</sup>L. Huguet<sup>1</sup>R. Lacaze-Labadie<sup>1</sup><sup>1</sup> Sorbonne universités, Université de Technologie de Compiègne, CNRS

Heudiasyc UMR 7253, CS60319, 60203 Compiègne Cedex, France

{domitile.lourdeaux - azzeddine.benabbou - lauriane.huguet - remi.lacaze-labadie}@hds.utc.fr

## Résumé

*Les environnements virtuels permettent aux apprenants d'expérimenter, de s'entraîner et de voir l'impact de leurs décisions. L'un des challenges informatiques est de pouvoir scénariser ces environnements en orientant le scénario vers des situations d'apprentissage. L'approche HUMANS vise pour cette scénarisation un ensemble d'objectifs souvent considérés comme contradictoires : la liberté d'action pour favoriser l'apprentissage par l'erreur, le caractère dynamique et l'efficacité du contrôle exercé sur le scénario pour garantir des apprentissages, la cohérence des comportements pour permettre l'explicabilité et l'adaptabilité du système nécessaire pour la variabilité des scénarios.*

## Mots Clef

Environnement virtuels, formation, scénarisation, situations complexes, personnalisation.

## Abstract

*Virtual environments can provide dynamic situational training by offering training subjects the possibility to experiment with different behaviors and outcomes. To enable several pedagogical strategies, such as a trial-and-error approach or a situated learning approach, the system has to be able to cope with user agency. Combining pedagogical control with user agency, all the while ensuring that the simulation stays coherent, is fundamentally problematic when complex situations are addressed. To tackle the authoring bottleneck, rather than making explicit the whole set of possible scenarios, the HUMANS approach involves the dynamic computation of situations that varies depending on pedagogical rules.*

## Keywords

Virtual environnements, training, orchestration, complex situations, personalization.

## 1 Introduction

La réalité virtuelle permet aux apprenants d'expérimenter, de s'entraîner et de voir l'impact de leurs décisions. Elle est souvent utilisée pour former à des gestes techniques ou à des compétences techniques, mais elle est encore peu exploitée pour former à des situations socio-techniques complexes

(e.g. pressions temporelles, risques, équipes hétérogènes, expertises diverses). Dans ces situations où il n'y a pas toujours de solutions idéales, les opérateurs doivent s'adapter rapidement et prendre des décisions difficiles. Pour gérer ce type de situations, il est important de former les opérateurs à des compétences techniques mais aussi à des compétences non-techniques (e.g. leadership, travail en équipe, gestion du stress, prise de décision).

Afin de faciliter l'apprentissage de ces compétences par l'expérience personnelle, il est nécessaire de confronter l'apprenant avec une diversité de situations potentielles de développement [24], i.e. adaptées à son niveau de connaissance et de maîtrise de la tâche à réaliser en lui fournissant des ressources pour l'accomplir. Pour ne pas mettre l'apprenant en situation d'échec mais dans des situations de développement potentiel, les situations doivent être contrôlées pour être suffisamment difficiles mais pas trop et pour cibler les compétences adaptées à chaque apprenant.

Pour supporter ce type d'apprentissage, l'écriture des scénarios nécessite un travail conséquent. Cet effort nécessaire pour le passage à l'échelle de scénarios cohérents et précisément contrôlés constitue ce qu'on appelle l'*authoring bottleneck*<sup>1</sup> [22], et montre la nécessité de mettre en place des systèmes de scénarisation génériques et indépendants des applications, capables de créer des environnements adaptables, sans avoir à définir explicitement l'intégralité des scénarios possibles. Nous nous intéressons au processus de scénarisation de ces situations en environnement virtuel. La scénarisation est un processus comprenant à la fois la spécification du ou des déroulements possibles ou souhaitables de la simulation, et le contrôle (exécution, suivi et correction) du déroulement des événements en temps interactif. Un système de scénarisation est composé d'un ou plusieurs langages de scénarisation, pour modéliser le contenu scénaristique et les objectifs scénaristiques, et d'un moteur de scénarisation pour gérer dynamiquement la réalisation du scénario [3].

Dans cet article nous présentons la suite logicielle HUMANS (HUMAN Models based Artificial eNvironments Software platform), un système de scénarisation visant un ensemble d'objectifs souvent considérés comme contradictoires :

1. goulet d'étranglement de l'écriture

- la liberté d'action permettant l'apprentissage par l'erreur,
- le caractère dynamique et l'efficacité du contrôle du scénario pour garantir des apprentissages,
- la cohérence des comportements représentatifs de ceux observés sur le terrain et le transfert depuis et vers des situations réelles,
- l'explicabilité pour favoriser des apprentissages réflexifs, i.e. l'amener à réfléchir sur sa tâche et ses apprentissages,
- et l'adaptabilité du système, nécessaire pour la variabilité des scénarios.

## 2 Travaux connexes

Deux grandes catégories d'approches de systèmes de scénarisation sont souvent opposées : les approches orientées scénario et les approches centrées personnages. Les approches dites orientées scénario, mettent l'accent sur la qualité globale du scénario. Un contrôleur de scénario modifie dynamiquement les caractéristiques des personnages parfois sans cohérence entre leurs états mentaux et ce qui se passe dans l'environnement [16]. Or, [21] ont montré que l'incohérence dans les motivations des personnages affectait l'expérience de l'utilisateur et sa compréhension. Par ailleurs, dans des contextes de formation, il est alors difficile d'avoir des comportements explicables pour le jeu. Dans les approches centrées personnages, le scénario émerge des interactions entre l'utilisateur et les personnages virtuels. Le contrôle est distribué, chaque personnage étant responsable de ses propres prises de décision. Ces approches se focalisent sur la création de personnages virtuels dotés de processus décisionnels évolués (I-Storytelling [10], EmoEmma [18]). Si ces personnages s'adaptent dynamiquement et offrent une grande liberté d'action, le contrôle scénaristique reste difficile. De nombreux travaux visent des systèmes hybrides avec des personnages semi-autonomes aux comportements pseudo-cognitifs centrés scénario comme dans Scenario Adaptor [17], ISAT [16], IN-TALE [19] ou centrés personnages comme FearNot! [2], MRE [13]. Dans ces cas encore, il est difficile d'assurer la cohérence des comportements et d'avoir des comportements explicables. Des travaux novateurs ont été menés par [11] qui propose une planification sociale très évoluée avec un contrôle motivationnel ou encore dans Thespian qui propose un système de calcul des caractéristiques des personnages virtuels et de planification des événements à l'initialisation. Thespian garantit le passage par des points clés de l'histoire [21]. Toutefois, le système de [11] nécessite de définir explicitement des macro-actions de manière très contraignantes et Thespian ne permet pas l'adaptation dynamique du scénario et la réorientation du scénario en temps interactif.

**Positionnement.** Notre système de scénarisation doit assurer à la fois la liberté d'action de l'apprenant, le contrôle dynamique de la simulation, la cohérence des comportements présentés et l'adaptabilité de l'environnement vir-

tuel. Le verrou lié à ces travaux naît de l'incompatibilité entre ces différents objectifs : le contrôle s'oppose à l'adaptabilité, la liberté d'action va à l'encontre du contrôle, l'alliance du contrôle et de l'adaptabilité met en péril la cohérence, et ainsi de suite. Notre système de scénarisation doit être générique et indépendant des applications. Dans le cadre de nos travaux, le contrôle vise à personnaliser les situations d'apprentissage pour un seul apprenant.

## 3 L'approche HUMANS centrée personnages

Pour conserver la liberté d'action et assurer l'adaptabilité et l'explicabilité des comportements, nous faisons l'hypothèse qu'il est possible de générer dynamiquement et automatiquement des situations d'apprentissage à partir de modèles de connaissances qui sous-tendent la simulation. L'originalité de nos travaux est de centrer notre approche sur l'activité humaine et les facteurs humains. La suite logicielle HUMANS vise à créer des **environnements virtuels informés** prenant en compte ou rendant compte des **comportements humains** variés, en particulier en situations dégradées impliquant des compromis, des actions pouvant être erronées ou non-optimales. HUMANS est composée de moteurs décisionnels s'articulant autour de modèles qui leur permettent de disposer de représentations communes (figure 1).

### 3.1 Moteurs HUMANS

HUMANS est composé de quatre moteurs décisionnels : REPLICANTS gère les personnages virtuels autonomes, SELDON la scénarisation à travers deux sous-moteurs (TAILOR et DIRECTOR), WORLD MANAGER l'exécution du monde et MONITOR les traces de l'apprenant.

**REPLICANTS : Personnages virtuels autonomes.** Pour faciliter l'adaptabilité, les personnages virtuels sont autonomes et dotés de processus cognitifs. Ils peuvent réaliser des activités individuelles ou collectives (co-activité, collaboration, coopération). Ils ont des émotions, des personnalités diverses, des relations sociales. Ils peuvent être stressés, fatigués, expérimentés. Ils ont des comportements « humains » et peuvent faire des compromis, transgresser des règles de sécurité, faire des erreurs, générer des perturbations, faciliter le travail d'équipe ou pallier les comportements erronés des autres personnages ou de l'apprenant [12], [15], [14], [8].

**SELDON : Scénarisation.** Pour contrôler les situations de développement et conserver la cohérence du monde, le modèle de scénarisation, SELDON, oriente indirectement le déroulement des événements en réalisant des ajustements ponctuels sur l'état du monde ou sur les personnages sans leur donner d'ordres [3]. SELDON se décompose en deux processus : TAILOR sélectionne les objectifs scénaristiques et DIRECTOR orchestre ces objectifs.

a. TAILOR utilise la trace d'activité de l'apprenant pour diagnostiquer un profil dynamique [9]. Pour créer ce



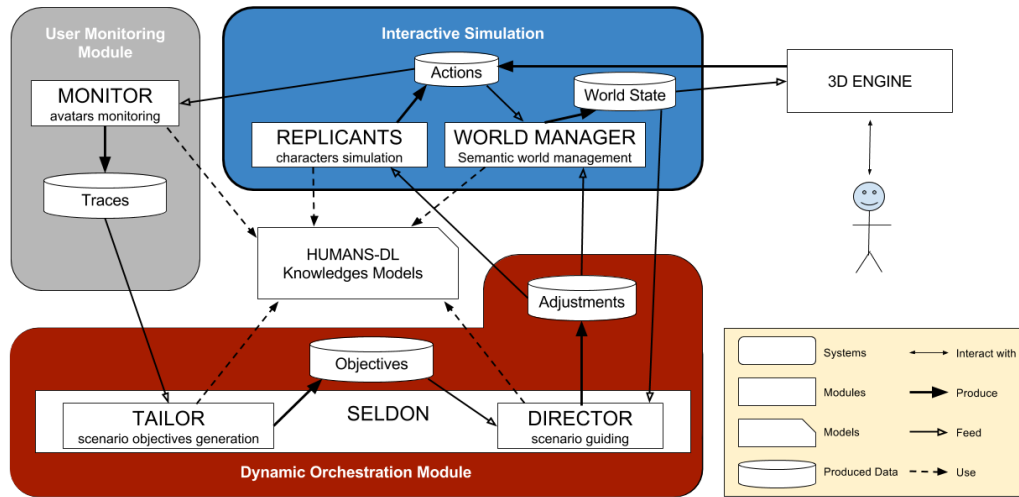


FIGURE 1 – Organisation générale de HUMANS

profil, nous proposons une formalisation mathématique et algorithmique de la théorie de Zone Proximale de Développement [24]. Un espace vectoriel de classes de situations est associé à des valeurs de croyance sur la capacité de l'apprenant à gérer ces situations. Les croyances sont calculées grâce à des mécanismes reposant sur la théorie des fonctions de croyances. À partir de ce modèle, un algorithme génétique sélectionne des objectifs scénaristiques sous la forme d'espaces de scénario dans la zone proximale de développement (figure 2). Il cherche à déterminer les situations et les propriétés du scénario jugées comme pertinentes pour l'apprenant. Par exemple, si l'apprenant n'arrive pas à gérer une classe de situation, le moteur sélectionne des situations du même ordre, voire plus simples. Au contraire, si l'apprenant gère correctement les situations, le moteur étend la zone proximale de développement avec des situations plus difficiles, plus critiques, etc. TAILOR propose alors à DIRECTOR des objectifs scénaristiques sous forme de valeurs de désirabilité sur des situations particulières et de contraintes globales sur les propriétés du scénario (e.g. complexité, criticité).

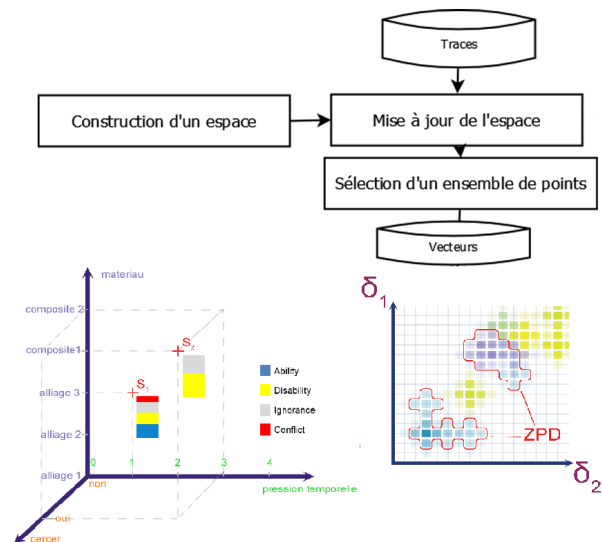


FIGURE 2 – TAILOR : sélection des espaces de scénario

**b.** Pour répondre à ces objectifs, DIRECTOR utilise les modèles qui sous-tendent la simulation pour prédire son évolution via un moteur de planification [3]. Il calcule un scénario à partir de prédictions et d'un ensemble d'ajustements possibles (figure 3). Trois types d'ajustements sont possibles : les happenings (déclenchement d'événements exogènes sans impact sur la cohérence du système), les late commitments et les contraintes d'occurrence. Le late commitment permet de préciser progressivement durant la simulation les états laissés incertains à l'initialisation. La modification des contraintes d'occurrence permet d'outrepasser les choix aléatoires des comportements probabilistes. Ces ajustements sont transmis aux moteurs d'exécution du monde et de génération des comportements des personnages virtuels pour influencer leur évolution. DIREC-

TOR compare l'évolution de l'état du monde avec le scénario planifié. Dans le cas où le scénario effectif dévie du scénario planifié, le moteur planifie un nouveau scénario tenant compte des changements. DIRECTOR repose sur des opérateurs de planification et les modèles de connaissances.

**MONITOR : suivi de l'apprenant.** MONITOR enregistre les actions de l'apprenant, de détecter les bonnes et les mauvaises pratiques, les conséquences effectives ou les probabilités de conséquences sur le système technique et humain et de faire un retour en temps réel au formateur en mode « jeu » (figure 8) à l'apprenant [1].

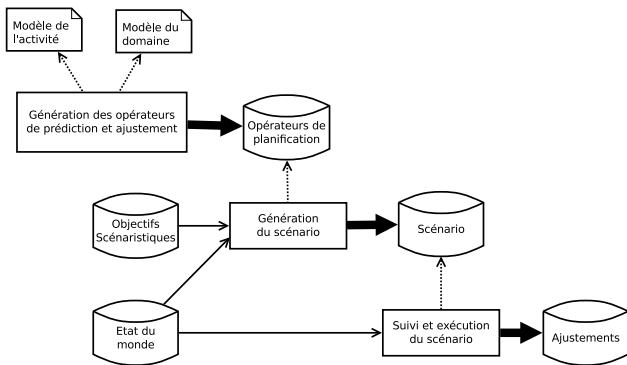


FIGURE 3 – Processus DIRECTOR

**WORLD MANAGER : gestion du monde.** Le moteur d'exécution du scénario, WORLD MANAGER maintient la cohérence et l'état du monde décrit par le modèle du monde. Il prend en compte les demandes d'action de l'apprenant ou des personnages virtuels pour mettre à jour l'état du monde en fonction des règles de fonctionnement [9]. Ce moteur offre également une interface aux autres moteurs décisionnels qui souhaitent obtenir des informations sur le modèle sémantisé du monde, ses règles de fonctionnement ou sur son état courant. Il crée automatiquement des faits qui sont valides avec l'état du monde courant. On dispose ainsi d'un pouvoir génératif suffisant pour générer procéduralement du contenu riche à la volée. Une ouverture des scénarios permise par l'indécision portée par les valeurs des états dits committables (non initialisés au démarrage) offre un mécanisme de résilience face aux actions de l'utilisateur, ou des entités autonomes, allant à l'encontre du scénario prescrit.

### 3.2 Modèles de connaissance

Dès lors que l'on souhaite introduire dans des environnements virtuels des entités devant interpréter leur environnement, la question du raisonnement sur des connaissances structurées se pose. Pour raisonner HUMANS s'appuie sur des modèles de connaissances regroupés sous l'appellation HUMANS-DL : le modèle d'activité, le modèle du monde, le modèle de causes/conséquences. Ils sont directement interprétables par les moteurs décisionnels et s'articulent entre eux.

**Approche orientée connaissance.** Les connaissances à représenter et à exploiter dans les environnements virtuels pour l'apprentissage humain concernent trois domaines : la scénarisation pédagogique (diagnostic, rétroactions, adaptation, etc.), la génération des comportements des personnages virtuels et les différents objets composant la scène virtuelle, les actions qui y sont possibles et les effets potentiels. Une contrainte forte réside dans le fait que les connaissances dans ces domaines doivent pouvoir être traitées de façon coordonnée, ce qui implique des représentations et des processus de traitement compatibles qui communiquent ou partagent certaines des don-

nées. Par exemple, l'état du monde virtuel et les actions autorisées à un instant constituent des données importantes, à la fois pour permettre aux personnages virtuels d'adapter leur comportement au contexte et pour identifier les actions entreprises par l'apprenant. Ces connaissances doivent être exprimées par des représentations symboliques sémantiquement interprétables. Ces modèles sont souvent renseignés par l'informaticien : il est l'interface entre l'environnement virtuel et les experts (du domaine, formateurs, etc.). Il doit transcrire dans des formalismes interprétables des connaissances complexes qui nécessitent des analyses très pointues. En informatique, plusieurs formalismes graphiques ou mathématiques ont été proposés pour modéliser ces connaissances dans les environnements virtuels (réseaux à propagation de marques, automates, réseaux de Petri, etc. voir un état de l'art complet dans [3]). Même si les formalismes tels que les automates rencontrent un grand succès dans le domaine de l'animation comportementale, notamment grâce à leur puissance d'expression et le caractère calculable et vérifiable, ils ne favorisent pas toujours la prise en compte des facteurs humains. Par ailleurs, ces modèles sont souvent illisibles dès que la tâche décrite devient un peu complexe : la taille des automates ou des réseaux de Petri augmentant de manière exponentielle en fonction de la taille de l'application, ils deviennent illisibles même par un spécialiste. Enfin, hormis les automates hiérarchiques parallèles ou réseaux de tâches hiérarchiques, le pouvoir d'abstraction de ce type de représentation est faible. Notre approche est un peu différente car nous souhaitons proposer des langages auteur pour permettre aux experts, spécialistes de l'analyse d'activité, du domaine ou encore de la pédagogie, de décrire ces expertises. Nous avons souhaité, ne pas fournir un seul modèle et un seul formalisme capable d'unifier toutes ces connaissances que seul un informaticien serait capable de renseigner. Nous proposons ainsi des modèles qui facilitent la manipulation de concepts communs avec des vues différentes pour chaque domaine (ergonomie, risque, didactique, pédagogie, fonctionnement technique, etc.). Ces modèles visent une conception collaborative. Chaque expert peut ainsi saisir ses modèles avec le même outil et sur la base du même modèle mais avec des approches différentes et des regards croisés.

**ACTIVITY-DL** est le modèle pour décrire l'activité de l'apprenant et des personnages virtuels autonomes [3]. L'activité est la mise en œuvre d'une procédure ou d'un ensemble d'actions en contexte réel. La modélisation de l'activité décrit l'observation de cette activité. L'activité se distingue de la procédure prescrite en ce qu'elle intègre également les comportements non nominaux mais qui sont malgré tout observables en contexte réel. Ainsi, l'activité décrit les écarts à la procédure s'illustrant par des arrangements ou des violations. Ces écarts nécessitent d'être caractérisés pour préciser dans quelles conditions ces comportements sont tolérables du point de vue de l'activité observée. ACTIVITY-DL est inspiré des langages de modélisation ergonomiques (MAD, GTA) et des HTN (voir

un état de l’art sur ces formalismes dans [7]). Il repose sur une représentation des tâches sous forme d’arbre. Les tâches abstraites sont des compositions de sous-tâches sur plusieurs niveaux. Au niveau le plus bas, les feuilles correspondent aux actions effectives réalisées par les opérateurs, et donc réalisables dans l’environnement. Celles-ci sont directement liées aux actions décrites dans le modèle du monde en WORLD-DL. Ce langage permet de définir d’une part chaque tâche avec ses conditions (préconditions, conditions d’arrêt) et son contexte d’exécution, et d’autre part une organisation hiérarchique entre ces tâches suivant une dimension temporelle en s’appuyant sur les opérateurs temporels d’Allen et une dimension logique (figure 4).

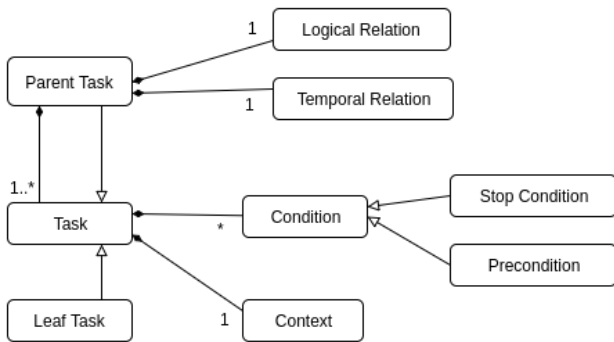


FIGURE 4 – ACTIVITY-DL

**WORLD-DL** permet de décrire à la fois les entités qui composent le monde, les actions possibles, les relations qu’elles entretiennent, mais également les comportements qui régissent l’évolution de ces entités, les conditions dans lesquelles les actions sont possibles et les effets qu’elles provoquent [9]. WORLD-DL s’appuie sur les standards de représentation OWL et RDF qui sont des formalismes informatiques, ce qui permet à la représentation d’être interprétable. En outre, la description des comportements du système dans une syntaxe telle que celle des règles JENA, un moteur de règles java, les rend explicites et donc accessibles et inspectables par des moteurs décisionnels tiers. Notamment, ces moteurs peuvent effectuer des requêtes et opérer des inférences sur le modèle pour générer leurs propres comportements. La posture ontologique, mise en œuvre au travers de OWL, offre une large expressivité. Et les primitives que nous avons définies pour la gestion de la partie dynamique du système offrent la possibilité de concevoir simplement des comportements non triviaux. Un des éléments important de ce méta-modèle est le late-commitment qui consiste à préciser progressivement durant la simulation des états laissés incertains à l’initialisation. Dans la plupart des environnements virtuels, le scénario de la simulation est limité par les conditions initiales. Cependant, il existe dans certains cas des états qu’il n’est pas nécessaire de définir dès le début de la simulation (états committable figure 5). Le late commitment, pro-

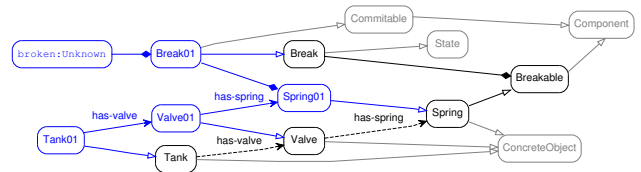


FIGURE 5 – Exemple d’état committable en WORLD-DL

posé par [23], consiste à ajouter, au cours de la simulation, un ensemble de faits qui seront considérés rétrospectivement comme faisant partie des conditions initiales. Ces faits sont ajoutés en fonction des besoins du scénario, à partir d’un ensemble de *framing operators* qui définissent les éléments pouvant y être ajoutés. Notre modèle du late commitment s’inspire de ce principe, mais se rapproche davantage des travaux de [20] sur l’Initial State Revision (ISR), en cela que ce dernier déclare explicitement les états considérés comme indéfinis et les ensembles de valeurs pouvant être prises par ces états. L’ISR propose la génération d’un scénario hors-ligne, la nécessité de spécifier les valeurs des états apparaît au cours du processus de planification, lorsque ces états interviennent dans les préconditions d’une action.

**CAUSALITY-DL** permet la description de chaînes de causalité significatives. Il ne s’agit pas de décrire exhaustivement les changements d’état du monde possibles, comme pourrait le faire un graphe d’état, mais uniquement ce qui est pertinent pour une application donnée. Il sert à spécifier les relations de causalité entre les événements d’intérêt qui peuvent survenir dans l’environnement virtuel, ainsi que les barrières qui peuvent empêcher la propagation de la causalité entre ces événements. CAUSALITY-DL utilise des éléments de l’ontologie du modèle du monde représenté en WORLD-DL, et est inspiré de formalismes d’analyses de risques : nœuds papillons et arbres de causes. Dans le cadre de la scénarisation, ce modèle permet de spécifier en amont un espace de trames scénaristiques d’intérêt (scénarios accidentels ressortis des analyses de risques, par exemple). Étant donné qu’il permet d’associer des valeurs aux événements (probabilité, complexité, etc.), il est également utilisé pour calculer les propriétés des différentes trames scénaristiques (crédibilité, complexité, gravité). La crédibilité est ici considérée comme la fréquence d’occurrence (événements courants, rares, improbables, etc.). Les modèles de causalité sont des graphes orientés acycliques dont les nœuds se réfèrent à des événements représentés dans le modèle du monde. Ces événements peuvent être liés par des relations de subsomption ou de causalité. Les relations de causalité sont également déterminées par des portes logiques indiquant la conjonction ou la disjonction, et sont conditionnées par la levée d’un ensemble de barrières, elles-aussi correspondant à des événements du modèle du monde (figure 6).

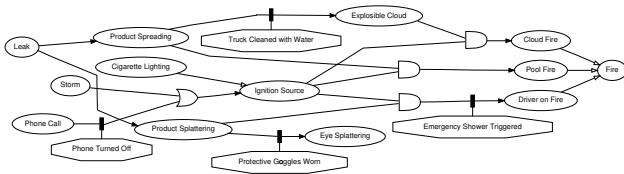


FIGURE 6 – Exemple représenté en CAUSALITY-DL

### 3.3 Généricité

HUMANS a été utilisée dans divers cas d’application de formation : prévention des risques, montage aéronautique, garde d’enfants, gestion de crise, sauvetage de blessés, gestion d’hémorragie post-partum et conduite automobile. HUMANS est modulaire, selon les applications ou besoins, tous les moteurs ne sont pas utilisés. Il est ainsi possible d’utiliser HUMANS comme soutien au formateur avec des conseils ou pour remplacer le formateur pour de la formation en autonomie. Nous détaillons ici certains de ces projets.

**Formation à la prévention des risques.** Le projet V3S a permis de développer deux environnements virtuels pour la formation à la prévention des risques : démontage de tuyaux et chargement de matières dangereuses sur des sites à haut risque [4]. Ce projet a permis de mettre en œuvre des personnages virtuels autonomes aux comportements erronés (e.g. compromis, violation, erreurs de connaissances), le suivi de l’apprenant et le moteur de planification DIRECTOR [3]. Pour illustrer le fonctionnement de DIRECTOR, imaginons que TAILOR souhaite déclencher une fuite. DIRECTOR pourrait choisir de déclencher une fuite de vanne. Il pourrait alors prévoir comme ajustements : un réservoir plein, un ressort usé pour empêcher la vanne de revenir et un événement pour stresser l’opérateur qui risque alors de ne pas accompagner la vanne (figure 7). DIRECTOR vérifie, ensuite, si tout se passe comme prévu .

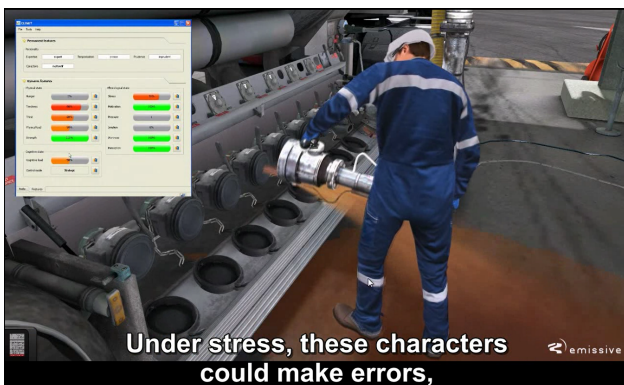


FIGURE 7 – Exemple d’ajustements - projet V3S

**Formation des assistantes de vie à la garde d’enfants.** Dans le projet SIMADVF, le moteur de suivi de l’apprenant

détecte les actions erronées de l’apprenant, de déterminer les conséquences plus ou moins graves de ses actions, de sélectionner des événements perturbateurs et de proposer un rejeu avec des explications sur les conséquences potentielles ou avérées des actions de l’apprenant [1] (figure 8).



FIGURE 8 – Rejeu - projet SIMADVF

**Formation au montage aéronautique.** Le projet NIKITA a permis d’élaborer le moteur de sélection d’objectifs scénaristiques, TAILOR. TAILOR identifie un profil dynamique de l’apprenant et sélectionne les situations mettant en jeu des compétences les plus adaptées (figure 9).

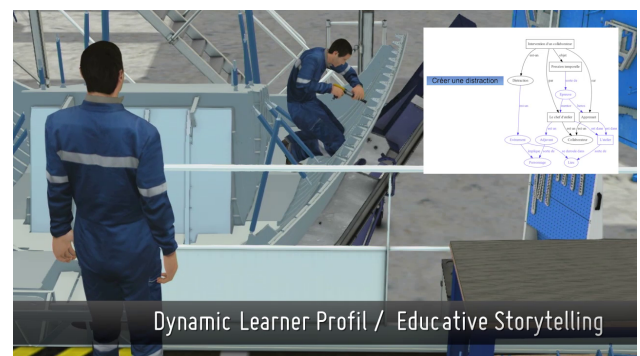


FIGURE 9 – Profil apprenant - projet NIKITA

**Formation au sauvetage de blessés.** Le projet VIC-TEAMS a pour objectif de former, en environnement virtuel, des leaders médicaux au sauvetage de blessés suite à un afflux massif de blessés (e.g. au combat, suite à des attentats ou catastrophes naturelles ou industrielles). Les membres de l’équipe sont joués par des personnages virtuels autonomes (figure 10). Dans ce projet les deux focus traités sont : la dimension collective et le stress. Le moteur décisionnel des personnages virtuels est capable de générer des comportements d’équipes variés. Ils peuvent reproduire des comportements erronés liés aux compétences non-techniques (e.g. croyances erronées sur le monde ou l’activité collective, mauvaises communications, comportements impactés par le stress). Le moteur de scénarisation

planifie, selon les compétences non-techniques de l'apprenant, des comportements "aidants" ou "perturbateurs" pour maintenir un certain niveau de difficulté et de stress. L'incertitude de l'évolution des situations est un facteur difficile pour prendre des décisions et notre objectif est non pas d'aider l'apprenant à trouver la solution idéale mais de l'aider à prendre conscience des choix qui s'offrent à lui et à prendre plus sereinement des décisions dans le futur.



FIGURE 10 – Équipes - projet VICTEAMS

**Formation aux situations critiques en conduite et en médecine.** MacCoy Critical est un projet pluridisciplinaire qui vise à étudier et à améliorer les dispositifs de formation utilisant la simulation et les environnements virtuels. Nous nous intéressons en particulier à l'entraînement aux compétences non-techniques dans les situations critiques. Le projet s'applique à deux domaines aux forts enjeux en santé publique et en sécurité qui sont la conduite automobile et la médecine. L'enjeu est de proposer un système de scénarisation capable de générer dynamiquement les situations critiques sans avoir à les écrire en amont. Nous avons défini, avec les partenaires du projet, des dimensions qui caractérisent ces situations. Parmi ces dimensions, nous trouvons les dilemmes [6]. Les situations dilemmatiques peuvent être définies comme étant des situations où il n'existe pas vraiment de "bonne" solution. C'est à dire des situations où le protagoniste est confronté à un choix exclusif dont les conséquences seront toujours négatives. Afin de générer ces situations, nous avons proposé une approche générative qui repose sur les modèles décrits précédemment [5]. Le modèle d'activité est utilisé pour identifier les tâches contradictoires, afin de garantir l'exclusivité du choix, et le modèle de causalité permet d'identifier les tâches dont la non-réalisation conduirait à des conséquences négatives. La figure 11 illustre un exemple de dilemme en conduite automobile.

## 4 Conclusion

Nous avons présenté l'approche HUMANS, un système de scénarisation d'environnements virtuels pour la formation qui répond à quatre objectifs : la liberté d'action de l'utilisateur, le caractère dynamique du contrôle, la cohérence

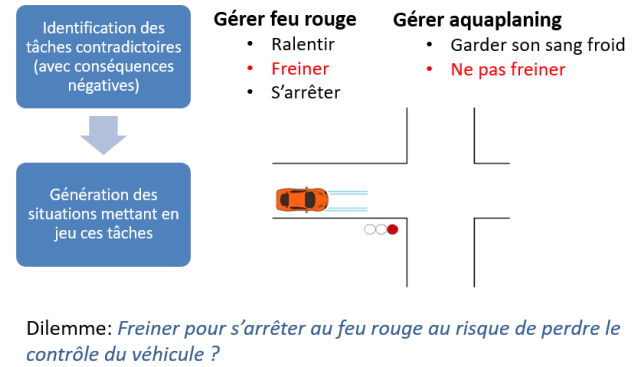


FIGURE 11 – Exemple de génération de dilemme - projet MacCoy Critical

des comportements et l'adaptabilité. HUMANS s'appuie sur une approche émergente et modulaire. La plate-forme a été exploitée dans des contextes de formation professionnelle variés.

Aujourd'hui, nous nous intéressons plus particulièrement à la scénarisation de situations critiques ou de crise impliquant des personnages virtuels autonomes engagés dans des décisions collectives. Nous visons à adapter ces situations selon le profil dynamique de l'apprenant (e.g. stress, compétences, émotions). La scénarisation que nous visons est une sorte de Kobayashi Maru inspiré de la simulation pour l'entraînement des officiers de Starfleet dans l'univers de fiction de Star Trek. Le but de cette simulation est d'analyser la réaction des officiers face à une situation sans issue. Le système cherche alors à générer des situations dans lesquelles il n'est pas possible de gagner. La différence ici est qu'en situation de formation, nous cherchons à générer des situations limites mais dans lesquelles l'apprenant ne doit pas sortir complètement perdant. Il faut donc trouver le compromis entre challenge, stress et apprentissage.

## Remerciements

Les auteurs remercient l'ANR, la DGA, la DGCIS, la Région Hauts-de-France et le FEDER financeurs de ces travaux. Les auteurs remercient les co-directeurs de thèse, tous les anciens doctorants qui ont contribué à construire HUMANS : Lydie Edward, Kahina Amokrane, Margot Lhommet, Camille Barot, Kevin Carpentier, Lucile Callebort et tous les étudiants qui ont contribué aux développements. Les auteurs remercient les partenaires des différents projets qui ont permis l'avancement de ces travaux.

## Références

[1] K. Amokrane-Ferka, D. Lourdeaux, and G. Michel. Tracking and dynamic scenario adaptation system in virtual environment. In Springer, editor, *Proceeding of The 16th International Conference on Artificial Intelligence in Education*, volume 7926, pages 758–761, Memphis, États-Unis, Jul 2013.

- [2] R. Aylett, R. Figueiredo, S. Louchart, J. Dias, and A. Paiva. Making it up as you go along-improvising stories for pedagogical purposes. In *Intelligent Virtual Agents*, pages 304–315, 2006.
- [3] Camille Barot. *Scénarisation d’environnement virtuel. Vers un équilibre entre contrôle, cohérence et adaptabilité*. PhD thesis, Université de Technologie de Compiègne, Compiègne, February 2014.
- [4] Camille Barot, Domitile Lourdeaux, Jean-Marie Burkhardt, Kahina Amokrane, and Dominique Lenne. V3S : a virtual environment for risk management training based on human-activity models. *PRESENCE : Teleoperators and Virtual Environments*, 2013.
- [5] Azzeddine Benabbou, Domitile Lourdeaux, and Dominique Lenne. Dynamic generation of dilemma-based situations in virtual environments. In *Intelligent Tutoring Systems*, page 526. Springer, 2016.
- [6] Jean Marie Burkhardt, Valentin Corneloup, Catherine Garbay, Yannick Bourrier, F Jambon, Vanda Luengo, A Job, Ph Cabon, A Benabbou, and D Lourdeaux. Simulation and virtual reality-based learning of non-technical skills in driving : critical situations, diagnostic and adaptation. *IFAC-PapersOnLine*, 49(32) :66–71, 2016.
- [7] J.M. Burkhardt, D. Lourdeaux, S. Couix, and M. Rouillé. La modélisation de l’activité humaine finalisée. In *Le traité de la réalité virtuelle : Volume 5, L’humain virtuel*. Les Presses-Mines Paris, 2009.
- [8] Lucile Callebert, Domitile Lourdeaux, and Jean-Paul A. Barthès. A trust-based decision-making approach applied to agents in collaborative environments. In *Proceedings of the 8th International Conference on Agents and Artificial Intelligence (ICAART 2016), Volume 1, Rome, Italy, February 24-26, 2016.*, pages 287–295, 2016.
- [9] Kevin Carpentier. *Scénarisation personnalisée dynamique dans les environnements virtuels pour la formation*. PhD thesis, Université de Technologie de Compiègne, Compiègne, January 2015.
- [10] M. Cavazza, F. Charles, and S. J Mead. Character-based interactive storytelling. *IEEE Intelligent Systems*, 17(4) :17–24, 2002.
- [11] Hsueh-Min Chang and Von-Wun Soo. Planning to influence other characters in agent-based narratives. In *Integrating Technologies for Interactive Stories Workshop, International Conference on Intelligent Technologies for Interactive Entertainment*, pages 12–17, 2008.
- [12] Lydie Edward. *Modélisation décisionnelle de personnages virtuels autonomes évoluant dans un environnement pour la présentation des risques sur les sites SEVESO*. PhD thesis, Université de Technologie de Compiègne, 2011.
- [13] R. Hill, J. Gratch, W. L. Johnson, C. Kyriakakis, C. LaBore, R. Lindheim, S. Marsella, D. Miraglia, B. Moore, J. Morie, J. Rickel, M. Thiébaux, L. Tuch, R. Whitney, J. Douglas, and W. Swartout. Toward the holodeck : integrating graphics, sound, character and story. In *Proceedings of the fifth international conference on Autonomous agents*, AGENTS ’01, pages 409–416, New York, NY, USA, 2001. ACM.
- [14] Lauriane Huguet, Nicolas Sabouret, and Domitile Lourdeaux. Simuler des erreurs de communication au sein d’une équipe d’agents virtuels en situation de crise. In *Rencontres des jeunes chercheurs en Intelligence Artificielle (RFIA 2016)*, Actes Rencontres des Jeunes Chercheurs en Intelligence Artificielle, Clermont-Ferrand, France, June 2016.
- [15] Margot Lhomme. *REPLICANTS : Humains virtuels cognitifs, émotionnels et sociaux. De l’empathie cognitive à l’empathie affective*. PhD thesis, Université de Technologie de Compiègne, 2012.
- [16] B Magerko, R. E. Wray, L. S. Holt, and B. Stensrud. Improving interactive training through individualized content and increased engagement. In *Interservice/Industry Training, Simulation & Education Conference (IITSEC)*, 2005.
- [17] J. Niehaus and M. Riedl. Scenario adaptation : An approach to customizing computer-based training games and simulations. In *14th International Conference on Artificial Intelligence in Education (AIED 2009) Workshops Proceedings*, page 89, 2009.
- [18] D. Pizzi and M. Cavazza. Affective storytelling based on characters’ feelings. In *Proceedings of the AAAI Fall Symposium on Intelligent Narrative Technologies*, 2007.
- [19] M. O Riedl, A. Stern, D. Dini, and J. Alderman. Dynamic experience management in virtual worlds for entertainment, education, and training. *International Transactions on Systems Science and Applications, Special Issue on Agent Based Systems for Human Learning*, 4(2) :23–42, 2008.
- [20] Mark Owen Riedl. *Narrative generation : balancing plot and character*. PhD thesis, North Carolina State University, 2004. AAI3154351.
- [21] Mei Si. *Thespian : a decision-theoretic framework for interactive narratives*. PhD thesis, University of Southern California, 2010.
- [22] Ulrike Spierling and Nicolas Szilas. Authoring issues beyond tools. In *Interactive Storytelling*, pages 50–61. Springer, 2009.
- [23] Ivo Martinus Theodorus Swartjes. *Whose story is it anyway? : how improv informs agency and authorship of emergent narrative*. PhD thesis, Enschede, May 2010. SIKS Dissertation Series No. 2010-20.
- [24] L. S. Vygotsky. *Mind in Society*. Harvard University Press, Cambridge, MA, 1978.

# Graphe de connaissances et folksonomie : leur performance comparative dans le calcul de l'affinité

Chun Lu<sup>1,2</sup>Philippe Laublet<sup>1</sup>Milan Stankovic<sup>1,2</sup>Filip Radulovic<sup>2</sup><sup>1</sup> Laboratoire STIH, Université Paris-Sorbonne, 28 rue Serpente, 75006 Paris<sup>2</sup> Sépage, 27 rue du Chemin Vert, 75011 Paris

## Résumé

*L'affinité est un élément essentiel dans bien des systèmes d'information centrés sur l'utilisateur comme les systèmes de recommandation. Le graphe de connaissances et la folksonomie sont respectivement des jalons importants pour le Web Sémantique et le Web Social. Nonobstant leur trait collaboratif partagé (du moins quelques grands graphes de connaissances le sont), les données codées diffèrent tant par la nature (fait versus expérience) que par la structure (formelle versus lâche). Dans ce papier, nous tentons d'éclaircir leur performance comparative dans la tâche du calcul de l'affinité à travers deux expériences dans le domaine du e-tourisme. Nos résultats montrent que le graphe de connaissances permet de calculer l'affinité avec plus de précision alors que la folksonomie augmente la diversité et la nouveauté. Ces constatations nous ont motivés à développer le Framework d'Affinité Sémantique pour bénéficier de leurs avantages respectifs. L'original de ce papier est publié par ESWC 2017.*

## Mots Clef

Affinité, recommandation, collaboratif, similarité, sémantique, graphe de connaissances, folksonomie, e-tourisme.

## Abstract

User-entity affinity is an essential component of many user-centric information systems. Knowledge graph and folksonomy are respectively the milestones of Semantic Web and Social Web. Despite their shared crowdsourcing trait (not necessarily all knowledge graphs but some major large-scale ones), the encoded data are different in nature and structure. Knowledge graph encodes factual data with a formal ontology. Folksonomy encodes experience data with a loose structure. In this paper, we are interested in observing their comparative performance in the affinity assessment task. To this end, we carried out two experiments within a travel destination recommendation scenario. Our main findings are that knowledge graph helps to assess more accurately the affinity but folksonomy helps to increase the diversity and the novelty. This interesting complementarity motivated us to develop a semantic affinity framework to harvest the benefits of both data spaces. The original paper is published by ESWC 2017.

## Keywords

Affinity, recommendation, collaborative, similarity, semantic, knowledge graph, folksonomy, e-tourism

## 1 Introduction

L'affinité entre un utilisateur et une entité (ex. film, musique, artiste) est la probabilité que l'utilisateur soit attiré par l'entité ou réalise une action liée à elle (ex. cliquer, acheter, aimer, partager). L'affinité est un élément essentiel dans bien des systèmes d'information centrés sur l'utilisateur comme les systèmes de recommandation, la publicité en ligne, la recherche exploratoire etc. Parmi les techniques de calcul de l'affinité, celles basées sur le contenu posent l'hypothèse que l'utilisateur aurait une affinité plus élevée avec les entités similaires à celles qu'il a appréciées dans le passé. Le graphe de connaissances et la folksonomie sont respectivement des jalons importants pour le Web Sémantique et le Web Social. Ils ont tous deux boosté les techniques basées sur le contenu grâce au grand nombre de données disponibles sur les entités. Sur le Web Sémantique, les utilisateurs contribuent à la création des graphes de connaissances universels comme DBpedia et Wikidata. Sur le Web Social, les utilisateurs annotent et catégorisent les entités avec des étiquettes libres formant des folksonomies. Nonobstant leur trait collaboratif partagé, les données codées diffèrent tant par la nature que par la structure. Les graphes de connaissances structurent des données factuelles avec une ontologie. Les folksonomies contiennent des données d'expérience avec une structure lâche. Nous donnons un exemple pour illustrer leur différence. Sur DBpedia, le film *dbr:Jumanji* est associé aux faits comme *dbr:Joe Johnston* par la propriété *dbo:director*, *dbr:Robin Williams* par *dbo:starring*. Dans la folksonomie de MovieLens<sup>1</sup>, le même film est abondamment annoté avec des étiquettes comme « nostalgic », « not funny », « natural disaster » etc. Ces étiquettes reflètent l'expérience qu'ont eue les différents utilisateurs et ainsi une sorte d'intersubjectivité qui n'est pas présente dans les graphes de connaissances.

Après une étude approfondie de la littérature (section 2), nous n'avons pas réussi à trouver des éclairages utiles sur l'efficacité comparative de ces deux espaces de données dans la tâche du calcul de l'affinité. Ces deux espaces de données continuant de proliférer sur le web,

<sup>1</sup> <https://movielens.org/>

il est plus que jamais nécessaire de faire la lumière sur cette question. Nous étudions cette question à travers deux expériences dans le domaine du e-tourisme. La première expérience hors ligne est décrite dans la section 3 dont les constatations ont motivé le développement du Framework d’Affinité Sémantique (section 4). La section 5 présente la deuxième expérience qualitative et évalue le framework proposé. La section 6 conclut le papier.

## 2 Travaux antérieurs

Depuis plus d’une décennie, les chercheurs étudient de près les liens entre le Web Sémantique et le Web Social. L’idée générale derrière ces efforts est d’augmenter la sémantique du Web Social à l’aide des technologies du Web Sémantique (Bontcheva et Rout, 2014). (Passant et Laublet, 2008) proposent l’ontologie *MOAT* et un framework collaboratif pour guider les utilisateurs à fournir la sémantique des tags (étiquettes) lors du processus d’annotation. (Mika, 2007) propose de construire des ontologies légères à partir des folksonomies. (Cantador et al., 2011) présentent une méthode utilisant le Web Sémantique et le traitement automatique des langues pour classer les en fonction de l’intention derrière l’application des tags. Certains auteurs essaient d’extraire des préférences utilisateurs à partir des folksonomies (Orlandi et al., 2012). D’autres essaient de prouver l’avantage de les utiliser dans les systèmes de recommandation (Semeraro et al., 2012). Du côté des graphes de connaissances, les auteurs les exploitent pour calculer la similarité sémantique entre les entités et les incorporent dans les systèmes de recommandation. Dans (Passant, 2010) et (Piao et Breslin, 2016), les auteurs présentent respectivement *Linked Data Semantic Distance* et *Resource Similarity* qui exploitent DBpedia pour recommander des artistes musicaux. Des variantes de *Spreading Activation* sont utilisées dans des systèmes de recommandation inter-domaines (Kaminskas et al., 2014) et de recherche exploratoire (Marie, 2014).

## 3 Première expérience

Nous avons mené une première expérience dans un scénario de recommandation des destinations de voyage.

### 3.1 Jeu de données

Nous avons adapté le jeu de données Yahoo! Flickr Creative Commons 100<sup>2</sup> (YFCC100M) (Thomee et al., 2016). YFCC100M contient 100 millions de photos et de vidéos géo-localisées et datées publiées sur Flickr. Les traitements principaux sont les suivants : 1. Retenir seulement les photos/vidéos avec la meilleure précision de géolocalisation 2. Associer chaque photo/vidéo à un lieu d’intérêt dans un graphe de connaissances (Lu et al., 2016) 3. Trier par ordre chronologique. A l’issue de ces traitements, nous avons obtenu, pour chaque utilisateur, une séquence de voyages contenant les villes

qu’il a visitées successivement. Le jeu de données final<sup>3</sup> utilisé dans cette expérience contient 3878 utilisateurs et 705 villes. Les séquences de voyages contiennent en moyenne 5,27 villes.

## 3.2 Traitements de la folksonomie et du graphe de connaissances

Nous utilisons une folksonomie collectée sur un site de voyage collaboratif. Elle contient 234 étiquettes sur 26,237 villes dans 154 pays. Nous l’avons modélisée dans un espace vectoriel de type *tag genome* (Vig et al., 2012) où les villes sont notées sur une échelle continue de 0 à 1 pour chaque étiquette. La mesure cosinus est utilisée pour calculer la similarité entre les villes. Pour chacune des 705 villes dans le jeu de données, nous avons exécuté des requêtes SPARQL avec toutes les propriétés sélectionnées (TABLE 1). Les ressources liées par la propriété *skos:broader* sont assimilées à celles liées par *dct:subject*. Nous avons ensuite éliminé les ressources qui sont liées à une seule ville car elles ne contribuent guère au calcul de la similarité. 501,365 ressources sont initialement obtenues et 29,743 d’entre elles sont finalement retenues. Nous avons adopté la mesure de Jaccard dont l’efficacité a été démontrée dans une comparaison avec des mesures plus sophistiquées (*VsmSim*, *GbkSim*, *FuzzySim*) dans un scénario de recommandation musicale (Nguyen et al., 2015).

Table 1 – Propriétés DBpedia sélectionnées pour calculer la similarité entre les villes

Entrant		Sortant	
dbo:birthPlace	dbo:broadcastArea	dbo:isPartOf	dbo:part
dbo:location	dbo:nearestCity	dbo:country	dbo:twinTown
dbo:deathPlace	dbo:ground	dbo:timeZone	dbo:saint
dbo:city	dbo:foundationPlace	dbo:Mayor	dbo:district
dbo:capital	dbo:assembly	dbo:region	dct:subject
dbo:hometown	dbo:restingPlace	dbo:province	(skos:broader)
dbo:recordedIn	dbo:place	dbo:leaderName	
dbo:residence	dbo:locationCity		
dbo:headquarter			

### 3.3 Calcul de l’affinité

Etant donné un profil utilisateur contenant une liste de villes visitées dans le passé, le score d’affinité d’une ville candidate  $v_i$  est la moyenne des scores de similarité qu’elle a avec chacune des villes de son profil.

$$\text{affinité}(u, v_i) = \frac{\sum_{v_j \in \text{profil}(u)} \text{sim}(v_i, v_j)}{|\text{profil}(u)|} \quad (1)$$

### 3.4 Protocole et métriques

Pour une séquence de voyages de  $n$  villes, les  $n-1$  premières villes constituent le profil utilisateur, la  $n$ -ième ville est considérée comme la vérité terrain. Chaque approche prend le profil utilisateur en entrée et génère trois listes de recommandations contenant respectivement 10, 20, 30 villes dans l’ordre décroissant des scores d’affinité. Dans le scénario de la recommandation, la capacité du calcul de l’affinité est reflétée par la précision. Nous utilisons deux métriques

3 Le jeu de données ainsi que certaines autres ressources mentionnées se trouvent à : <https://bitbucket.org/sepage/semantic-affinity-framework>

2 <http://webscope.sandbox.yahoo.com/>



pour la mesurer : Succès (2) et Mean Reciprocal Rank (MRR) (3). Depuis quelques années, les chercheurs qui travaillent sur les systèmes de recommandation portent un intérêt particulier sur la diversité et la nouveauté. Ces deux qualités sont mesurées avec les formules (5) et (6). Par analogie avec (Di Noia et al., 2014), la diversité intra-liste (ILS) est calculée par rapport à deux propriétés : *dbo:country* et *dct:subject*. Nous utilisons les valeurs pagerank des ressources DBpedia comme les scores de popularité. Comme dans (Nguyen et al., 2015), nous considérons que 20% des villes ayant les meilleurs scores sont populaires.

$$\text{Succès} = \frac{\sum_{u \in U} rel_{g,u}}{|U|} \text{ où } rel_{g,u} = \begin{cases} 1, & \text{si vérité terrain } g \text{ est dans top-}N \\ 0, & \text{sinon} \end{cases} \quad (2)$$

$$\text{MRR} = \frac{1}{|U|} \sum_{u \in U} \frac{1}{rank_u} \quad (3)$$

$$\text{ILS}_u@N = \sum_{i \in L_u^N} \sum_{j \in L_u^N} \frac{sim(i,j)}{|paires|} \quad (4)$$

$$\text{ILS}@N = \frac{1}{|U|} \sum_{u \in U} \text{ILS}_u@N \quad (5)$$

$$\text{Nouveauté}@N = \frac{\text{nombre de villes recommandées impopulaires}}{N * |U|} \quad (6)$$

### 3.5 Résultats et discussion

Les résultats sont présentés dans la TABLE 2. Les tests t appariés montrent que les différences entre les deux approches sont statistiquement significatives sur toutes les métriques dans toutes les configurations. Nous observons un avantage net de GC sur FOLK en termes de succès et de MRR, ce qui reflète la capacité de détecter les villes qui sont en affinité élevée avec l'utilisateur et de mieux les ordonner. Les recommandations générées par FOLK sont plus diverses et plus nouvelles. Dans la folksonomie, certains aspects ne sont pas présents, tels que la géographie (*dbo:country*, *dbo:region*), les personnes (*dbo:birthPlace*, *dbo:residence*), les catégories connexes (*dct:subject*, *skos:broader*). La folksonomie contient des traits tels que « Luxury Brand Shopping », « Clean Air » et « Traditional food ». Ces traits peuvent être partagés par de différentes villes du monde même si elles sont moins populaires. Ces constatations nous ont motivés pour développer le Framework d'Affinité Sémantique qui profite de la complémentarité des deux approches pour parvenir à un compromis équilibré sur la pertinence, la diversité et la nouveauté.

TABLE 2 – Résultats de la première expérience, GC : Graphe de connaissance, FOLK : Folksonomie

	Top-10		Top-20		Top-30	
	GC	FOLK	GC	FOLK	GC	FOLK
Succès						
MRR	<b>0.232</b>	0.06	<b>0.33</b>	0.116	<b>0.386</b>	0.166
ILS	<b>0.047</b>	0.003	<b>0.047</b>	0.003	<b>0.047</b>	0.003
Nouveauté	0.257	<b>0.089</b>	0.208	<b>0.072</b>	0.176	<b>0.065</b>

### 4 Framework d'Affinité Sémantique

Nous proposons un Framework d'Affinité Sémantique (FAS) qui intègre, agrège, enrichit et nettoie les données sur les entités en provenance des graphes de connaissances et des folksonomies. Son pipeline est décrit dans FIGURE 1. Une explication plus détaillée se trouve dans (Lu et al., 2017). Un graphe d'affinité est

généralisé à l'issue du processus. Nous avons également développé un mécanisme pour expliquer les recommandations. Par exemple, on peut expliquer la recommandation de *dbr:Ljubljana* par *dbc:Capitals in Europe*. Etant donné une liste d'entités, nous cherchons les caractéristiques les plus fréquentes tout en contrôlant leur diversité par le biais des propriétés qui relient les caractéristiques aux entités.

### 5 Deuxième expérience

Nous avons mené une deuxième expérience qualitative pour évaluer l'utilité et l'efficacité du Framework d'Affinité Sémantique. Nous avons généré un graphe d'affinité avec les données décrites dans 3.2 que nous nommons GA. A part la recommandation, nous nous sommes intéressés à la capacité d'explication des différentes approches.

You submitted:	You might like:	We recommend you:
dbr:Rome	dbc:Clothing	dbr:The_Hague
dbr:Florence	dbr:Food	dbr:Haarlem
dbr:Amsterdam	dbr:David_de_Haen	dbr:Naples
	dbr:Italy	dbr:Milan
	dbr:History	dbr:Turin

Figure 2 – Exemple de recommandations et d'explications générées par GA

37 personnes âgées de 25 à 38 ans ont participé à cette expérience dont 19 hommes et 18 femmes. Ils travaillent tous dans des sociétés sises dans une pépinière d'entreprise à Paris. Nous avons demandé aux participants de se mettre dans le scénario de rechercher pour la prochaine destination de voyage. Ils sont allés sur l'interface de test où les 705 villes étaient affichées dans un ordre aléatoire. Ils ont choisi plusieurs villes qui leur paraissaient intéressantes à première vue. Une fois que ces choix ont été soumis, ils ont reçu trois listes de recommandations contenant cinq villes accompagnées d'une explication avec cinq entités/étiquettes/catégories. FIGURE 2 montre un exemple généré par GA. Ils ont noté la pertinence, la diversité et la nouveauté/intérêt des recommandations et des explications sur une échelle de 1 à 5. Nous considérons que les notes supérieures à 3 comme notes positives. Nous utilisons comme métrique le pourcentage des notes positives.

Sur la recommandation, les résultats présentés dans FIGURE 3 sont en phase avec ceux de la première expérience. Sur l'explication, les résultats montrent que celle générée par FOLK est la mieux appréciée. Notre folksonomie a été créée de manière collaborative par les voyageurs. Elle couvre plusieurs aspects du voyage (nourriture, activité, transport). La capacité d'explication de GA a été boostée par l'inclusion de la folksonomie, ce qui l'a permis de surperformer GC. Les participants sont en général sceptiques envers GC. Certains trouvent ses explications assez générales, par exemple *dbr:Leisure*. D'autres les trouvent difficiles à comprendre, par exemple *dbr:China\_Record\_Corporation*. En effet, ces problèmes pourraient être résolues à travers l'utilisation de l'arbre des catégories DBpedia (définir un seuil en deçà duquel une catégorie peut être considérée comme

étant trop générale) ou le filtrage sur *rdf:type* (mettre certaines classes sur liste noire). Ces filtrages pourraient être intégrés dans le pipeline du Framework d’Affinité Sémantique.

## 6 Conclusion

Dans ce papier, nous avons étudié la performance comparative du graphe de connaissance et de la folksonomie dans la tâche du calcul de l’affinité à travers deux expériences dans le domaine du e-tourisme. Les résultats montrent que le graphe de connaissances permet de calculer l’affinité avec plus de précision alors que la folksonomie augmente la diversité et la nouveauté. Nous avons développé le Framework d’Affinité Sémantique pour bénéficier de leurs avantages respectifs. La combinaison des deux espaces de données aboutit à une performance équilibrée tant pour la recommandation que pour l’explication. Plus de détails se trouvent dans l’original de ce papier (Lu et al., 2017).

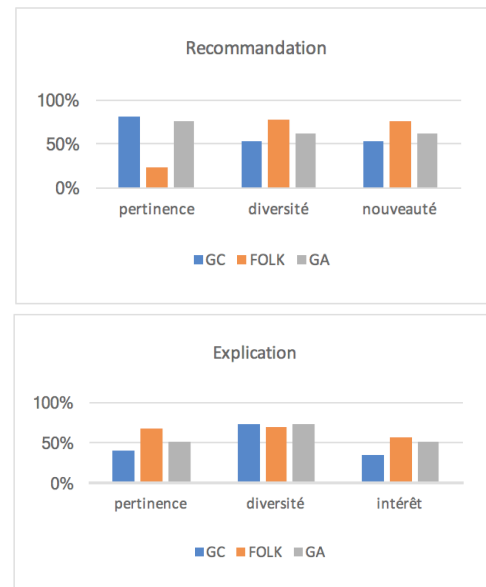


Figure 3 – Résultats de la deuxième expérience

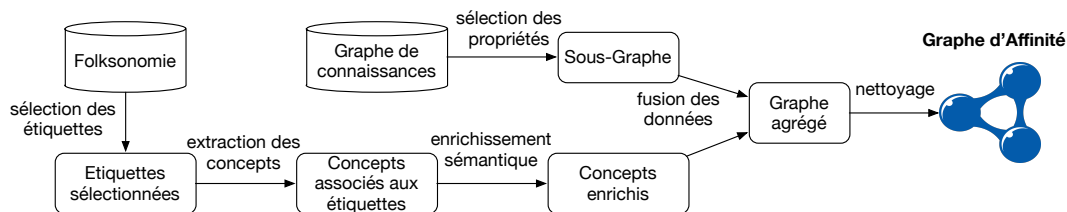


FIGURE 1 – Pipeline du Framework d’Affinité Sémantique

## Bibliographie

- [1] Bontcheva, K., Rout, D. (2014). Making sense of social media streams through semantics: a survey. *Semantic Web*, 5(5), 373-403
- [2] Cantador, I., Konstas, I., Jose, J. M. (2011). Categorising social tags to improve folksonomy-based recommendations. *Web Semantics: Science, Services and Agents on the World Wide Web*, 9(1), 1-15
- [3] Di Noia, T., Cantador, I., & Ostuni, V. C. (2014, May). Linked open data-enabled recommender systems: ESWC 2014 challenge on book recommendation. In *Semantic Web Evaluation Challenge* (pp. 129-143). Springer International Publishing.
- [4] Kaminskis, M., Fernández-Tobías, I., Ricci, F., Cantador, I. (2014). Knowledge-based identification of music suited for places of interest. *Information Technology & Tourism*, 14(1), 73-95
- [5] Lu, C., Laublet, P., Stankovic, M. (2016). Travel attractions recommendation with knowledge graphs. In: *Proceedings of the 20th International Conference on Knowledge Engineering and Knowledge Management*. Bologna, Italy
- [6] Lu, C., Stankovic, M., Radulovic, F., & Laublet, P. (2017, May). Crowdsourced Affinity: A Matter of Fact or Experience. In *European Semantic Web Conference* (pp. 554-570). Springer, Cham.
- [7] Marie, N. (2014). Linked data based exploratory search. Doctoral dissertation. Université de Nice Sophia-Antipolis
- [8] Mika, P. (2007). Ontologies are us: A unified model of social networks and semantics. *Web semantics: science, services and agents on the World Wide Web*, 5(1), 5-15
- [9] Nguyen, P. T., Tomeo, P., Di Noia, T., & Di Sciascio, E. (2015). Content-based recommendations via DBpedia and Freebase: a case study in the music domain. In *International Semantic Web Conference* (pp. 605-621). Springer International Publishing.
- [10] Orlandi, F., Breslin, J., Passant, A. (2012). Aggregated, interoperable and multi-domain user profiles for the Social Web. In: *Proceedings of the 8th International Conference on Semantic Systems* (pp. 41-48). ACM
- [11] Piao, G., and Breslin, J. (2016). Measuring semantic distance for linked open data-enabled recommender systems. In: *Proceedings of the 31st Annual ACM Symposium on Applied Computing*. ACM
- [12] Passant, A.: dbrec—music recommendations using DBpedia. In: *Proceedings of the 9th International Semantic Web conference* (pp. 209-224). Springer Berlin Heidelberg (2010)
- [13] Passant, A., Laublet, P. (2008). Meaning Of A Tag: A collaborative approach to bridge the gap between tagging and Linked Data. In: *Proceedings of Linked Data on the Web workshop*

- [14] Semeraro, G., Lops, P., De Gemmis, M., Musto, C., & Narducci, F. (2012). A folksonomy-based recommender system for personalized access to digital artworks. *Journal on Computing and Cultural Heritage (JOCCH)*, 5(3), 11
- [15] Thomee, B., Shamma, D. A., Friedland, G., Elizalde, B., Ni, K., Poland, D., and Li, L. J. (2016). YFCC100M: The new data in multimedia research. *Communications of the ACM*, 59(2), (pp. 64-73)
- [16] Vig, J., Sen, S., & Riedl, J. (2012). The tag genome: Encoding community knowledge to support novel interaction. *ACM Transactions on Interactive Intelligent Systems (TiiS)*, 2(3), 13.



# 3DRESC-TF : Apprentissage par transfert pour la réutilisation de connaissances en classification d'objets 3D

François Meunier<sup>1,2</sup>Christophe Marsala<sup>1</sup>Laurent Castanié<sup>2</sup>

<sup>1</sup> Sorbonne Universités, UPMC Univ Paris 06, CNRS, LIP6 UMR 7606  
4 place Jussieu, 75005 Paris

<sup>2</sup> Total Exploration-Production, Tour Coupole, La Défense  
2 Place Jean Millier, 92800 Puteaux

Francois.Meunier@lip6.fr

## Résumé

Dans cet article, nous proposons une amélioration basée sur du transfert de connaissances d'un algorithme de classification d'objets 3D existant, afin de pouvoir faire face aux manques de données, problème fréquent en apprentissage.

L'idée maitresse réside dans le fait de transférer directement les sous-parties constructrices d'attributs d'un domaine vers un autre, afin de conserver les avantages d'interprétabilité et de souplesse de la méthode d'origine.

## Mots Clef

Classification supervisée, objets 3D, apprentissage par transfert, sélection d'attributs.

## Abstract

In this article, we propose an improvement of an algorithm dealing with classification of 3D objects, using transfer learning, in order to face the lack of data, a frequent problem in learning systems.

The main idea is to directly transfer some sub-parts which are attributes builder from one domain to another in order to keep the advantages of interpretability and flexibility of the original method.

## Keywords

Supervised classification, 3D objects, transfer learning, feature selection.

## 1 Introduction

Pour l'entreprise pétrolière Total, lors de l'étude géologique d'un sous-sol potentiellement riche en hydrocarbures, il est nécessaire de bien comprendre sa structure. C'est un élément fondamental de la chaîne de traitement de l'information en géosciences, dont la qualité et l'efficacité d'exécution se répercuteront sur le reste de l'étude. C'est en effet au cours de cette étape que le géologue peut obtenir, pour les décideurs, les éléments dont ces derniers

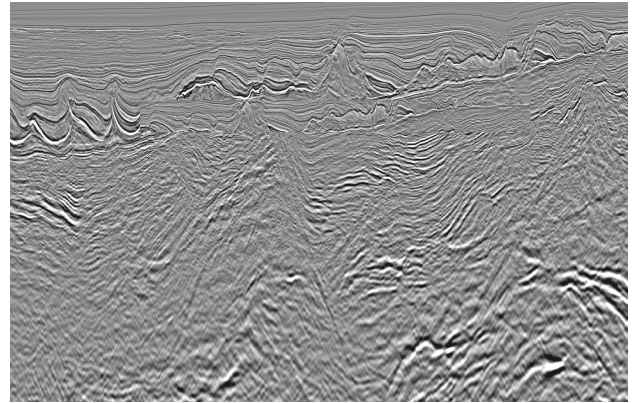


FIGURE 1 – Image sismique

ont besoin lors des choix de permis d'exploitation.

Pour cela, les géologues créent, à partir d'images sismiques reconstituées grâce à des ondes acoustiques envoyées dans le sol (Figure 1), un géo-modèle 3D censé représenter les principales couches et failles de la zone (Figure 2). Le géo-modèle, composé de surfaces 3D correspondant aux limites entre les couches sédimentaires, ne décrit que la structure géométrique du sous-sol.

Ce dernier est ensuite, en fonction des phénomènes que l'on y trouve, classé par le géologue selon les ensembles pertinents auxquels il se rattache. Les ensembles dont on parle correspondent à la présence ou à l'absence d'un certain phénomène géologique, comme par exemple des failles (normales, inverses, ...) ou des phénomènes d'interaction entre différentes couches (subduction, ...). Il s'agit donc, pour un géologue, de déterminer l'absence ou la présence de phénomènes géologiques locaux.

Jusqu'à maintenant, cette classification, comme le reste de l'étude, est réalisée manuellement par un géologue d'expérience qui connaît la zone. Cela représente une

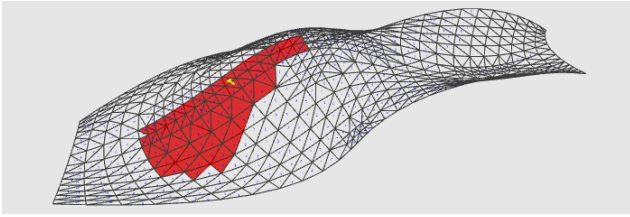


FIGURE 2 – Maillage triangulaire 3D récupéré pour chaque couche géologique

charge de travail conséquente, qu'une automatisation peut grandement faciliter. De plus, certains géologues disposent de moins d'expérience, tant dans le métier en lui-même que dans les personnes les plus utiles à contacter, entraînant une mauvaise évaluation des données en présence.

Néanmoins, afin de permettre une adoption par des professionnels de méthodes changeant leur manière de travailler, il est nécessaire, en plus de les convaincre par les performances des modèles présentés, de leur fournir un moyen d'appréhender de façon interprétable la méthode et/ou les résultats de cette dernière.

L'algorithme de classification supervisé de formes 3D de [8] est une première approche du problème issue de la notion de shapelets de [10], en utilisant des sous-parties discriminantes de la structure étudiée. Basé sur un système de classification supervisé traditionnelle, les résultats sont convaincants.

Pourtant, il nous est apparu comme essentiel de chercher à rendre ce système moins dépendant des données spécifiques disponibles comme base d'apprentissage de la classification souhaitée. Pouvoir réutiliser d'autres classifications, dont les probabilités de distribution des données et des classes ne seraient pas trop éloignées de celle visée, est un bon moyen d'y parvenir. A cette fin, il est proposé une adaptation de cette méthode afin de pouvoir améliorer la performance du classifieur en se reposant sur de l'expérience accumulée dans d'autres contextes, ce par transfert de connaissances.

Dans cet article, après une présentation de la méthode d'origine en Section 2, un état de l'art sur le transfert de connaissances est présenté en Section 3, puis en Section 4 la méthode proposée, avant que la Section 5 ne donne les perspectives de cette dernière.

## 2 Approche existante

En classification supervisée, certaines structures de données nécessitent un traitement particulier, impossible à réaliser par les méthodes classiques. Dans le cas des objets 3D, les méthodes actuelles ne permettent pas, sans connaissance préalable, de comprendre quelles sous-parties d'un objet peuvent être discriminantes dans le cadre de la présence du phénomène étudié, et pourquoi elles le sont. En effet, ces méthodes font le plus souvent

appel à des descripteurs [1], qui extraient des particularités de l'objet, ce qui tend à caler l'objet sur un modèle global, et donc peut ne pas tenir compte de sous-parties discriminantes.

Un certain nombre de conditions pour une nouvelle méthode se posent alors, avec notamment, pour ne citer que les plus importantes :

- comment représenter les phénomènes locaux que l'on souhaite détecter ?
- comment classifier les objets 3D ?
- comment tenir compte des sous-parties comme attributs essentiels ?
- comment permettre à l'utilisateur de comprendre le choix du classifieur ?

L'avantage évident de la méthode des shapelets appliquée aux séries temporelles est qu'il est possible de comparer des éléments de tailles très variées sans connaissance *a priori* sur ces derniers. Ce ne sont plus les éléments dans leur globalité qui peuvent être utiles, mais la juxtaposition de quelques-unes des sous-parties de ce dernier, sous-parties pertinentes possédant des caractéristiques différentes, qui, ensemble, permettent de déduire les classes d'appartenance.

L'idée maîtresse est de chercher à adopter une méthode, efficace pour les séries temporelles, à un autre type de données que sont les objets 3D constitués de maillages triangulaires de points. En utilisant des descripteur locaux [1] [3] et des mesures de distances de répartition [4] [7], les résultats s'avèrent supérieurs à l'état de l'art, et permettent en plus une visualisation des extraits les plus pertinents utilisés lors de cette classification.

Le but est de classifier les formes selon la présence ou non d'angles, le processus est représenté sur la Figure 3, et se décompose comme suit :

1. on extrait aléatoirement des sous-parties ;
2. on les trie par pertinence ;
3. on utilise les plus discriminantes pour classer de nouveaux objets.

De la méthode de [13], on adapte la technique des shapelets appliquée aux séries temporelles combinée à l'approche de [11] dans laquelle la sélection aléatoire de sous-parties potentiellement discriminantes est effectuée (sélectionner, pour ces tests, l'ensemble des candidats possibles devient bien trop long en termes de temps de calcul).

Dans l'algorithme 1, nommé 3DRESC pour *3D Random Extraction of Sub-parts for Classification*, on a :

- *Extraction(objet, tailles, nombres)* qui extrait d'un objet 3D un certain nombre de sous-parties de tailles prédéfinies ;
- *CalculDescripteur(sous – sur face)* qui calcule le descripteur pour chaque point de la sous-surface ;

**Algorithme 1 : 3DRESC [8]**

```

1  Entrées :
2  Objets : objets 3D; Classes : classes correspondantes;
   k : nombre d'attributs souhaité; tailles : tailles des
   extraits; nombres : nombre d'extraits pour chaque taille
   ;
3  Sorties :
4  SousParties : Liste de sous-parties discriminantes
   d'objets 3D ;
5  Début
6  i ← 0;
7  pour Objet dans Objets faire
8  |   i ← i + 1;
9  |   j ← 0;
10 |   candidats ←
11 |   |   Extraction(Objet, tailles, nombres);
12 |   pour candidat dans candidats faire
13 |   |   j ← j + 1;
14 |   |   DescObjet3D[i][j] ←
15 |   |   |   CalculDescripteur(candidat);
16 |   fin
17 fin
18 obj1 ← 0;
19 pour Objet1 dans Objets faire
20 |   obj1 ← obj1 + 1;
21 |   obj2 ← 0;
22 |   pour Objet2 dans Objets faire
23 |   |   obj2 ← obj2 + 1;
24 |   |   si Objet1 différent de Objet2 alors
25 |   |   |   Distances ← min(CalculDistances(
26 |   |   |   |   DescObjet3D[obj1], DescObjet3D[obj2]);
27 |   |   |   Gains ←
28 |   |   |   |   CalculGains(Distances, Classes);
29 |   |   |   ListeGain[i] ← max(Gains);
30 |   fin
31 fin
32 ListeGain ← Ordonner(ListeGain);
33 SousParties ← ListeGain[1, ..., k];
34 Fin

```

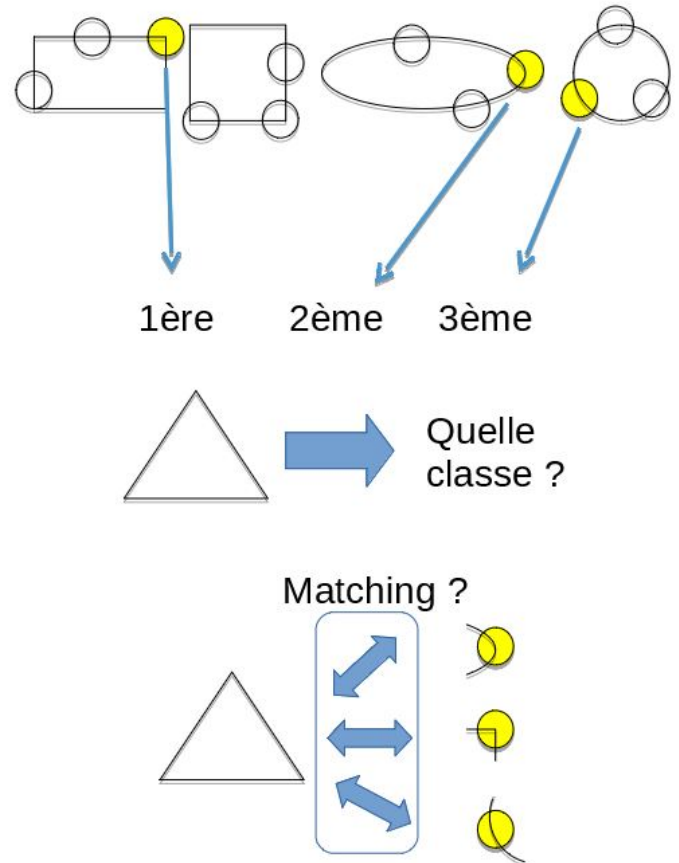


FIGURE 3 – Classification des formes

- *CalculDistances*(*Desc1*, *Desc2*) qui calcule, à partir de deux descripteurs de 2 objets, la similitude de leurs histogrammes de répartition de valeurs.
- *CalculGains*(*distances*, *classes*) qui calcule par la formule *f-stat* le gain d'information de l'attribut (créé par une sous-surface particulière) correspondant pour la classification souhaitée.

En présence de données en quantité suffisante, cette méthode s'est avérée utile aux géologues, que ce soit pour la classification en elle-même que pour la compréhension des zones choisies car pertinentes. Les performances des techniques de l'état de l'art ont bien été surpassées, et l'interprétabilité par la visualisation de données pertinentes a permis de plus facilement intégrer la méthode à la chaîne de traitement existante en géosciences.

### 3 L'apprentissage par transfert

Une hypothèse des méthodes traditionnelles d'apprentissage automatique est que les données d'apprentissage et les données de test sont issues du même domaine, de sorte que l'espace des attributs en entrée et la distribution des données sont les mêmes. Cependant, dans beaucoup de scénarios d'apprentissage supervisé, cette hypothèse ne

tient pas.

Il est donc nécessaire de créer des méthodes d'apprentissage performantes formées à partir de données plus facilement récupérables, voir simplement les seules disponibles. Ce concept est appelé apprentissage par transfert, ou *transfer learning* [2] [6], et est illustré par la Figure 4.

Pour un domaine d'apprentissage particulier, si l'on considère les notations de [9], alors :

- $D = \{\chi, P(X)\}$ , avec  $\chi$  l'espace d'attributs et  $X = \{x_1, x_2, \dots, x_n\} \in \chi$ ;
- $T = \{\Upsilon, f(\cdot)\}$  avec  $\Upsilon$  un espace de *labels*, et  $f(\cdot)$  une fonction de prédiction.

Nous pouvons par ailleurs décomposer l'apprentissage par transfert en deux ensembles, que sont, selon [12] :

1. Si  $X_S = X_T$  alors il est défini comme un apprentissage par transfert homogène. La majorité des solutions d'apprentissage par transfert homogène utilisent l'une des trois principales stratégies qui consistent à corriger, entre les données source et cible, la différence de distribution marginale, la différence de distribution conditionnelle, ou à la fois les différences de distribution marginale et conditionnelle.
2. Sinon, on parle d'apprentissage par transfert hétérogène. La majorité des solutions d'apprentissage par transfert hétérogènes sont axées sur l'alignement des espaces d'entrée des domaines source et cible, avec comme hypothèse que les distributions de domaines restent les mêmes.

Dans notre domaine, et dans ces deux cas, la principale façon de procéder reste l'extraction d'un modèle de cartographie des éléments ou des distributions d'un ou plusieurs domaines (appelé(s) domaine(s) source(s)) vers un autre (appelé domaine cible). Ce faisant, nous perdons notre capacité initiale qui consistait à pouvoir visualiser les raisons qui ont amené notre algorithme à donner un résultat particulier.

## 4 L'adaptation de la méthode initiale à un contexte de transfert de connaissances

### 4.1 Description de l'approche

Supposons que l'on dispose d'un domaine source unique avec beaucoup d'exemples d'apprentissage (en fait une tâche de classification réalisée sur des données particulières), et un domaine cible avec seulement quelques données déjà annotées (dans notre cas, environ 25% du nombre total des objets dans le domaine cible). Ce cas d'étude est illustré sur la Figure 4.1. En utilisant la méthode précédemment proposée par [8], nous n'apprenons que sur les quelques exemples disponibles du domaine

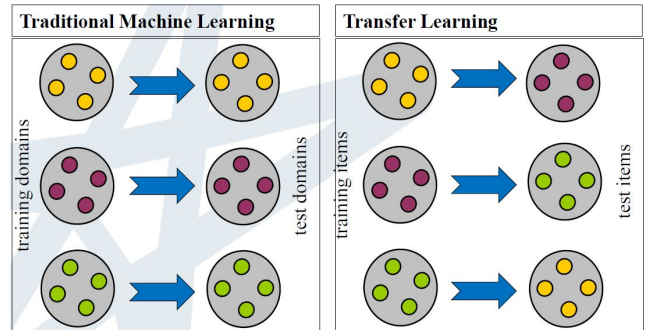


FIGURE 4 – Une définition intuitive du transfer learning - [9]

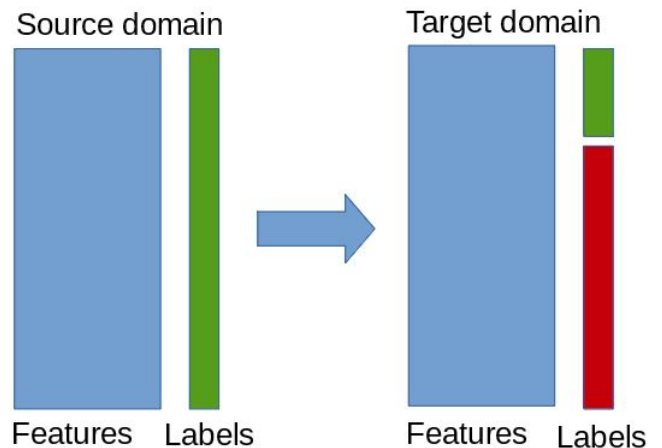


FIGURE 5 – Notre problème d'apprentissage par transfert : en vert, les éléments labellisés, et en rouge ceux qui ne le sont pas



cible et prédisons le reste. Cependant, avec le transfert de connaissances, le domaine source devient précieux, car il permettrait de sélectionner également les sous-parties qui peuvent rester pertinentes en passant du domaine source vers le domaine cible et de les ajouter aux sous-parties déjà sélectionnées grâce la méthode précédente (donc exclusivement issues du domaine cible).

C'est la raison pour laquelle il est proposé de transférer simplement certaines des sous-parties les plus pertinentes extraites d'un domaine vers un autre, après avoir vérifié que la pertinence de cette sous-partie dans le domaine cible est bonne. Ces sous-parties transférées restent les constructeurs d'attributs sur lesquels l'apprentissage est réalisé.

Le processus se déroule comme suit :

1. tout d'abord, la méthode classique de sélection des sous-parties pertinentes dans le domaine cible est réalisée ;
2. parallèlement, la même chose est faite pour le domaine source ;
3. ensuite, les sous-parties pertinentes du domaine source sont testées au sein du domaine cible afin d'évaluer leur capacité à y être réutilisées ;
4. enfin, les sous-parties de l'extraction originale du domaine cible et les sous-parties choisies dans le domaine source sont toutes deux utilisées pour classer l'ensemble de données du le domaine cible.

Soient  $O_1, \dots, O_n$  l'ensemble des objets labellisés qui constituent notre base d'apprentissage :

- $\forall O_{i=1, \dots, n}$ , on extrait aléatoirement  $m$  sous-surfaces (pour chaque objet  $O_i$ ) que sont  $S_{i,1}, \dots, S_{i,m}$ .
- On considère un descripteur d'objet 3D particulier  $Desc$ , et une distance  $Dist$  (qui peut être une distance entre répartition de valeurs, pour le cas des descripteurs d'objets).  $\forall i \in (1, \dots, n)$  et  $\forall j \in (1, \dots, m)$ , on calcule le descripteur  $Desc(S_{i,j})$  de la sous-surface  $S_{i,j}$ .

Puis, pour chaque sous-surface  $S_{i,j}$  à évaluer :

- $\forall i' \in (1, \dots, n) \setminus i$  et  $\forall j' \in (1, \dots, m)$ , on calcule la proximité définie comme

$$Prox(S_{i,j}, S_{i',j'}) = Dist(Desc(S_{i,j}), Desc(S_{i',j'}))$$

- Les degrés correspondant à ses valeurs d'imbrication dans chacun des objets  $O_1, \dots, O_n$  valent :  $\forall i' \in (1, \dots, n)$ ,

$$Degré(s_{i,j}, O_{i'}) = \inf_{l=1, \dots, m} (Prox(S_{i,j}, S_{i',l}))$$

Ce degré de matchage ou d'appartenance d'une sous-surface à un objet permet d'établir une forme de proximité entre la sous-surface étudiée et chaque objet. C'est à partir de ces valeurs que le score de pertinence de chacune des sous-surfaces est évalué, afin de ne sélectionner que les plus discriminantes.

Ce score, proposé la première fois par [5], et dont la pertinence pour le domaine étudié a été confirmée par [8], est la mesure de  $f-stat$  :

$$f-stat = \frac{\frac{1}{C-1} \sum_{i=1}^C (\bar{D}_i - \bar{D})^2}{\frac{1}{n-C} \sum_{i=1}^C \sum_{d_j \in D_i} (d_j - \bar{D}_i)^2} \quad (1)$$

avec  $C > 1$  le nombre de classes,  $n$  le nombre de sous-parties ( $n > C$ ),  $\bar{D}_i$  la moyenne des degrés entre les surfaces de classe  $i$  et la sous-partie concernée, et  $\bar{D}$  la moyenne des degrés entre la sous-partie et les objets 3D.

Par la suite, les  $k$  premiers candidats selon ce critère sont utilisés pour le calcul des attributs,  $k$  est défini expérimentalement en moyenne à environ 0.1% du nombre total d'extraits.

Dans notre cas, pour le domaine source, tous les extraits ne seront pas conservés tels quels pour la construction d'attributs, il est en effet d'abord nécessaire de les tester pour s'assurer de leur pertinence. La valeur  $f-stat$  calculée pour chaque sous-surface du domaine source au sein du domaine cible doit être au minimum du même ordre de grandeur que les extraits du domaine cible qui ont été sélectionnés.

Pour chaque extrait dont le test est positif, ce dernier est ajouté au domaine cible (avec ses propres extraits), et le transfert se base donc sur l'extrait constructeur d'attribut, et non pas l'attribut lui-même.

Soit  $O_S$  l'ensemble des sous-parties sélectionnées au sein du domaine source,  $O_C$  ceux du domaine cible, et  $Transfert_{Extrait}$  une fonction de seuil qui prend comme arguments une sous-surface extraite et un domaine possédant une classification. La fonction renvoie une réponse positive si l'extrait a une pertinence (fonction  $f-stat$ ) supérieure au seuil choisi, et est donc ajouté (transféré) aux extraits  $O_C$ , et sera ignoré sinon.

Avec, dans l'algorithme 2 :

- $Capacit.Separation$ , fonction qui calcule le score d'une sous-partie selon le test  $f-stat$  ;
- $Seuil_{TF}$ , le seuil défini empiriquement comme étant le minimum de pertinence de sous-parties sélectionnées de l'ensemble d'entraînement, et la capacité de séparation utilisée reste la même que dans l'algorithme d'origine, c'est-à-dire le  $f-stat$ .

Les sous-parties finalement utilisées pour la construction des attributs sont l'union de celles issues du domaine cible

**Algorithme 2** : Transfert d’extraits pertinents

---

```

1 Entrées :
2  $Objets_{Source}$  : objets 3D du domaine source ;
    $Objets_{Cible}$  : objets 3D du domaine cible ;
    $SousParties$  : sous-parties issues du domaine source ;
    $Classes$  : classes correspondantes pour chaque objet ;
    $Seuil_{TF}$  : limite arbitraire de pertinence pour le transfert
   des sous-parties du domaine source ;
3 Output :
4  $SousPartiesPertinentes$  : Liste de sous-parties
   pertinentes d’objets 3D utilisable pour le transfert
   learning, donc venant du domaine source ;
5 Début
6  $SousPartiesPertinentes_{Transferees} \leftarrow 0$ ;
7 pour  $SousPartie \in SousParties$  faire
8   |  $Pertinence \leftarrow$ 
   |    $CapacitSeparation(SousPartie, Classes, Objets_{Cible})$ ;
9   | si  $Pertinence \geq Seuil_{TF}$  alors
10  | |  $SousPartiesPertinentes_{Transferees} \leftarrow$ 
   | |  $SousPartiesPertinentes \cup SousPartie$ ;
11 fin
12 Fin

```

---

et celles sélectionnées au sein du domaine source qui sont aussi pertinentes dans le domaine cible, à savoir :

$$SousParties_{Cible} \cup SousParties_{Pertinentes_{Transferees}}$$

De cette manière, nous obtenons des résultats qui comparent les approches en termes de performances. Ils ont été rassemblés à partir de tests multiples en utilisant la sélection aléatoire des objets, les seules contraintes étaient d’avoir deux classes dans chaque domaine.

L’algorithme global correspondant est *3DRESC-TF*, pour *3D Random Extraction of Sub-parts for Classification with Transfer Learning*.

## 4.2 Expérimentations

*3DRESC-TF* est testé avec des données libres du *Princeton Shape Benchmark*<sup>1</sup> (PSB), dont les données proviennent de multiples ensembles d’objets 3D déjà existants. Ces derniers sont ceux qui, en termes de formats et de types, se rapprochent le plus des géo-modèles 3D dont nous disposons en entreprise. Les classes les plus fines sont retenues (3 niveaux étant possibles), afin de rajouter encore davantage de contraintes et de difficultés à la prédiction lors des tests.

Le classifieur utilisé une fois les attributs récupérés est celui des forêts aléatoires du package Python *scikit-learn*. Les tests sont effectués sur un processeur *Intel Core i7 vPro* avec 16 Go de RAM.

Ainsi, nous réutilisons la connaissance de plusieurs domaines disponibles et pertinents dans le but d’améliorer

notre prédiction. Dans le cas de domaines très différents, ce transfert d’apprentissage pourrait faire face à un léger apprentissage dit négatif (baisse de la performance en raison d’informations transmises mais non pertinentes, voir nuisibles), et donc devrait logiquement rester inutilisées, ce qui reviendrait à n’employer que l’algorithme d’origine.

Face au potentiel déséquilibre de classe au sein de ces nouvelles expériences, nous choisissons d’utiliser la *AUC* (aire sous la courbe ROC - *Receiver Operating Characteristic*) pour évaluer les performances du modèle.

La première chose à souligner, et qui s’avère intuitive, est le fait que plus le nombre d’éléments dans le domaine cible est faible, plus notre méthode d’apprentissage par transfert sera bénéfique vis-à-vis du processus d’apprentissage classique. En effet, moins d’exemples signifie un apprentissage plus difficile, et donc l’intérêt pour le transfert des connaissances venant d’autres domaines augmente fortement. Nous gardons néanmoins l’avantage de la méthode originale qui est d’être en mesure d’apprendre sur un jeu de données restreint.

Cependant, on peut mentionner le fait qu’en ajoutant simplement de nouvelles fonctionnalités du domaine source, on améliorerait statistiquement la performance si on utilise un bon algorithme de classification (forêts aléatoires, par exemple).

Pour supprimer ce possible biais dans nos résultats et pouvoir observer le seul impact de notre proposition d’apprentissage de transfert, nous avons tout simplement décidé de comparer notre méthode d’apprentissage mécanique classique en utilisant les 200 extraits par objet plus le nombre de sous-parties transférées possibles dans l’apprentissage par transfert, avec l’approche d’apprentissage par transfert elle-même. Ainsi, les résultats de la Table 1 confirment que cet algorithme d’apprentissage par transfert améliore clairement la manière classique de calculer notre méthode.

Ensuite, et toujours en essayant de donner à l’utilisateur une explication des résultats donnés, nous proposons de faire une visualisation d’une tâche d’apprentissage par transfert particulière.

A partir de deux classifications d’avions représentés sous forme de maillages 3D, la première étant le domaine source et la seconde le cible, nous donnons le nombre d’éléments utilisés comme apprentissage dans le domaine cible et montrons les sous-parties les plus pertinentes (transférées et du domaine cible).

La première classification, c’est-à-dire le domaine source, vise à séparer le biplan des avions commerciaux, tandis que la seconde différencie les chasseurs des planeurs.

En utilisant ces deux domaines pour la visualisation, nous obtenons finalement les sous-parties suivantes comme les plus pertinentes extraites du domaine source et réutilisées au sein de la cible.

1. <http://shape.cs.princeton.edu/benchmark/> consulté le 05/02/2017

TABLE 1 – Classification en fonction du nombre d’objets pour les domaines cible et source

Nombre d’objets du domaine source	Nombre d’objets du domaine cible	Sans apprentissage par transfert	Avec apprentissage par transfert
60	40	$0.616 \pm 0.037$	$0.642 \pm 0.034$
	30	$0.581 \pm 0.049$	$0.617 \pm 0.047$
	20	$0.545 \pm 0.055$	$0.583 \pm 0.065$
80	40	$0.611 \pm 0.035$	$0.666 \pm 0.036$
	30	$0.580 \pm 0.041$	$0.625 \pm 0.041$
	20	$0.548 \pm 0.054$	$0.601 \pm 0.060$
100	40	$0.616 \pm 0.032$	$0.640 \pm 0.034$
	30	$0.579 \pm 0.045$	$0.615 \pm 0.048$
	20	$0.544 \pm 0.058$	$0.581 \pm 0.069$

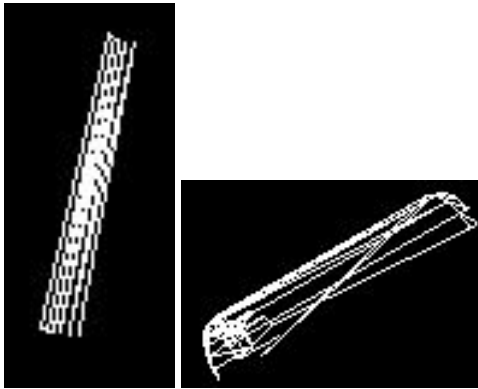


FIGURE 6 – A gauche, l’aile d’un planeur et à droite, le cockpit d’un biplan

Dans un cas particulier, à partir des formes d’un planeur et d’un biplan, les sous-parties récupérées pour le transfert car pertinentes sont données dans la Figure 6.

Nous remarquons que la première sous-partie représente des ailes très minces, et il est logique de la transférer, car ces ailes très particulières sont caractéristiques d’un planeur (longues et minces), et si on les compare à un chasseur à réaction, permettent d’améliorer la classification en utilisant ce critère.

Le second est un peu plus centré sur le corps de l’avion. On peut aisément reconnaître un cockpit de biplan, qui apporte une connaissance sur les moyens utilisés par le classifieur pour reconnaître la classe d’appartenance.

Testée sur des données réelles de Total, la méthode a encore besoin, pour être validée, de voir son intérêt confirmé par le métier, que la mise en pré-production devrait permettre d’ici peu.

## 5 Conclusions et perspectives

Dans cet article, nous avons proposé un nouvel algorithme, appelé 3DRESC-TF, qui vient compléter de précédents travaux sur la classification d’objets 3D. Afin de faire face à un manque de données d’apprentissage, cette adaptation

visée à transférer de la connaissance issue d’autres apprentissages, dans le but d’améliorer le taux de classification.

Proposant des performances supérieures aux méthodes de l’état de l’art, nous valorisons de cette manière les apprentissages réalisés par le passé, afin, d’une part, d’améliorer la classification en cours, et d’autre part de pouvoir faire face à des données différentes de celles utilisées pour l’apprentissage.

Néanmoins, les sous-parties extraites et récupérées pour la visualisation ne sont à l’heure actuelle pas beaucoup plus interprétables avec cette méthode que celles obtenues par la méthode d’origine, et la pertinence d’une utilisation systématique de cette nouvelle fonctionnalité devra être, à l’avenir, étudiée pour le cas d’application concrète qui est le notre à Total.

## Références

- [1] L. A. Alexandre. 3D Descriptors for Object and Category Recognition : a Comparative Evaluation. In *Workshop on Color-Depth Camera Fusion in Robotics at the IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS)*, Vilamoura, Portugal, October 2012.
- [2] W. Dai, Q. Yang, G.-R. Xue, and Y. Yu. Boosting for transfer learning. In *Proceedings of the 24th International Conference on Machine Learning, ICML ’07*, pages 193–200, New York, NY, USA, 2007. ACM.
- [3] Q. V. Dang. *Similarités dans des modèles BRep paramétriques : detection et applications*. PhD thesis, Université de Toulouse, 2014.
- [4] S. Kullback and R. A. Leibler. On information and sufficiency. *The Annals of Mathematical Statistics*, 22(1) :79–86, march 1951.
- [5] J. Lines, L. M. Davis, J. Hills, and A. Bagnall. *A Shapelet Transform for Time Series Classification*. KDD ’12. ACM, New York, NY, USA, 2012.
- [6] J. Lu, V. Behbood, P. Hao, H. Zuo, S. Xue, and G. Zhang. Transfer learning using computational intelligence : A survey. *Knowledge-Based Systems*, 80 :14 – 23, 2015. 25th anniversary of Knowledge-Based Systems.

- [7] K. Matusita. Decision rules, based on the distance, for problems of fit, two samples, and estimation. *Annals of Mathematical Statistics*, 26(4) :631–640, Dec. 1955.
- [8] F. Meunier, C. Marsala, L. Castanié, and B. Conche. Classification d’objets 3d par extraction aléatoire de sous-parties discriminantes pour l’étude du sous-sol en prospection pétrolière. In *Extraction et Gestion des Connaissances*, pages 225–236, 2017.
- [9] S. J. Pan and Q. Yang. A survey on transfer learning. *IEEE Trans. on Knowl. and Data Eng.*, 22(10) :1345–1359, Oct. 2010.
- [10] A. Refregier. Shapelets : I. a method for image analysis. *Mon. Not. Roy. Astron. Soc.*, 338 :35, 2001.
- [11] X. Renard, M. Rifqi, and M. Detyniecki. Random-shapelet : an algorithm for fast shapelet discovery. In *Proceedings of the IEEE International Conference on Data Science and Advanced Analytics (DSAA ’2015)*, pages 1–10, Oct. 2015.
- [12] K. Weiss, T. M. Khoshgoftaar, and D. Wang. A survey of transfer learning. *Journal of Big Data*, 3(1) :9, 2016.
- [13] L. Ye and E. Keogh. Time Series Shapelets : A New Primitive for Data Mining. In *Proceedings of the 15th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD ’09*, pages 947–956, New York, NY, USA, 2009. ACM.

# Affectation distribuée d'individus à des activités avec des préférences additivement séparables

M. Morge

A. Nongaillard

Univ. Lille, CNRS, Centrale Lille, UMR 9189 - CRISAL  
Centre de Recherche en Informatique Signal et Automatique de Lille, F-59000 Lille, France

{Maxime.Morge,Antoine.Nongaillard}@univ-lille.fr

## Résumé

*Nous souhaitons proposer un réseau social numérique afin que les utilisateurs forment des groupes pour pratiquer ensemble des activités. Dans cet article, nous introduisons un modèle formel de formation de coalitions correspondant à ce cas d'usage. Nous nous restreignons à des préférences additivement séparables pour proposer un algorithme distribué. Nous démontrons que le résultat est Pareto-optimal. Nos expérimentations montre que la solution atteinte par notre algorithme est meilleure que celle obtenue via les techniques classiques de recherche locale et que sa distribution permet d'accélérer son exécution.*

## Mots Clef

Résolution distribuée de problème, négociation, comportement d'agents, problème d'appariement.

## Abstract

*We aim at providing a social network such that users form groups to practice together some activities. In this paper, we introduce a formal framework for coalition formation which is suitable for our usecase. We restrict ourselves to additively separable preferences in order to propose a distributed matching algorithm. We demonstrate that its outcome is a Pareto-optimum. Our experiments shows we reach a better outcome than the classical local search techniques and that the distribution of our algorithm speeds up its runtime.*

## Keywords

Distributed problem solving, Negotiation, Agent behavior, Matching problem.

## 1 Introduction

Les Systèmes Multi-Agents (SMA) constituent un paradigme de premier ordre pour l'analyse, la conception et

l'implémentation de systèmes composés d'entités autonomes en interaction. Afin de concevoir des environnements socio-techniques médiateurs ou simulés, les SMA permettent de modéliser des boucles de rétroaction entre acteurs hétérogènes dont les prises de décision locales font émerger des phénomènes globaux. Un des défis auquel la communauté SMA fait face consiste à faciliter l'élicitation des préférences des utilisateurs.

Nos travaux s'inscrivent dans le projet « Partir Ensemble » (PartENS). Ce projet vise à comprendre et modéliser les dynamiques de rétroaction se produisant dans un collectif en interaction à la fois au sein d'un réseau social virtuel et d'un réseau social réel. Le cas d'usage sur lequel nous nous concentrons ici concerne un collectif de seniors âgés de 60 à 70 ans qui entrent en relation afin d'effectuer des activités ponctuelles. Ce terrain d'expérimentation implique plusieurs milliers d'individus inscrits au CCAS (Caisse Centrale d'Activités Sociales du Personnel) de la ville de Lille. L'objectif est ici de maximiser les activités organisées afin d'améliorer la cohésion sociale et lutter contre l'isolement chez certains seniors. Nous souhaitons proposer un réseau social numérique afin que les utilisateurs forment des groupes pour pratiquer ensemble des activités. Ce système a pour objectif de suggérer à chaque utilisateur des individus avec lesquels pratiquer ces activités.

Dans cet article, nous introduisons un modèle formel de formation de coalitions correspondant à notre cas d'usage. Un ensemble d'individus doivent être appariés pour pratiquer une des activités qui leur ait proposé en fonction des préférences des individus pour leurs alter egos et pour les activités. Nous nous restreignons à des préférences additivement séparables afin de proposer un algorithme d'affectation. Nous démontrons que le résultat de notre algorithme est valide et Pareto-optimal. En adoptant le modèle d'acteur [2], nous sommes en mesure de distribuer cet algorithme. Nos expérimentations montre que la solution atteinte par notre algorithme est de meilleure qualité que celles obtenues avec des techniques classiques de re-

cherche locale. De plus, sa distribution permet d'accélérer son exécution (jusqu'à 3,5 fois).

La section 2 compare notre approche aux travaux scientifiques connexes. Nous introduisons notre problème d'affectation dans la section 3. Nous proposons un algorithme d'affectation dans la section 4. La distribution de cet algorithme est présentée sous la forme de comportements d'agents dans la section 5. La section 6 exhibe nos résultats expérimentaux. Finalement, nous synthétisons nos travaux puis nous dressons quelques perspectives (cf Section 7).

## 2 Travaux connexes

La théorie du choix social vise à construire et à analyser des processus pour la décision collective où un ensemble d'agents sélectionnent ou classent conjointement un sous-ensemble d'alternatives parmi celles disponibles. Contrairement à l'Économie, l'Informatique se préoccupe des questions algorithmiques afin de rendre opérationnelles ces méthodes. Nous nous focalisons ici sur un problème d'appariement en particulier.

Dans le problème de formation de coalition hédonique introduit par [4], chaque joueur est caractérisé par une relation de préférences sur l'ensemble des coalitions auxquelles il peut appartenir. Notre problème est une spécialisation de ce problème générale. Nous pouvons le représenter comme un jeu hédonique en associant à chaque activité un joueur qui préfère ne pas être surchargé et en munissant chacun des autres joueurs des préférences agrégeant celle d'un individu sur les activités et sur les alter egos. Nous considérons ici que ces deux préférences sont indépendantes, puis nous les agrégeons dans une fonction d'utilité. Sous cette hypothèse, notre configuration a des propriétés structurelles utiles qui se distinguent de celle d'un jeu hédonique générique. Le fait que les activités soient des points focaux [11] et la représentation succincte des préférences nous permet de proposer un algorithme pour atteindre de "bon" appariements.

Le problème de sélection d'activité de groupe a été proposé dans [3]. Dans un tel problème, chaque agent participe à au plus une activité et ses préférences portent uniquement sur les activités et elles dépendent du nombre de participants à l'activité. C'est une généralisation d'un jeu hédonique anonyme. Même si ce problème a été étendu dans [7] pour prendre en compte les relations entre individus, ces relations sont encodées par un réseau social, i.e. un graphe non orienté où les nœuds correspondent aux individus et les arêtes représentent des liens de communication entre eux. À l'inverse, les individus sont munis, dans notre problème, d'une relation de préférence sur leur alter egos.

Le problème des hôpitaux-internes a été introduit dans [6]. Ce problème peut être considéré comme une spécialisation du problème de formation de coalition hédonique où un ensemble de résidents doivent être affectés à des hôpitaux étant donnés les préférences des résidents vis-à-vis

des hôpitaux et les préférences des hôpitaux pour les résidents. Le problème HR a fait l'objet de nombreuses extensions [8]. À notre connaissance, aucune n'est adaptée à notre cas d'usage.

Comment un agent évalue ses penchants vis-à-vis des alternatives? Nous optons ici pour des fonctions d'utilité, i.e. des préférences cardinales. Par souci de simplification, nous supposons que l'évaluation des activités et celle des groupes sont comparables. Bien que leur expressivité soit limitée, nos préférences sont linéaires par rapport au nombre d'individus et d'activités.

Quelle est la « meilleure » solution à un problème de choix collectif? Dans la littérature, on trouve principalement deux types de règles qui dérivent le choix collectif de la satisfaction individuelle des agents : les premières se basent sur les propriétés désirables de la solution (e.g. la stabilité) alors que les secondes se basent sur l'agrégation de la satisfaction individuelle des agents (e.g. le bien-être utilitaire). Dans cet article, nous adoptons la seconde approche car aucun concept de la première approche ne convient.

Comment atteindre un appariement maximisant le bien-être utilitaire? Comme affirmé dans cet article, ce problème peut être NP-difficile. C'est la raison pour laquelle, nous pouvons considérer les algorithmes de Problème d'Optimisation sous Contraintes Distribuées (DCOP) ou les méthodes de recherche locale (LST). Il a été montré dans [5] que les algorithmes de DCOP ne passent pas nécessairement à l'échelle pour les problèmes d'appariement. Nous montrons ici que les techniques de LST ne conviennent pas car la fonction à optimiser possède de nombreux optima locaux. C'est la raison pour laquelle nous avons adopté une méthode de résolution multi-agents et notamment une modélisation multi-niveaux comme préconisé par [9] où chaque agent « activité » représente un groupe d'individus.

## 3 Problème IA

Notre objectif ici est de proposer un modèle formel de formation de coalitions adapté à notre cas d'usage. Nous introduisons ici le problème des individus/activités (IA). Dans une instance de problème IA, les individus privilégient les activités qui leur plaisent avec les partenaires qu'ils apprécient.

**Définition 1** (Problème IA). *Un problème individus/activités (IA) de taille  $(m, n)$ , avec  $m \geq 1$  et  $n \geq 1$ , est un couple  $IA = \langle I, A \rangle$  avec  $m$  individus et  $n$  activités, où :*

- $A = \{a_1, \dots, a_n\}$  est un ensemble de  $n$  activités. Chaque activité  $a_j$  peut accueillir au plus  $c_j$  membres ;
- $I = \{1, \dots, m\}$  est un ensemble de  $m$  individus. À chaque individu  $i$  correspond,
  1. une préférence sur les activités, i.e. une relation réflexive, complète et transitive  $\succeq_i$  sur les activités  $A \cup \{\theta\}$ , notamment l'activité nulle (notée

$\theta$ ). La relation de préférence stricte correspondante est notée  $\triangleright_i$ ,

2. une préférence purement hédonique, i.e. une relation de préférence réflexive, transitive et complète  $\succsim_i$  sur l'ensemble des groupes auxquels il appartient  $G(i) = \{G \subseteq I; i \in G\}$ . La relation de préférence stricte correspondante est dénotée  $\succ_i$ .

L'activité nulle correspond à ne rien faire. Les préférences sur les groupes sont des préférences purement hédoniques car la satisfaction d'un individu dépend uniquement des autres membres de son groupe et non pas des autres groupes constitués par les individus restants.

Nous souhaitons former des coalitions autour des activités.

**Définition 2** (Coalition). Soit  $IA = \langle I, A \rangle$  un problème  $IA$ . Une **coalition** est un couple  $C = \langle a, G \rangle$  où  $a \in A \cup \{\theta\}$  et  $G \subseteq I$ . L'activité de la coalition  $C$ , éventuellement l'activité nulle, est  $a_C$  avec une capacité<sup>1</sup>  $c_C$  et le groupe  $G_C$ . Une coalition non-vide  $C$  est telle que  $G_C \neq \emptyset$  et  $C$  est pour  $i$  si  $i \in G_C$ .

Nous espérons que le nombre d'individus est considérablement plus important que le nombre d'activités ( $m \gg n$ ).

**Définition 3** (Appariement). Un **appariement**  $M$  pour le problème  $IA = \langle I, A \rangle$  est représenté par les fonctions  $a_M : I \rightarrow A \cup \{\theta\}$  et  $g_M : I \rightarrow \mathcal{P}(I)$  telles que :

$$\forall i \in I, a_M(i) \in A \cup \{\theta\} \quad (1)$$

$$\forall i \in I, i \in g_M(i) \subseteq I \quad (2)$$

$$\forall i \in I, a_M(i) = \theta \Rightarrow g_M(i) = \{i\} \quad (3)$$

$$\forall i \in I \forall j \in g_M(i), a_M(j) = a_M(i) \quad (4)$$

$$\forall i, j \in I, i \neq j \wedge a_M(i) = a_M(j) \neq \theta \Rightarrow g_M(i) = g_M(j) \quad (5)$$

L'affectation d'un individu est une activité, éventuellement l'activité nulle (cf équation 1). Chaque individu est associée à un groupe auquel il appartient (cf équation 2). Tous les individus qui sont affectés à l'activité nulle sont seuls (cf équation 3). Tous les individus associés les uns aux autres ont la même activité (cf équation 4) et réciproquement tous les individus qui sont affectés à la même activité, à l'exception de l'activité nulle, sont associés les uns aux autres (cf équation 5).

Pour simplifier les notations, nous introduisons la fonction de poste d'un appariement  $M$  qui retourne l'ensemble des individus affectés à chaque activité :

$$p_M : A \cup \{\theta\} \rightarrow \mathcal{P}(I) \\ p_M(a) = \{i \in I; a_M(i) = a\} \quad (6)$$

Les postes d'une activité peuvent être vide. Si  $a_M(i) = \theta$ , on dit que  $i$  n'est pas affecté. Une activité  $a \in A$  est :

1. l'activité nulle a une capacité infinie.

surchargée si  $|p_M(a)| > c_a$ ; ii) pleine si  $|p_M(a)| = c_a$ ; iii) sous-chargée sinon. Un appariement est dit **valide** si aucune activité n'est surchargée.

Un appariement est une structure de coalition.

**Proposition 1** (Partition). Soit  $M$  un appariement pour le problème  $IA = \langle I, A \rangle$ .

$$\forall a \in A \cup \{\theta\}, \langle a, p_M(a) \rangle \text{ est une coalition} \quad (7)$$

$$\bigcup_{a \in A \cup \{\theta\}} p_M(a) = I \wedge \quad (8)$$

$$\forall a_i, a_j \in A \cup \{\theta\} p_M(a_i) \cap p_M(a_j) = \emptyset$$

**Preuve 1** (Partition). Des définitions 2 et 3.

$C_M(i)$  est la coalition dans  $M$  qui contient  $i$ .

Chaque individu évalue les coalitions et donc les appariements en fonction du groupe auquel il appartient et de son activité.

**Définition 4** (Rationalité). Soit  $IA = \langle I, A \rangle$  un problème  $IA$ .

- Une coalition  $C$  pour  $i$  est **individuellement rationnelle** pour  $i$  ssi :

$$(a_C \succeq_i \theta) \wedge G_C \succsim_i \{i\} \quad (9)$$

- Un appariement  $M$  est **individuellement rationnel** ssi :

$$\forall i \in I, (a_M(i) \succeq_i \theta) \wedge g_M(i) \succsim_i \{i\} \quad (10)$$

- Soient  $C$  et  $C'$  deux coalitions qui sont individuellement rationnelles pour  $i$ .

- L'individu  $i$  **préfère**  $C$  à  $C'$  (dénoté  $C \succsim_i C'$ ) ssi :

$$a_C \succeq_i a_{C'} \wedge G_C \succsim_i G_{C'} \quad (11)$$

- L'individu  $i$  **préfère strictement**  $C$  à  $C'$  (noté  $C \succ_i C'$ ) ssi :

$$C \succsim_i C' \wedge (a_C \triangleright_i a_{C'} \vee G_C \succ_i G_{C'}) \quad (12)$$

- Soient  $M$  et  $M'$  deux appariements valides pour  $IA$ . L'individu  $i$  **préfère**  $M$  à  $M'$  (noté  $M \succsim_i M'$ ) ssi :

$$C_M(i) \succsim_i C_{M'}(i) \quad (13)$$

La relation de préférence stricte sur les appariements est notée  $\succ_i$ . Un individu choisit une coalition et donc un appariement tel que son activité est préférée à l'activité nulle et il préfère ses partenaires plutôt que de rester seul (cf équations 9 et 10). Il préfère une coalition à une autre s'il préfère l'activité et le groupe de la première (cf équation 11). Un individu préfère un appariement à un autre s'il préfère sa coalition dans le premier (cf équation 13). La relation de préférence sur les appariements valides est réflexive, transitive et éventuellement partielle.

Une première propriété désirable pour l'évaluation d'un appariement est la stabilité du cœur.

**Définition 5** (Stabilité du cœur). Soit  $M$  un appariement pour le problème  $IA = \langle I, A \rangle$ . Une coalition non-vide  $C$  bloque l'appariement  $M$  ssi :

1. l'activité n'est pas surchargée :  $|G_C| \leq c_C$
2. tous les individus de la coalition préfèrent celle-ci à leur affectation dans  $M$  :

$$\forall i \in G_C, C \succsim_i C_M(i) \quad (14)$$

3. au moins un individu de la coalition préfère strictement celle-ci à son affectation selon  $M$  :

$$\exists i \in G_C C \succ_i C_M(i) \quad (15)$$

L'appariement  $M$  est **stable de cœur** s'il est valide et qu'il n'existe pas de coalition bloquante :

$$\forall C \subseteq (A \cup \{\theta\}) \times (2^I \setminus \emptyset) \quad (16)$$

ce n'est pas le cas que  $C$  bloque  $M$

Les individus dans une coalition bloquante souhaite se désolidariser et former leur propre coalition, ce qui rend le partitionnement instable.

Un appariement est Nash stable si aucun individu n'a intérêt à dévier unilatéralement de sa coalition vers une autre (éventuellement vide).

**Définition 6** (Nash stabilité). Soit  $M$  un appariement pour le problème  $IA = \langle I, A \rangle$ . L'appariement  $M$  est **Nash stable** s'il est valide, rationnel et :

$$\forall i \in I \forall a \in A, a \neq a_M(i) \Rightarrow p_M(a) = c_a \vee C_M(i) \succsim_i \langle a, p_M(a) \cup \{i\} \rangle \quad (17)$$

Un appariement Nash stable est à l'abri des mouvements individuels car les activités sont pleines ou la coalition de chaque individu est au moins aussi bonne que les autres.

Comme l'illustre l'exemple suivant, même si la rationalité individuelle d'un appariement est une condition nécessaire pour être soit stable de cœur soit Nash stable, aucune de ces stabilités n'implique l'autre. De plus, un problème  $IA$  n'a pas nécessairement d'appariement stable de cœur ou Nash stable.

**Exemple 1** (Stabilité). Considérons le problème  $IA$  avec 3 individus (1, 2 et 3) et une activité  $a$  telle que  $a \succeq_i \theta$  avec  $i \in \{1, 2, 3\}$ .

Nous supposons que la capacité de  $a$  est 2 et que les préférences sociales sont circulaires :

- $\{1, 2\} \succ_1 \{1\} \succ_1 \{1, 3\}$  ;
- $\{2, 3\} \succ_2 \{2\} \succ_2 \{1, 2\}$  ;
- $\{1, 3\} \succ_3 \{3\} \succ_3 \{2, 3\}$ .

Cette instance n'a pas d'appariement stable de cœur car l'appariement  $M_1$  (avec  $p_{M_1}(a) = \{1, 2\}$ ) est bloqué par la coalition  $\langle a, \{2, 3\} \rangle$ , l'appariement  $M_2$  (avec  $p_{M_2}(a) = \{2, 3\}$ ) est bloqué par  $\langle a, \{1, 3\} \rangle$  et  $M_3$  (avec  $p_{M_3}(a) = \{1, 3\}$ ) est bloqué par  $\langle a, \{1, 2\} \rangle$ . De plus, il n'y a pas d'appariement Nash stable. En particulier,  $M_1$ ,  $M_2$  et  $M_3$

ne sont pas individuellement rationnels.

Supposons maintenant que la capacité de  $a$  est 3 et que l'individu 3 est "indésirable", i.e. les coalitions avec 3 ne sont pas individuellement rationnelles pour les autres :

- $\{1, 2\} \succ_1 \{1\} \succ_1 \{1, 2, 3\} \succ_1 \{1, 3\}$  ;
- $\{1, 2\} \succ_2 \{2\} \succ_2 \{1, 2, 3\} \succ_2 \{2, 3\}$  ;
- $\{1, 2, 3\} \succ_3 \{2, 3\} \succ_3 \{1, 3\} \succ_3 \{3\}$  ;

L'appariement  $M_1$  tel que  $p_{M_1}(a) = \{1, 2\}$  est stable de cœur mais n'est pas Nash stable. L'appariement  $M_2$  tel que  $p_{M_2}(a) = \{3\}$  est Nash stable mais n'est pas stable de cœur.

Une autre propriété désirable pour évaluer un appariement est la Pareto-optimalité.

**Définition 7** (Pareto-optimal). Soient  $M$  et  $M'$  deux appariements valides pour le problème  $IA = \langle I, A \rangle$ .  $M'$  **domine au sens de Pareto**  $M$  ssi :

$$\forall i \in I, C_{M'}(i) \succsim_i C_M(i) \quad (18)$$

$$\exists i \in I, C_{M'}(i) \succ_i C_M(i) \quad (19)$$

Un appariement est **Pareto-optimal** s'il n'est pas dominé au sens de Pareto.

Un appariement est Pareto-optimal s'il n'existe pas d'alternative pour laquelle tous les agents sont dans une position meilleure ou équivalente.

La stabilité de cœur est une condition suffisante mais pas nécessaire pour la Pareto-optimalité.

**Proposition 2** (Pareto-optimum). Tous les appariements stable de cœur sont Pareto-optimaux.

**Preuve 2** (Pareto-optimum). Nous prouvons par contradiction qu'un appariement stable de cœur est Pareto-optimal. Nous supposons que  $M$  est un appariement stable de cœur qui n'est pas Pareto-optimal. Donc, il existe un appariement  $M'$  qui domine  $M$  au sens de Pareto. Considérons la coalition  $C_{M'}(i)$  avec  $i$  qui satisfait l'équation 19. On vérifie :

1.  $M'$  est valide ;
2. l'équation 14 par l'équation 18 ;
3. l'équation 15 par l'équation 19.

En conséquence, par la définition 5, on conclut que  $C_{M'}(i)$  bloque  $M$  qui est en contradiction avec notre hypothèse.

Dans notre exemple précédent, quand la capacité de  $a$  est 2 et que les préférences sociales sont circulaires, tous les appariements où l'activité est pleine sont Pareto-optimaux. Quand la capacité est 3 et que l'individu 3 est indésirable, l'appariement où l'activité est pleine est également Pareto-optimal.

Même si la stabilité est une propriété souhaitable, il n'existe pas nécessairement une telle solution. En revanche, la Pareto-optimalité semble ne pas être discriminante. Une autre façon d'évaluer la qualité d'un appariement réside dans la notion de bien-être social. Pour cela,



nous supposons que les individus ont des préférences cardinales. De plus, dans un problème IA, chaque individu évalue ses préférences vis-à-vis des  $2^{m-1}$  groupes. La représentation de telles préférences est exponentielle en espace. En revanche, la représentation de préférences additivement séparables est linéaire par rapport au nombre d'individus.

**Définition 8** (Additivement Séparable IA). *Soit  $IA = \langle I, A \rangle$  un problème IA de taille  $(m, n)$ . Le problème est **additivement séparable** (ASIA) si chaque individu  $i \in I$  est muni :*

1. d'une fonction de valuation  $v_i : A \cup \{\theta\} \rightarrow [-1; 1]$  représentant ses préférences sur les activités, éventuellement nulle ;
2. d'une fonction de valuation  $w_i : I \setminus \{i\} \rightarrow [-1; 1]$  représentant ses préférences sur les partenaires potentiels.

La fonction d'utilité pour un individu  $i$  est la fonction  $u_i : G(i) \times A \cup \{\theta\} \rightarrow [-1; 1]$  définie telle que :

$$\forall g \in G(i) \forall a \in A \cup \{\theta\} \quad (20)$$

$$u_i(g, a) = \frac{\frac{1}{m-1} \sum_{j \in g, j \neq i} w_i(j) + v_i(a)}{2} \quad (21)$$

Nous supposons que les préférences sur les individus et les préférences sur les activités sont comparables. En particulier, l'utilité pour un individu d'être seul dépend uniquement de son évaluation de l'activité. En outre, nous supposons que plus l'individu a de partenaires, plus il est satisfait car nous visons à améliorer la cohésion sociale et éviter l'isolement.

Dans la suite, nous ne considérons pas que les utilités sont définies pour correspondre aux préférences comme dans [1] mais nous supposons un accès direct aux utilités des individus car cela sera le cas dans notre application pratique.

Nous adoptons ici l'approche utilitaire inspirée par Bentham. En d'autres termes, notre objectif est de maximiser la somme des utilités individuelles.

**Définition 9** (Bien-être). *Soit  $IA = \langle I, A \rangle$  un problème ASIA de taille  $(m, n)$ . Le **bien-être utilitaire** d'un appariement  $M$  est défini tel que :*

$$U_I(M) = \frac{1}{m} \sum_{i \in I} u_i(g_M(i), a_M(i)) \quad (22)$$

Plus le bien-être est important, meilleur est l'appariement.

**Exemple 2** (problème ASIA). *Considérons l'exemple précédent où la capacité de l'activité est 3 et l'individu 3 est indésirable. On définit les fonctions de valuations telles que  $v_1(a) = v_2(a) = v_3(a) = 0$ ,  $w_1(2) = w_2(1) = \frac{1}{2}$ ,  $w_1(3) = w_2(3) = -1$ ,  $w_3(1) = \frac{1}{2}$  et  $w_3(2) = 1$ . En conséquence, les préférences sur les groupes sont conformes à la formulation précédente. L'appariement où les individus 1 et 2 sont affectés à l'activité maximise le bien-être.*

## 4 Résolution

Afin de maximiser le bien-être utilitaire, nous pouvons considérer la programmation quadratique, i.e. une méthode d'optimisation avec un modèle mathématique représentable par une fonction quadratique. Une instance de problème ASIA peut être modélisé par  $n \times m$  variables  $x_{ia} \in \{0, 1\}$  telles que  $x_{ia} = 1$  si l'individu  $i$  est affecté à l'activité  $a$  et  $x_{ia} = 0$  sinon,  $m$  contraintes  $\sum_{a \in A} x_{ia} \leq 1$  représentant l'exclusion mutuelle de l'affectation d'un même individu aux activités et  $n$  contraintes  $\sum_{i \in I} x_{ia} \leq c_a$  garantissant la validité de l'appariement. En considérant les fonctions de valuations  $(w_i(j))_{i \in I}$  et  $(v_i(a))_{i \in I, I}$ , la fonction objectif correspondant au bien-être utilitaire, qui est à maximiser, est :

$$\sum_{a \in A} \sum_{i \in I} x_{ia} (v_i(a) + \frac{1}{m-1} \sum_{j \neq i} w_i(j) \times x_{ja})$$

Quand on écrit sous une forme standard ce problème, i.e. en minimisant une fonction objectif de la forme  $\frac{1}{2} x^T Q x^T + c^T x$  avec une matrice symétrique  $Q$ , on peut remarquer que cette dernière n'est pas nécessairement définie positive et donc le problème peut être NP-difficile [10].

C'est la raison pour laquelle nous proposons ici un algorithme (cf algorithme 1) qui permet de calculer un "bon" appariement. Initialement, tous les individus sont seuls, affectés à l'activité nulle et libres. Tour à tour, chaque individu libre  $i$  considère l'activité qui l'attire et qu'il préfère  $a$ . Si aucun autre individu n'est affecté à cette activité,  $i$  est affecté. Sinon, l'algorithme tente d'améliorer le bien-être de ce groupe, éventuellement en renvoyant les individus dont la présence contribue le moins au bien-être du groupe. Si l'affectation de  $i$  n'améliore pas le bien-être du groupe, alors  $i$  doit concéder et donc considérer l'activité attractive suivante. Les individus qui sont remplacés par  $i$  doivent concéder. Si la capacité de  $a$  n'est pas atteinte alors le groupe peut croître (ligne 21). Un agent qui est rejeté par toutes les activités qu'il considère comme attrayantes reste seul et il est définitivement inactif. Un algorithme d'approximation consiste à exclure qu'un seul individu à chaque étape (ligne 20).

On peut remarquer que notre algorithme (éventuellement d'approximation) retourne toujours un appariement valide.

**Proposition 3** (Terminaison). *Notre algorithme (d'approximation) appliqué à ASIA se termine et l'appariement retourné est valide.*

**Preuve 3** (Terminaison). *Soit  $IA = \langle I, A \rangle$  un problème ASIA. Nous considérons l'invariant de boucle  $\sum_{i \in I} |\text{concessions}(i)| + |\text{Free}|$ . Cet invariant est positif ou nul. Il décroît strictement à l'issue de chacune des boucles car :*

1. un individu affecté est retiré de *Free* ;
2. un individu qui n'est pas affecté concède jusqu'à être éventuellement affecté à l'activité nulle ;

---

**Algorithme 1** : Calcul d'un appariement pour un problème ASIA
 

---

```

1 . Entrées :  $IA = \langle I, A \rangle$ 
   Sorties : un appariement  $M$ 
2  $Free = I$ ;
3 pour chaque  $i \in I$  faire
4    $concessions(i) = A.SortWith(v_i(\_) > v_i(\_) > 0)$ ;
5    $a_M(i) = \theta$  ;
6    $g_M(i) = \{i\}$  ;
7 tant que  $Free \neq \emptyset$  faire
8   pour chaque  $i \in Free$  faire
9     si  $concessions(i) = \emptyset$  alors  $Free \setminus = \{i\}$  ;
10    sinon
11       $a = concessions(i).head$  ; //  $a$  est
12      l'activité préférée
13       $g = p_M(a)$  ;
14       $g' = g \cup \{i\}$  ;
15      si  $g = \emptyset$  alors
16        /* les postes de  $a$  sont
17        vides et donc  $i$  et
18        affecté */
19         $a_M(i) = a$  ;
20         $g_M(i) = \{i\}$  ;
21         $Free \setminus = \{i\}$  ;
22      sinon
23         $u_{max} = -\infty$  ;
24         $bg = \emptyset$  ;
25         $SG = \{sg \subsetneq g' ; sg \neq \emptyset\}$  ;
26        /* éventuellement
27         $SG = \{sg \subsetneq g' ; |sg| = |g'| - 1\}$  */
28      si  $c_a > |g|$  alors  $SG \cup = g'$  ;
29      /*  $g$  peut croître */
30      pour chaque  $sg \in SG$  faire
31         $u = \sum_{k \in sg} u_k(sg, a)$  ;
32        si  $u > u_{max}$  alors
33           $u_{max} = u$  ;
34           $bg = sg$  ;
35        /*  $bg$  est le meilleur groupe */
36      pour chaque  $j \in bg$  faire  $g_M(j) = bg$  ;
37      pour chaque  $j \in g \setminus bg$  faire
38        /*  $j$  est désaffecté */
39         $a_M(j) = \theta$  ;
40         $g_M(j) = \{j\}$  ;
41         $Free \cup = \{j\}$  ;
42         $concessions(j) =$ 
43         $concessions(j).tail$  ;
44      si  $i \in bg$  alors
45        /*  $i$  est affecté */
46         $a_M(i) = a$  ;
47         $g_M(i) = bg$  ;
48         $Free \setminus = \{i\}$  ;
49      else
50        /*  $i$  est rejeté */
51         $concessions(i) =$ 
52         $concessions(i).tail$  ;
53    retourner  $M$ 

```

---

3. un individu, qui est désaffecté, concède et au moins un autre individu est affecté (donc retiré de  $Free$ );

L'appariement obtenu est valide car les activités ne sont jamais surchargées.

Le resultat de notre algorithme exact est Pareto-optimal.

**Proposition 4** (Pareto-optimalité). *Notre algorithme appliqué à un problème ASIA retourne un appariement Pareto-optimal.*

**Preuve 4** (Pareto-optimalité). *Soit  $IA = \langle I, A \rangle$  un problème ASIA. Nous montrons par contradiction que le résultat de notre algorithme est un appariement Pareto-optimal  $M$ . Nous supposons que  $M$  est dominé au sens de Pareto par un appariement valide  $M'$ . Par l'équation 19, il existe un individu  $i$  tel que  $C_{M'}(i) \succ_i C_M(i)$ . Par l'équation 12,  $C_{M'}(i) \succ_i C_M(i)$  et :*

- soit  $g_{M'}(i) \succ_i g_M(i)$  et  $a_{M'}(i) = a_M(i)$ . Comme  $C_{M'}(i) \succ_j C_M(i)$  et par l'équation 11,  $g_{M'}(i) \succ_j g_M(i)$  pour tous les partenaires de  $i$ . Par l'équation 20,  $\sum_{j \in g_{M'}(i)} u_j(g_{M'}(i), a_{M'}(i)) > \sum_{j \in g_M(i)} u_j(g_M(i), a_M(i))$ . Ceci est une contradiction avec notre algorithme (ligne 27) qui calcule les postes qui maximisent le bien-être du groupe ;
- soit  $a_{M'}(i) \triangleright_i a_M(i)$  et donc  $a_{M'}(i)$  précède  $a_M(i)$  dans  $concessions(i)$ . Selon l'algorithme,  $i$  a été rejeté ou désaffecté par  $a_{M'}(i)$ . En conséquence, il existe un individu  $j \in C_{M'}(i)$  tel que  $C_M(j) \succ_j C_{M'}(j)$ . Ceci est une contradiction avec l'équation 18.

**Exemple 3** (Algorithme). *Considérons l'exemple 2. Notre algorithme opère les affectations/désaffectations suivantes sur la liste d'individus libres (3, 2, 1) :*

1. 3 est affecté car les postes de  $a$  sont initialement vide ;
2. 2 est affecté car la capacité de 3 n'est pas atteinte et  $\{2, 3\}$  est le meilleur sous-groupe de  $\{2, 3\}$  ;
3. 1 remplace 3 car  $\{1, 3\}$  est le meilleur sous-groupe de  $\{1, 2, 3\}$  ;
4. 3 concède et il est définitivement affecté à l'activité vide.

L'appariement obtenu est Pareto-optimal et maximise le bien-être utilitaire.

On peut noter que dans le cas général notre algorithme ne maximise pas nécessairement le bien-être utilitaire.

## 5 Comportement d'agents

Nous considérons ici le modèle de programmation concurrente par passage de messages asynchrones, appelé modèle d'acteur, proposé dans [2]. Selon cette perspective, les primitives sont les agents et les événements. Un agent représente un programme indépendant qui s'exécute sur

son propre processeur. Un événement consiste en la création d'un agent ou l'émission/la réception d'un message. On peut noter qu'un tel système est distribué car le délai de transmission des messages est arbitraire mais non négligeable. On suppose que les canaux de communication sous-jacents sont fiables (un message est délivré une et une seule fois) et que les messages peuvent arriver dans un ordre différent de celui de l'émission.

Afin de proposer un solveur distribué à partir de ce modèle, nous distinguons trois types d'agents :

1. l'agent « solveur » qui crée les autres agents et enregistre les affectations ;
2. les agents « individu » tous munis du même comportement mais avec des préférences qui les distinguent ;
3. les agents « activité » tous munis du même comportement mais qui gèrent des coalitions et des capacités différentes.

Le comportement de l'agent « solveur » consiste à : i) créer les agents ; ii) lancer la résolution ; iii) enregistrer les affectations et les désaffectations ; iv) renvoyer l'appariement quand tous les individus sont affectés.

Le comportement de l'agent « individu » consiste à créer sa liste de concessions puis à se proposer à l'activité préférée qui l'attire. Quand l'agent est affecté ou désaffecté, il en informe l'agent « solveur ». Lors d'une désaffectation, l'agent « activité » attend une confirmation avant d'affecter un nouvel individu afin que l'appariement ne soit pas retourné prématurément par l'agent « solveur ». Si un agent « individu » devient libre, il concède, i.e. il se propose à l'activité suivante qui l'attire jusqu'à être affecté à l'activité nulle.

Le comportement de l'agent « activité » est décrit par l'automate déterministe représenté dans la figure 1. Quand une proposition est reçue, elle est acceptée si le groupe courant est vide (dans l'état *Available*). Sinon, l'agent traite les propositions une à une (dans l'état *Raising*) en identifiant le sous-groupe qui maximise le bien-être utilitaire. Si la capacité est atteinte, les sous-groupes de taille  $c_a$  sont considérés. Dans le cas contraire, le nombre de postes peut croître. Si le proposant n'est pas dans le nouveau groupe, alors la proposition est rejetée. Sinon la proposition est acceptée quand le membre écarté a confirmé la désaffectation (dans l'état *Firing*). Quand une proposition a été traitée, l'agent « activité » est prêt à évaluer les propositions suivantes, éventuellement celles stockées.

## 6 Évaluation empirique

Nos expérimentations ont pour objectifs d'évaluer la qualité de la solution atteinte par notre algorithme et l'accélération due à sa distribution.

Nous avons implémenté notre prototype avec le langage de

programmation Scala<sup>2</sup> et la boîte à outils Akka<sup>3</sup>. Cette dernière, en s'appuyant sur le modèle d'acteur [2], nous permet de réduire la distance entre les spécifications du SMA et son implémentation. Afin d'envisager un grand nombre d'individus, nous considérons notre algorithme d'approximation.

Nous avons implémenté un algorithme de recherche locale pour le comparer à notre algorithme. Cet algorithme de montée en gradient qui, à partir d'un appariement valide aléatoire cherche à améliorer le bien-être utilitaire, est itératif. Deux appariements sont voisins s'ils sont identiques à l'exception d'un individu affecté à l'activité nulle ou ayant changé d'activité. Si la nouvelle activité est pleine, alors tous les échanges d'individus sont envisagés.

Nous considérons des problèmes *ASIA* constitués de  $n$  activités et  $m$  individus. Pour une même instance de problème, toutes les activités ont le même capacité ( $m/n$ ). Pour chaque jeu de paramètres ( $n$  et  $m$ ), nous avons généré (pseudo)-aléatoirement 100 instances.

Dans un premier temps, nous avons comparé le bien-être utilitaire de l'appariement obtenu par notre algorithme avec celui obtenu par recherche locale. La figure 2 présente les bien-utilitaires moyens obtenus pour chaque jeu de paramètres (avec  $2 \leq n \leq 10$  et  $2 \times n \leq m \leq 10 \times n$ ). Le bien-être obtenu par notre algorithme surpasse sans conteste celui la recherche locale. En effet, le bien-être utilitaire est pour le problème *ASIA* une fonction qui possède de nombreux optima locaux.

Dans un second temps, nous avons comparé les temps d'exécution des versions distribuée et centralisée de notre algorithme d'approximation. La figure 2 présente les temps d'exécution moyen obtenus pour chaque jeu de paramètres (avec  $2 \leq n \leq 10$  et  $2 \times n \leq m \leq 400$ ). Alors que l'algorithme centralisé est plus rapide quand le nombre d'individus est faible ( $\sim 40$ ), son temps d'exécution croît rapidement avec le nombre d'individus (18 ms pour 100 individus et 10 activités) alors que le temps d'exécution de la version distribuée est moindre (9 ms dans ce dernier cas). De plus, le temps d'exécution de l'algorithme de montée en gradient est très important (1660 ms pour 100 individus et 10 activités). On peut s'attendre à un temps d'exécution encore plus important si on adopte une méthode de recherche locale du type recuit simulé sans pour autant garantir d'obtenir un optimum global.

En résumé, la solution atteinte par notre algorithme semble de bonne qualité. De plus, ce dernier est distribuable ce qui permet d'accélérer (jusqu'à 3,5 fois) son exécution.

## 7 Conclusions

Nous avons introduit ici le problème générique individus/activités où des individus doivent être affectés à des

2. <http://www.scala-lang.org/>

3. <http://akka.io>

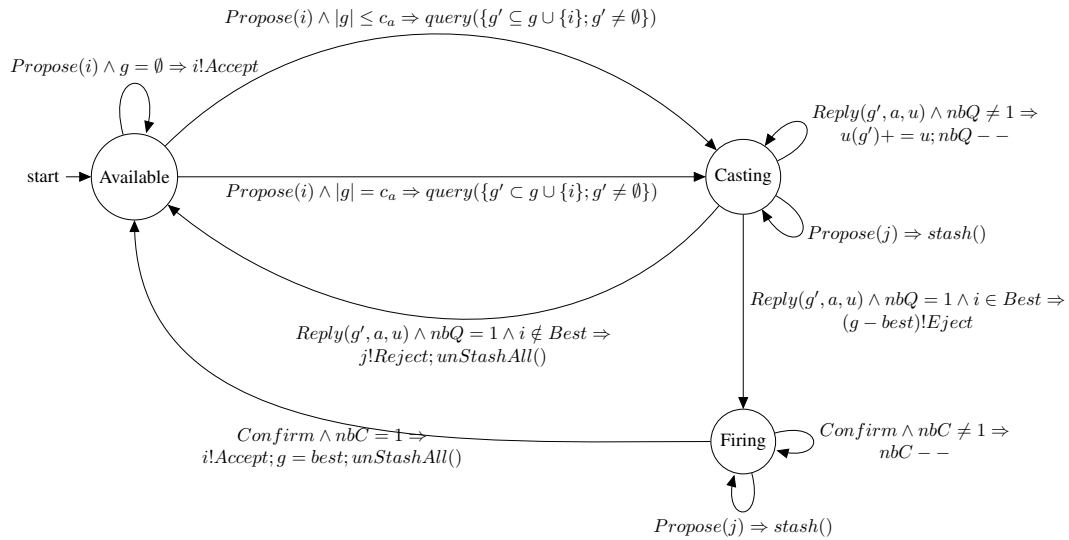


FIGURE 1 – Comportement d’un agent « activité »

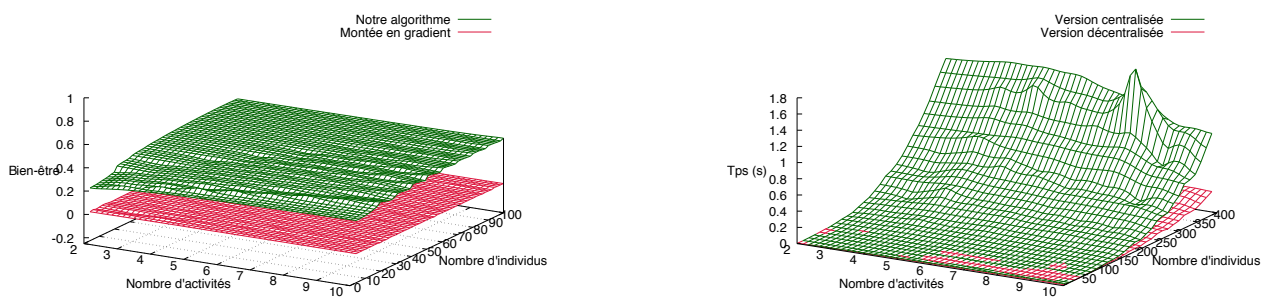


FIGURE 2 – Bien-être utilitaire (à gauche) et temps d’exécution (à droite)

activités en privilégiant les activités qui leur plaisent avec les partenaires qu'ils apprécient. Bien que souhaitable, la stabilité d'une solution n'est pas garantie. À l'inverse, la Pareto-optimalité n'est pas discriminante. C'est la raison pour laquelle nous avons adopté des préférences additivement séparables afin d'évaluer la qualité d'une solution à l'aide du bien-être utilitaire des individus. De plus, la représentation des préférences additivement séparable est linéaire par rapport au nombre d'individus et d'activités. Maximiser le bien-être utilitaire peut être NP-difficile. C'est la raison pour laquelle nous avons proposé un heuristique où les individus se proposent aux activités qu'ils préfèrent quitte à concéder. Les individus qui contribuent le moins au bien-être du groupe sont désaffectés et concèdent. Nous avons montré que cet algorithme renvoie systématiquement un appariement valide et Pareto-optimal. En adoptant le modèle d'acteur, nous avons distribué cet algorithme. La difficulté réside dans : i) la détection de l'arrêt (comme les messages peuvent arriver dans un ordre différent de l'envoi, un agent « activité » doit attendre la confirmation d'une désaffectation avant d'affecter un nouvel individu) ; ii) la synchronisation (comme l'acceptation d'un individu dépend du groupe, un agent « activité » doit traiter les propositions les unes après les autres). Nos expérimentations montrent que le bien-être utilitaire obtenu par notre algorithme surpasse celui atteint par recherche locale et que la distribution permet d'accélérer notre algorithme (jusqu'à 3,5 fois).

Dans de futurs travaux, nous souhaitons évaluer notre moteur d'appariement avec des données réelles dans le cadre du projet PartENS. Afin de proposer des appariements qui soient socialement plus juste, nous envisageons de modifier le critère local de décision de notre algorithme pour maximiser le bien-être égalitaire.

## Remerciements

Ce travail s'inscrit dans le projet PartENS soutenu par le programme chercheur-citoyen de la région Nord Pas de Calais. Nous remercions le comité de programme des JFSMA 2017 qui, par ses remarques, nous a permis d'améliorer cet article.

## Références

- [1] Boutilier, C., Caragiannis, I., Haber, S., Lu, T., Procaccia, A. D., and Sheffet, O. (2015). Optimal social choice functions : A utilitarian view. *Artificial Intelligence*, 227 :190–213.
- [2] Clinger, W. D. (1981). *Foundations of actor semantics*. PhD thesis, Massachusetts Institute of Technology.
- [3] Darmann, A., Elkind, E., Kurz, S., Lang, J., Schauer, J., and Woeginger, G. (2012). Group activity selection problem. In *Proc. of the 8th International Conference on Internet and Network Economics*, pages 156–169, Liverpool, UK. Springer Berlin Heidelberg.
- [4] Dreze, J. and Greenberg, J. (1980). Hedonic coalitions : Optimality and stability. *Econometrica*, 48 :987–1003.
- [5] Everaere, P., Morge, M., and Picard, G. (2012). Casanova : un comportement d'agent respectant la privacité pour des mariages stables et équitables. *RIA*, 26(5) :471–494.
- [6] Gale, D. and Shapley, L. S. (1962). College admissions and the stability of marriage. *The American Mathematical Monthly*, 69 :9–14.
- [7] Igarashi, A., Peters, D., and Elkind, E. (2017). Group activity selection on social networks. In *Proc. of AAAI*, pages 565–571.
- [8] Manlove, D. F. (2014). *Algorithmics of Matching Under Preferences*. World Scientific.
- [9] Nongaillard, A. and Picault, S. (2016). Modélisation multi-niveaux des problèmes d'affectation et d'appariement. In *JFSMA'16*, pages 75–84, Rouen, France. Cépaduès.
- [10] Sahni, S. (1974). Computationally related problems. *SIAM Journal on Computing*, 3(4) :262–279.
- [11] Schelling, T. C. (1980). *The strategy of conflict*. Harvard university press.



# Suivi et détection des idéations suicidaires dans les médias sociaux

Bilel Moulahi<sup>1</sup>

Jérôme Azé<sup>1</sup>

Sandra Bringay<sup>1,2</sup>

<sup>1</sup> LIRMM, Université de Montpellier, CNRS

<sup>2</sup> AMIS, Université Paul Valéry Montpellier

bilel.moulahi@lirmm.fr

## Résumé

*L'utilisation croissante des médias sociaux permet un accès sans précédent aux comportements, aux pensées et aux sentiments des individus. Nous nous intéressons ici à l'évolution des états émotionnels des individus captés au travers des services de microblogging de type Twitter. Notre objectif est de prédire l'apparition d'idéations suicidaires. Dans ce travail, nous avons mis en place une chaîne de traitements permettant d'extraire des caractéristiques à partir des messages reflétant l'état émotionnel. Puis, nous appliquons un modèle basé sur les Conditionnal Random Fields pour prédire un nouvel état. L'originalité de l'approche est de prendre en compte l'historique des états émotionnels pour prédire le nouvel état. Une expérimentation préliminaire nous a permis d'évaluer notre approche sur des cas réels d'utilisateurs de Twitter. Ces type d'approche permet de mieux comprendre les liens entre expressions dans les médias sociaux et idéations suicidaires ainsi que les transitions entre états émotionnels.*

## Mots Clef

Média sociaux, Suicide, Conditionnal Random Fields.

## Abstract

*The increasing use of social media allows an unprecedented access to individuals' behaviors, thoughts and feelings. In this paper, we are interested in the evolution of the emotional states of individuals captured through the microblogging service Twitter in order to predict suicidal ideations. We propose a pipeline to extract the characteristics of the messages describing the emotional state. Then, we apply a model based on the Conditional Random Fields to predict the new psychological state by taking into account the history of emotional states. Preliminary experiments allowed us to evaluate our approach on real cases of Twitter users. This type of approach helps to better understand the links between expressions in social media and suicidal ideations as well as transitions between emotional states.*

## Keywords

Example, model, template.

## 1 Introduction

Les médias sociaux sont de plus en plus utilisés par les professionnels de santé pour détecter et diagnostiquer des troubles dépressifs majeurs [1, 6, 14, 16, 20]. Les plateformes comme Twitter et Facebook facilitent l'auto-présentation sélective de comportements indésirables, tels que l'automutilation, l'anorexie, ainsi que l'expression d'émotions négatives liées aux idéations suicidaires, en particulier chez les jeunes.

Afin de mieux comprendre ces nouvelles pratiques, de nombreuses études comme celles de [4, 5] ont porté sur la recherche dans le discours des individus des références à la dépression et aux idéations suicidaires. Des auteurs comme [4, 5] soulignent une corrélation entre les taux de tentatives de suicide et le volume de messages liés aux idéations suicidaires publiés dans les médias sociaux. Si les médias sociaux peuvent affecter les individus en répandant des pensées suicidaires, ils peuvent aussi avoir un rôle positif en aidant ces individus à trouver du soutien moral. Par exemple, l'activité intense en ligne, notamment nocturne, est un signe précoce permettant d'anticiper une dégradation de l'état émotionnel d'un individu.

De méthodes efficaces sont désormais disponibles pour analyser les sentiments exprimés dans les réseaux sociaux [2, 10]. Plusieurs études récentes tirent partie de ces travaux pour la détection et la surveillance des idéations suicidaires [19, 7, 18] et d'états dépressifs [15, 9]. Toutefois, la plupart des méthodes de l'état de l'art prédisent les émotions véhiculées par les utilisateurs au niveau d'un message ou de l'individu. Elles ne prennent pas en compte l'évolution du comportement de l'individu. Or, les idées suicidaires sont incluses dans un continuum de séquences d'événements influençant les états émotionnels et qui peuvent éventuellement conduire à une tentative de suicide [1]. Étant donné la nature séquentielle des contenus produits par les individus dans les médias sociaux, nous utilisons dans ces travaux un modèle basé sur les Conditional Random Fields [11, 21] qui permet de capturer l'évolution de l'état émotionnel des individus au fil du temps, en tenant compte du contexte passé et de l'activité du moment. Les états émotionnels sont au préalable inférés à partir des messages, via une analyse de textes permettant d'identifier

les facteurs de risque [6].

Nous avons évalué notre approche sur un corpus de tweets annotés manuellement, publiés par des individus ayant exprimé des références au suicide. La collection d'individus a été validée par un psychiatre pour n'inclure que les utilisateurs ayant présenté de réels symptômes. Les résultats expérimentaux montrent que le système prédit correctement des séquences d'états mentaux.

Dans le reste de l'article, nous présentons un état de l'art succinct, puis décrivons notre approche pour la détection d'idéations suicidaires. Nous présentons et discutons les résultats, avant de conclure sur des perspectives.

## 2 Etat de l'art

En France, près de 10 000 personnes mettent fin à leurs jours chaque année, soit environ 25 par jour, selon le dernier rapport de l'OMS<sup>1</sup>. Deuxième cause de mortalité chez les 15-24 ans, après les accidents, le suicide est un fléau qui touche des adolescents souvent fragilisés par cette période charnière de la vie. Avec l'avènement des médias sociaux, les personnes à risque et notamment les jeunes, utilisent des outils comme Facebook, Twitter et Reddit pour exprimer des idéations suicidaires. Il est possible d'utiliser ces médias pour détecter de manière précoce les individus vulnérables et intervenir rapidement. En 2015, Facebook<sup>2</sup> a introduit un nouveau service permettant aux utilisateurs de rapporter un comportement suicidaire. Très récemment, ce service a évolué<sup>3</sup> pour permettre aux personnes qui visionnent un *live-stream* Facebook<sup>4</sup> d'interpeller son auteur ou de faire un signalement.

De nombreuses études ont porté sur les notes laissées par les individus avant un suicide. Ces notes ont été analysées en développant des classifieurs supervisés et non supervisés pour identifier les sujets discutés ainsi que les émotions exprimées par des personnes étant passées à l'acte [19, 17]. Plus récemment, plusieurs études se sont intéressées à l'évaluation des facteurs de risque suicidaires dans les médias sociaux avec l'objectif de mieux comprendre ou de prévenir le suicide en détectant les idéations suicidaires de manière précoce. Par exemple, le projet Durkheim<sup>5</sup> étudie les activités des anciens combattants américains sur Twitter, Facebook et LinkedIn. L'objectif de ce projet est d'identifier les marqueurs de comportements à risque. [18] ont développé des modèles de prédiction en utilisant les textes des notes. Les résultats montrent que les personnes qui se sont suicidées expriment souvent de la peur et une certaine agitation avant de passer à l'acte. Les modèles de prédiction proposés ont montré des taux d'exactitude

proche de 65%. [7] ont analysé les messages Twitter d'une jeune fille qui venait de se suicider, publiés les vingt-quatre heures précédant son décès. Ils ont trouvé une augmentation des émotions positives et un changement de la focalisation de soi à d'autres lorsque le moment du décès s'est approché. Les auteurs ont également étudié un éventail plus large de tweets. Pour cela, ils ont utilisé le logiciel Linguistic Inquiry and Word Count (LIWC)<sup>6</sup> pour identifier dans le discours, des mots porteurs d'émotions ainsi que des processus cognitifs. [20] ont utilisé un panel en ligne de 250 jeunes d'une vingtaine d'années, utilisant régulièrement Twitter, pour examiner l'association entre les tweets liés au suicide et les passages à l'acte. Les auteurs ont étudié les caractéristiques linguistiques de l'idéation suicidaire et ont identifié les marqueurs les plus fréquents. Par exemple, des phrases comme "*I want to commit suicide*" sont fortement associées aux tentatives de suicide, alors que des phrases suggérant une intention suicidaire, comme "*I want to die*" y sont moins associés. Contrairement aux techniques populaires d'apprentissage, [9] ont combiné plusieurs méthodes de TAL pour filtrer les utilisateurs de forums et identifier les symptômes de dépression. Ces auteurs ont mis en correspondance les questionnaires traditionnels de dépistage de la dépression avec un ensemble de termes associés aux symptômes. Ensuite, ils détectent ces termes dans les textes et en déduisent un score au niveau du message.

La littérature actuelle manque de modèles efficaces pour prédire les tentatives de suicide. Actuellement, peu d'approches intègrent l'évolution du comportement de l'individu. L'analyse porte généralement sur un message ou sur l'ensemble des messages d'un individu. L'analyse ne permet pas de prédire à quel moment une personne présente un risque suicidaire. [14] explorent une approche basée sur les concepts drift pour identifier un temps à risque. Une limite à leur approche est de ne pas expliciter les transitions entre les états émotionnels comme nous souhaitons le faire dans ce travail.

## 3 Un modèle basé sur le contexte pour le suivi et la détection des idéations suicidaires dans les médias sociaux

Dans cette section, nous reformulons le problème et décrivons le modèle utilisé pour monitorer les idéations suicidaires dans les médias sociaux. Étant donné une séquence de messages pouvant traiter de thèmes jugés à risque tels que la dépression, le suicide, l'automutilation ou l'anorexie, mais également contenir des thèmes sans rapport ou même des blagues dans un intervalle de temps très court, avec quelle précision pouvons-nous prédire qu'un individu présente un réel risque suicidaire ? Un modèle d'analyse de sentiments typique traite ce problème comme une tâche de classification multi-classes et prédit une étiquette pour

1. Observatoire national du suicide <http://www.who.int/topics/suicide/fr/>

2. <https://www.facebook.com/help/suicideprevention>

3. <https://newsroom.fb.com/news/2017/03/building-a-safer-community-with-new-suicide-prevention-tools/>

4. vidéo en direct

5. <http://www.durkheimproject.org/research/>

6. <https://liwc.wpengine.com>



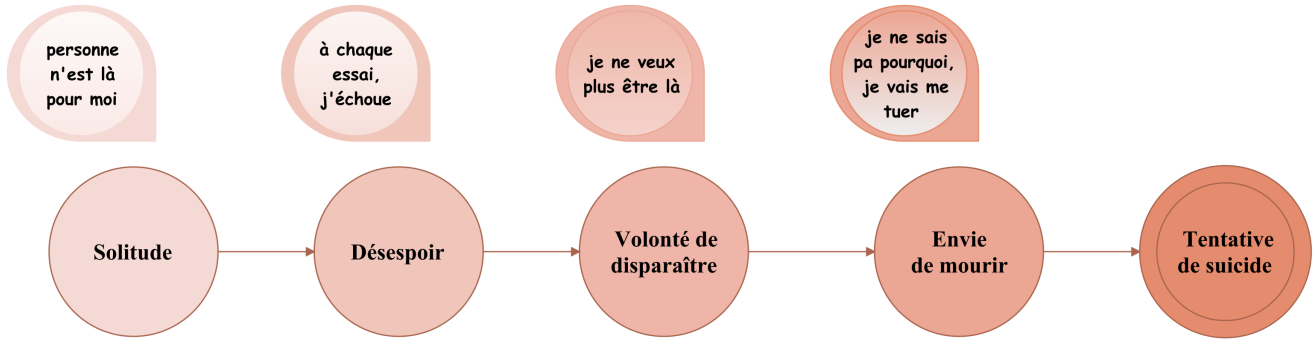


FIGURE 1 – Exemple d'évolution de l'état émotionnel d'un individu.

chaque message indépendamment de la séquence entière. Dans ce travail, nous supposons que l'état émotionnel déduit d'un message au temps  $t$ , dépend des états émotionnels précédents. Notre hypothèse principale est que les états émotionnels peuvent être modélisés comme des observations dépendantes et continues, qui peuvent être capturées via des méthodes de traitement automatique de la langue puis prédites. L'état émotionnel peut être représenté soit par un état positif, neutre ou négatif [2], soit par des modèles plus complexes incluant des émotions comme la tristesse, l'espoir, l'excitation, etc. [12, 22, 10]. Dans la suite, nous considérons trois niveaux d'états émotionnels mais notre approche peut être facilement généralisée indépendamment du nombre d'états émotionnels.

### 3.1 Description du problème

Soit  $P = \langle p_1, p_2, \dots, p_n \rangle$  un flux continu de messages (tweets, messages facebook, etc.) ordonnés dans le temps dans une fenêtre temporelle  $W$ . Le problème consiste à prédire un vecteur  $Y = \langle y_1, y_2, \dots, y_n \rangle$  d'états émotionnels associés à la séquence de messages observée  $P$ . Les observations  $P$  en entrée sont représentées par des vecteurs d'attributs. Chaque observation  $p_j$  contient différentes informations à propos du message au temps  $t_j$ . Chaque variable  $y_j$  est un état émotionnel inféré à partir de l'observation  $p_j$ . La Figure 1 décrit une série de messages impliquant un changement de l'état émotionnel.

Ces changements d'états sont inspirés par un travail récent en psychologie cognitive [1]. Les auteurs ont étudié différents types de comportements suicidaires en explorant des signes cognitifs et les changements de comportements pouvant indiquer une tentative suicide. Par exemple, comme le montre la figure 1, un individu évoque sa *volonté de disparaître* avec le message "*Je ne veux plus être là*". Il est susceptible de voir son état émotionnel évoluer vers un niveau plus risqué comme *Envie de mourir* qu'il exprime avec le message "*Je vais me tuer*", avant de réaliser une *tentative de suicide*. Chaque nœud de la figure représente un état émotionnel qui est calculé en fonction d'une observation, le message de l'utilisateur. Les transitions (arêtes) entre les états encodent la séquence des changements d'état. Les modèles de réseaux de Markov permettent de représenter

ces changements émotionnels séquentiels. Une généralisation intéressante est donnée par des modèles graphiques tels que les Conditional Random Fields (CRF) [11].

Dans la suite, nous présentons les deux étapes de notre méthode. La première étape permet d'associer un ou plusieurs états émotionnels à un message à partir de l'analyse du texte de ce message. La deuxième étape permet de modéliser les interactions entre états émotionnels et peut être utilisée pour prédire le prochain état émotionnel d'un individu.

### 3.2 Extraction des caractéristiques

Avant d'extraire des caractéristiques des messages, ces derniers sont prétraités (mise en minuscules, suppression des ponctuations multiples, des caractères spéciaux, des mentions à d'autres utilisateurs et des URL). Nous avons extrait des caractéristiques largement utilisées dans la littérature [19].

- Le premier ensemble de caractéristiques inclut les caractéristiques lexicales du texte. Nous utilisons les étiquettes grammaticales (POS) pour capturer l'auto-référence (pronoms personnels à la première personne "I", "My", "Je", "Mon", etc), les noms, les verbes et les adverbess. [6] ont montré que l'utilisation de la première personne au singulier ou au pluriel peut révéler le bien-être ou le mal-être mental. Nous prenons également en compte l'intensité des émotions, en considérant la présence d'*intensifieurs* comme *très*, *complètement*, *intensément* surtout lorsqu'ils sont utilisés avec des pronoms personnels (eg., "je suis très triste"). Quand une phrase contient une négation suivie d'un symptôme, nous inversons cette caractéristique (eg., "je ne vais pas bien !").
- Le second ensemble de caractéristiques est lié aux lexiques. Nous cherchons dans les messages des termes couramment utilisés par les personnes à risque dans les médias sociaux. Nous considérons la fréquence des termes faisant référence aux émotions négatives, à la dépression, à l'automutilation, à la tristesse, à la santé mentale et au suicide. Pour ce faire, nous nous sommes inspirés des travaux de [6], qui ont exploité le lexique ANEW [3] contenant un classement d'émotions pour un large

nombre de mots<sup>7</sup>. Ensuite, nous enrichissons ces caractéristiques en incluant un autre lexique qui se réfère aux mots d'injure. En effet, [6] ont montré que ces caractéristiques véhiculent des informations importantes dans le contexte de l'analyse des états émotionnels.

### 3.3 Modèle basé sur les Conditional Random Fields

CRF est un type de modèle graphique probabiliste non dirigé qui a été appliqué avec succès dans de nombreux problèmes de traitement de textes et de visualisation [21]. Un avantage de ce modèle réside dans sa capacité à capturer les dépendances complexes entre les observations, en plus des interprétations complètes de la relation entre les caractéristiques qu'il fournit. Dans notre contexte, cette propriété est très importante étant donné que la transition d'un état émotionnel à un autre dépend fortement des états observés précédemment. Par exemple, comme représenté dans la figure 1, il est très improbable que l'état émotionnel d'un individu saute soudainement de *solitude* (non risqué) à un état *tentative de suicide* (très risqué). La modélisation CRF est un modèle puissant qui aide à apprendre les comportements des utilisateurs et à prédire la séquence des états mentaux des utilisateurs.

Soit une séquence de messages observée  $P = \langle p_1, p_2, \dots, p_n \rangle$  et une séquence d'états émotionnels cachés  $Y = \langle y_1, y_1, \dots, y_n \rangle$ , CRF modélise la probabilité conditionnelle comme suit :

$$p(Y|P) = \frac{1}{Z(P)} \exp\left(\sum_{i=1}^n \sum_{k=1}^F w_k f_k(y_{i-1}, y_i, P, i)\right) \quad (1)$$

où  $Z$  est un facteur de normalisation (aussi appelé la fonction de partition) pour que  $p(Y|P)$  soit une probabilité valide pour toutes les séquences étiquetées.  $Z$  est défini comme la somme de l'exponentielle du nombre de séquences :

$$Z(P) = \sum_P \exp\left(\sum_{i=1}^n \sum_{k=1}^F w_k f_k(y_{i-1}, y_i, P, i)\right) \quad (2)$$

Les paramètres  $w_k$  sont les poids des caractéristiques  $f_i$ . Ils sont appris par des techniques d'optimisation comme les approches par gradient. Les fonctions caractéristiques  $f_k(y_{i-1}, y_i, P, n)$  prennent en compte une paire d'états émotionnels adjacents  $y_{i-1}, y_i$ , la séquence entière de messages  $P$  et la position courante dans la séquence  $i$ .

Notons que l'utilisation de CRF nous permet de définir un grand nombre de fonctions dépendantes ou indépendantes sans nous soucier de la relation statistique complexe entre ces fonctions. L'utilisation de chaque fonction dépend du poids  $w_k$  qui agit comme facteur d'activation de la fonction.

7. Ce lien contient le code et les scripts utilisés pour générer ce lexique : [https://github.com/sbma44/begin\\_aneu](https://github.com/sbma44/begin_aneu).

## 4 Expérimentations

Dans ce qui suit, nous détaillons la préparation du jeu de données utilisé, puis nous analysons les performances de l'approche en utilisant les caractéristiques détaillées dans 3.2. Nous explorons aussi l'importance des caractéristiques extraites à partir des messages et nous montrons les thèmes importants pour chaque état émotionnel en utilisant le modèle Latent Dirichlet Allocation [8].

### 4.1 Préparation des données

En raison de l'absence de base de données librement accessible pour l'évaluation des méthodes de détection des risques de suicide dans les médias sociaux, nous avons utilisé l'API en temps réel Twitter<sup>8</sup> pour collecter des tweets contenant des références à des thèmes tels que la dépression, l'automutilation, l'anorexie et le suicide. La liste de mots clés utilisée pour récupérer les tweets a été définie manuellement à partir de la liste des facteurs de risque définie par l'APA (American Psychological Association<sup>9</sup>) et la liste des signes avant-coureurs définie par l'AAS (American Association of Suicidology<sup>10</sup>).

Parmi les tweets recueillis, nous n'avons conservé que ceux pour lesquels un symptôme lié au suicide a été validé par un psychiatre. 60 individus ont ainsi été choisis. Pour éviter un ajustement excessif du modèle, nous avons également inclus 60 comptes Twitter d'individus non à risque en utilisant les mêmes mots clés. Nous avons sélectionné au hasard les 50 derniers tweets de chacun de ces groupes. La collection totale de données contient 5976 tweets. Huit chercheurs et un psychiatre ont manuellement annoté 507 tweets de la collection pour déterminer les états émotionnels associés. Pour cette étude préliminaire, nous avons considéré trois états émotionnels. Le choix de ces classes est motivé par les travaux de [13]. Ces classes sont définies comme suit :

- *Aucune détresse* : le message traite d'événements quotidiens tels que le travail, les sorties, les activités du week-end, etc.
- *Détresse minimale/modéré* : le message exprime un niveau de détresse qui pourrait être considéré comme commun pour la plupart des individus tels que un examen, une présentation pour le travail, une dispute avec un ami, etc.
- *Détresse importante* : le message mentionne des références à l'auto-mutilation, aux idéations suicidaires, des excuses, des sentiments négatifs comme l'inutilité, la haine de soi, la culpabilité, etc.

Chaque tweet a été examiné par au moins deux annotateurs, avec un sous-ensemble de 55 tweets validés par le psychiatre. Nous avons calculé un kappa de Cohen de 69,1%, qui souligne un accord substantiel entre les annotateurs.

8. <https://dev.twitter.com/streaming/overview>

9. <http://www.apa.org/topics/suicide/>

10. <http://www.suicidology.org>

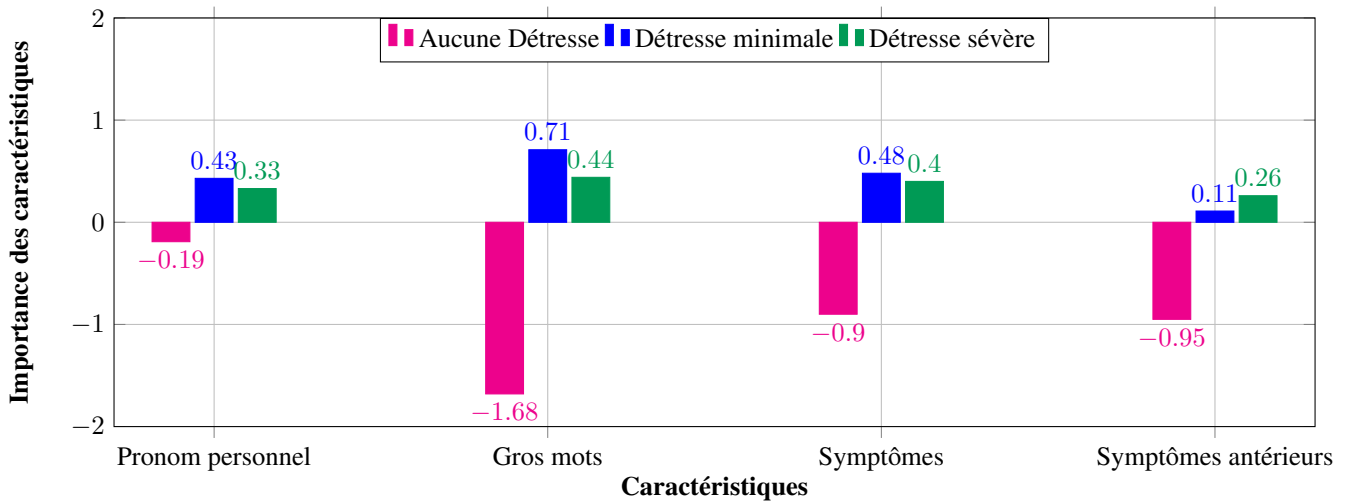


FIGURE 2 – Importance des caractéristiques dans les trois classes.

Nous avons également calculé un kappa pondérée, qui tient compte des différents niveaux de désaccord, de l'ordre de 71,5% qui est un taux largement satisfaisant pour juger l'accord entre les annotateurs. Le processus d'annotation a donné 141 instances de la classe *aucune détresse*, 110 instances de la classe *détresse minimale* et 256 instances de la classe *détresse sévère*.

## 4.2 Protocole d'évaluation

Afin d'évaluer notre approche, nous avons adopté une méthodologie entièrement automatisée basée sur une validation croisée ( $k=5$ ) sur l'ensemble des données annotés afin d'apprendre et tester le modèle proposé. Pour ce faire, à chaque itération, nous avons partitionné l'ensemble des 507 tweets annotés en échantillons d'apprentissage (70%) et de test (30%). Chaque instance est constituée par le tweet d'un utilisateur avec comme contexte l'ensemble des tweets formant la séquence des publications de l'utilisateur. L'objectif principal de la phase d'apprentissage est d'apprendre les paramètres de notre modèle ainsi que ceux des méthodes de référence (*baselines*). Nous comparons notre approche avec les deux modèles SVM et Random Forest en utilisant les mêmes ensembles d'apprentissage et de test. Nous avons utilisé les mesures d'évaluation : Rappel, Précision et F-mesure.

## 4.3 Résultat et discussion

**Analyse des caractéristiques.** Nous avons exploré l'importance des caractéristiques extraites à partir des messages en fonction des trois états émotionnels considérés dans la figure 2. Cette analyse a été effectuée sur l'ensemble des données d'apprentissage. Il est à noter que cette analyse est exploratoire, et nous ne l'avons pas utilisée pour la sélection des attributs les plus pertinents.

Chaque poids donne l'importance de la caractéristique pour l'état émotionnel. Alors que les poids élevés (positifs ou négatifs) indiquent une association forte, des pondéra-

tions nulles indiquent que la caractéristique a peu ou pas d'impact sur l'état émotionnel.

Dans ce travail, nous avons considéré 39 caractéristiques pour chaque message, mais pour des raisons de présentation, nous ne rapportons ici que les caractéristiques les plus importantes. Les *pronoms à la première personne* (lexical), les *mots d'injures* (lexique), les *Symptômes* (lexique) et *Symptômes antérieurs* (lexique) ont un impact négatif sur l'état émotionnel "Aucune détresse". En effet, les individus les moins à risque n'utilisent généralement pas de références aux symptômes ni de mots d'injures dans leurs messages. Au contraire, les injures et les symptômes sont des marqueurs importants pour la classe *détresse minimale* avec une légère différence pour la classe *détresse sévère*. D'autre part, la caractéristique *Symptômes antérieurs* qui reflète les symptômes antérieurs observés, est importante pour la classe *détresse sévère*. En effet, si à l'instant  $t-1$ , si un individu publie un message ayant une valeur importante pour la caractéristique *Symptôme*, il est probable que l'utilisateur présente une détresse émotionnelle sévère au moment  $t$ . La possibilité de prendre en compte ce dernier point est un avantage clé des modèles CRF.

Les tableaux 1 et 2 montrent les thèmes extraits des tweets appartenant aux états émotionnels *Aucune détresse* et *Détresse sévère*, en utilisant le modèle Latent Dirichlet Allocation [8]. Par souci de simplicité, nous avons fixé le nombre de thèmes à 2. Dans le tableau 1, les messages de l'état émotionnel *Aucune détresse* portent clairement sur la *famille*, les *voyages*, les *relations*, etc. alors que dans la Table 2, les thèmes abordés sont liés aux pensées suicidaires (e.g. suicide, meurtre, etc.). On remarque l'importance des intensificateurs appliqués aux termes reflétant des idées liées à la fin de vie (e.g. *assez*, *sans valeur*, *plus*, *fin*). Les figures 3 et 4 présentent les termes les plus utilisés pour les deux états émotionnels. Ces dernières corroborent les conclusions du modèle LDA.

Thème	Description
1	crazy tonight today morning hot <b>game</b> road <b>excitingtimes</b> haftflmdic <b>familytravel</b> arkan- sas cards poolside be life want up if day good everyone my yourself blessed hidline <b>girlideas</b> really <b>lol</b> know finally
2	<b>love</b> my me <b>happy</b> im bed <b>relationships</b> can always new better we few little snapchat going wait <b>excited</b> school suck drift tired makes exactly pat older attractive <b>dearly</b> travel catch

TABLE 1 – Thèmes extraits des tweets de la classe *Aucune détresse*.

Thème	Description
1	me if would up one really <b>fucking</b> will <b>no</b> eve- ryone am give <b>enough</b> days <b>everything</b> good <b>suicidal</b> still my seems <b>suicide</b> like <b>never</b> be <b>fuck</b> alive <b>die</b> say <b>help</b> <b>sleep</b>
2	<b>want</b> <b>myself</b> <b>tired</b> can my don <b>anymore</b> be wish not <b>kill</b> <b>worthless</b> never <b>feel</b> <b>die</b> will am <b>hate</b> end time <b>sorry</b> much me like <b>enough</b> live everything many <b>fight</b> <b>depressionproblems</b>

TABLE 2 – Thèmes extraits des tweets de la classe *Détresse sévère*.

**Analyse des changements d'états émotionnels.** Pour mieux comprendre les changements de comportement des utilisateurs, nous exploitons la puissance des CRF pour analyser les changements entre les états émotionnels. La figure 5 montre les transitions entre les 3 états émotionnels que nous avons considérés, en se basant sur l'ensemble des données d'apprentissage. Cette figure 5 permet d'identifier que les transitions les plus probables entre deux états différents vont de la classe *Aucune détresse* à la classe *Détresse minimale*, avec une probabilité inférieure pour la transition opposée. Les individus passant à un état émotionnel plus risqué sont peu susceptibles de revenir à un état normal. Les utilisateurs dans l'état *Aucune détresse* et *Détresse sévère* tendent à rester dans le même état avec des valeurs de 0,54 et 1,41, respectivement. La valeur de probabilité très faible pour la transition de *Détresse sévère* à *Détresse minimale* s'explique car l'état émotionnel *Détresse sévère* est généralement atteint lorsque l'individu se focalise sur une possible tentative de suicide [1].

La Table 3 fournit les résultats des mesures de précision, de rappel et le F1-score pour toutes les classes, ainsi qu'une comparaison avec les deux modèles d'apprentissage auto-

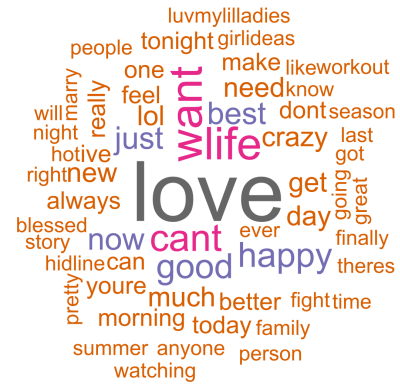


FIGURE 3 – *Aucune Détresse*.

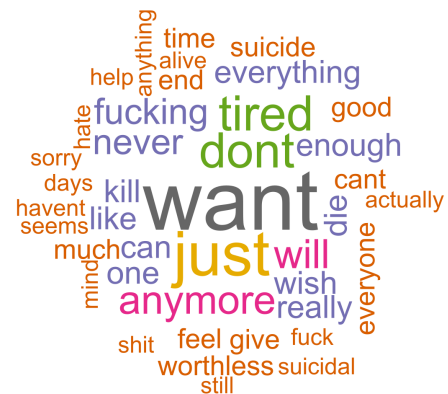


FIGURE 4 – *Détresse sévère*.

matique SVM et Random Forest. Dans la phase d'apprentissage, ces modèles sont construits en utilisant les paramètres par défaut de WEKA<sup>11</sup>. Pour apprendre notre modèle CRF, nous exploitons une descente de gradient en utilisant la méthode L-BFGS. Les coefficients de régularisation L1 et L2 du modèle sont fixés par défaut à 0.1 et 0.1, respectivement.

Les meilleurs résultats sont ceux des classes *Aucune détresse* et *Détresse sévère* selon le score F1, en dépit de la valeur importante de précision pour l'état émotionnel *Détresse minimale*. En effet, le taux élevé de faux négatifs pour l'état de détresse minimale peut s'expliquer par le fait que : (i) cet état se trouve entre deux états ; (ii) l'auto-transition pour cet état est très faible (0.01) par rapport aux transitions entrantes et sortantes, en particulier depuis l'état *Aucune détresse*. D'autre part, le modèle permet d'identifier tous les messages appartenant à l'état *Aucune détresse*, ce qui n'est pas surprenant compte tenu de l'analyse de l'importance des caractéristique présentée dans la figure 2 (cf. section 4.3). Les résultats présentés dans le tableau 3 sont prometteurs car la plupart des méthodes de classification des textes pour des applications liées au suicide ou à la dépression atteignent à peine 0,7 [4, 16]. Les résul-

11. [www.cs.waikato.ac.nz/ml/weka/](http://www.cs.waikato.ac.nz/ml/weka/)

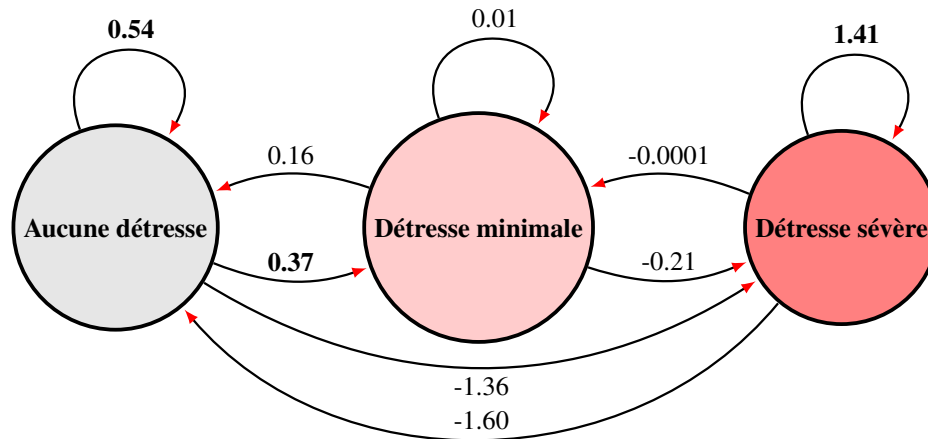


FIGURE 5 – Changement des émotions des utilisateurs selon les trois classes. Les valeurs en gras sont les plus probables. La couleur symbolise la sévérité de l'état émotionnel.

tats obtenus par les méthodes d'apprentissage automatique sont moins importants avec une grande différence en comparaison avec Random Forest. Les valeurs sont plus faibles en termes de Rappel pour les deux méthodes. Cette différence de performance peut être expliquée par l'absence de sélection d'attributs qui pourrait être considéré comme un désavantage dans les tâches de classification de texte.

	Précision	Rappel	F1-score
Aucune détresse	0.706	<b>1.000</b>	0.828
Détresse minimale	<b>1.000</b>	0.176	0.300
Détresse sévère	<b>0.941</b>	0.571	0.711
<b>Notre approche</b>	<b>0.816</b>	<b>0.752</b>	<b>0.711</b>
SVM	0.446	0.227	0.301
Random Forest	0.500	0.127	0.202

TABLE 3 – Évaluation des résultats du système de monitoring.

## 5 Conclusion et Perspectives

Dans cet article, nous avons proposé une approche pour le dépistage des idéations suicidaires basée sur un modèle probabiliste appelé Conditional Random Fields qui permet de modéliser et prédire le comportement en ligne de l'individu comme une séquence d'états émotionnels évoluant au fil du temps. Cette représentation permet d'incorporer un ensemble riche de caractéristiques complexes intégrant le contexte des messages précédents. L'efficacité de l'approche a été évaluée sur des données réelles, sur une collection de tweets publiés par des individus ayant montré des symptômes graves liés au suicide. Ces évaluations préliminaires ont montré que notre modèle est capable de fournir des interprétations complètes de la relation entre les états émotionnels et les résultats en termes de prédictions sont encourageants par rapport à la littérature.

Un avantage de l'approche est que nous pouvons facilement incorporer de nouvelles caractéristiques liées au texte en incluant notamment de nouveaux lexiques mais également non liées aux textes comme des informations contextuelles : l'heure de la rédaction du message, sa longueur, etc mais encore des caractéristiques liées à d'autres médias (images, vidéos...) associés aux messages. En effet, le modèle CRF permet d'incorporer un ensemble riche de caractéristiques représentant le contexte sans se soucier de leurs relations a priori (corrélations positives ou négatives). Cette flexibilité nous permet d'intégrer dans le modèle un ensemble de caractéristiques dont les dépendances peuvent être assez complexes et mal connues.

Un point important consiste à filtrer les références réelles liées au suicide par rapport aux messages de support et de condoléances, ou encore les campagnes de prévention du suicide [4]. Par ailleurs, nous pouvons intégrer une représentation plus complexe de l'état émotionnel que seulement trois états. Par exemple, dans une étude clinique qui a été menée sur des jeunes étudiants, [15] ont établi un ensemble de critères cliniques de suicide qui peuvent être présents dans les publications Facebook. Nous citons à titre d'exemple la dépression, perte d'intérêt/plaisir dans les activités, changements d'appétit, problèmes de sommeil, agitation psychomotrice ou retard, perte d'énergie, sentiment d'inutilité ou de culpabilité, diminution de la concentration, et idées suicidaires.

## Références

- [1] A. Adler, A. Bush, F. K. Barg, G. Weissinger, A. T. Beck, and G. K. Brown. A mixed methods approach to identify cognitive warning signs for suicide attempts. *Archives of Suicide Research*, 20(4) :528–538, 2016.
- [2] L. Barbosa and J. Feng. Robust sentiment detection on Twitter from biased and noisy data. In *Proceedings*

- of the 23rd International Conference on Computational Linguistics : Posters, COLING 2010, pages 36–44, Beijing, China, 2010. Association for Computational Linguistics.
- [3] M. M. Bradley and P. J. Lang. Affective norms for English words (ANEW) : Stimuli, instruction manual, and affective ratings. Technical report, Center for Research in Psychophysiology, University of Florida, Gainesville, Florida, 1999.
- [4] P. Burnap, W. Colombo, and J. Scourfield. Machine classification and analysis of suicide-related communication on twitter. In *Proceedings of the 26th ACM Conference on Hypertext & Social Media*, HT '15, pages 75–84, New York, NY, USA, 2015. ACM.
- [5] G. B. Colombo, P. Burnap, A. Hodorog, and J. Scourfield. Analysing the connectivity and communication of suicidal users on twitter. *Computer Communications*, 73, Part B :291 – 300, 2016. Online Social Networks.
- [6] M. De Choudhury, S. Counts, and E. Horvitz. Predicting postpartum changes in emotion and behavior via social media. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*, CHI '13, pages 3267–3276, New York, NY, USA, 2013. ACM.
- [7] J. F. Gunn and D. Lester. Twitter postings and suicide : An analysis of the postings of a fatal suicide in the 24 hours prior to death. *Suicidologi*, 17(3) :28–30, 2012.
- [8] M. D. Hoffman, D. M. Blei, and F. Bach. Online learning for latent dirichlet allocation. In *Proceedings of the 23rd International Conference on Neural Information Processing Systems*, NIPS'10, pages 856–864, USA, 2010. Curran Associates Inc.
- [9] C. Karmen, R. C. Hsiung, and T. Wetter. Screening internet forum participants for depression symptoms by assembling and enhancing multiple NLP methods. *Computer Methods and Programs in Biomedicine*, 120(1) :27–36, 2015.
- [10] S. Kim, F. Li, G. Lebanon, and I. A. Essa. Beyond sentiment : The manifold of human emotions. In *Proceedings of the Sixteenth International Conference on Artificial Intelligence and Statistics*, pages 360–369, 2013.
- [11] J. D. Lafferty, A. McCallum, and F. C. N. Pereira. Conditional random fields : Probabilistic models for segmenting and labeling sequence data. In *Proceedings of the Eighteenth International Conference on Machine Learning*, ICML '01, pages 282–289, San Francisco, CA, USA, 2001. Morgan Kaufmann Publishers Inc.
- [12] R. J. Larsen and E. Diener. Promises and problems with the circumplex model of emotion. *Review of Personality and Social Psychology*, 13(13) :25–59, 1992.
- [13] M. T. Lehrman, C. O. Alm, and R. A. Proaño. Detecting distressed and non-distressed affect states in short forum texts. In *Proceedings of the 2012 Workshop on Language in Social Media*, LSM 2012, pages 9–18, Montreal, Canada, 2012. Association for Computational Linguistics.
- [14] C. Maigrot, S. Bringay, and J. Azé. Concept drift vs suicide : comment l'un peut prévenir l'autre ? In *16ème Journées Francophones Extraction et Gestion des Connaissances, EGC 2016*, volume E-30, pages 219–230, 2016.
- [15] M. Moreno, L. Jelenchick, K. Egan, E. Cox, H. Young, K. Gannon, and T. Becker. Feeling bad on Facebook : depression disclosures by college students on a social networking site. *Depression and Anxiety*, 28(6) :447–455, 2011.
- [16] B. O'Dea, S. Wan, P. J. Batterham, A. L. Calear, C. Paris, and H. Christensen. Detecting suicidality on twitter. *Internet Interventions*, 2(2) :183 – 188, 2015.
- [17] J. P. Pestian, P. Matykiewicz, and J. Grupp-Phelan. Using natural language processing to classify suicide notes. In *Proceedings of the Workshop on Current Trends in Biomedical Natural Language Processing*, BioNLP '08, pages 96–97, Stroudsburg, PA, USA, 2008. ACL.
- [18] C. Poulin, B. Shiner, P. Thompson, L. Vepstas, Y. Young-Xu, B. Goertzel, B. Watts, L. Flashman, and T. McAllister. Predicting the risk of suicide by analyzing the text of clinical notes. *PLoS ONE*, 9(1), 2014.
- [19] I. Spasic, P. Burnap, M. Greenwood, and M. Arribas-Ayllon. A Naïve Bayes Approach to Classifying Topics in Suicide Notes. *Biomedical Informatics Insights*, 5(1) :87–97, 2012.
- [20] H. Sueki. The association of suicide-related twitter use with suicidal behaviour : A cross-sectional study of young internet users in japan. *Journal of Affective Disorders*, 170 :155 – 160, 2015.
- [21] C. Sutton and A. McCallum. An introduction to conditional random fields. *Found. Trends Mach. Learn.*, 4(4) :267–373, 2012.
- [22] M. Yik, J. A. Russell, and J. H. Steiger. A 12-point circumplex structure of core affect. *Emotion*, 11(4) :705–731, 2011.

# Négociation « one-to-many » adaptative pour améliorer l'acceptabilité des services d'un fournisseur SaaS

Amro Najjar<sup>1</sup>Olivier Boissier<sup>1</sup>Gauthier Picard<sup>1</sup><sup>1</sup> Laboratoire Hubert Curien UMR CNRS 5516, Saint-Etienne, France

prenom.nom@emse.fr

## Résumé

Le taux d'acceptabilité et la satisfaction utilisateur sont devenus des facteurs clés pour éviter le désabonnement des clients et assurer le succès de tout fournisseur de logiciel en tant que service (ou SaaS). Néanmoins, le fournisseur doit également minimiser les coûts de location de services cloud. Pour faire face à ces objectifs contradictoires, la plupart des travaux considèrent la gestion de ressources de manière unilatérale par le fournisseur. Ainsi, les préférences utilisateur et leur acceptabilité subjective sont ignorées. Des études récentes dans le domaine de la qualité d'expérience (QoE) recommandent aux fournisseurs d'utiliser des métriques comme les quantiles pour jauger plus précisément l'acceptabilité des services. Dans cet article, nous proposons un mécanisme de négociation « one-to-many » adaptatif pour améliorer l'acceptabilité des services d'un fournisseur SaaS. Se basant sur une estimation par quantiles de l'acceptabilité des services et sur l'apprentissage du modèle de négociation de l'utilisateur, ce mécanisme ajuste le processus de négociation du fournisseur afin de garantir un taux d'acceptabilité désiré tout en respectant des contraintes budgétaires. Ce mécanisme est mis en œuvre et ses résultats sont analysés au regard d'approches comparables.

## Mots Clef

négociation, adaptation, taux d'acceptabilité, SaaS, cloud computing.

## Abstract

Service acceptability rate and user satisfaction are becoming key factors to avoid client churn and secure the success of any Software as a Service (SaaS) provider. Nevertheless, the provider must also minimize the cost it pays to rent resources from the cloud. To address these contradicting concerns, most of existing works carry out resource management unilaterally by the provider. Consequently, end-user preferences and her subjective acceptability of the service are mostly ignored. In order to assess user satisfaction and service acceptability, Mean Opinion Score (MOS) used to be the most popular metric. However, recent studies

in the domain of Quality of Experience (QoE) recommend providers to use better metrics such as quantiles to gauge user service acceptability more precisely. In this article we propose an adaptive one-to-many negotiation mechanism to improve the service acceptability of an open SaaS provider. Based on quantile estimation of service acceptability rate and a learned model of the user negotiation strategy, this mechanism adjusts the provider negotiation process in order to guarantee the desired service acceptability rate while meeting the budget limits of the provider. The proposed mechanism is implemented and its results are examined and analyzed in light of comparable results.

## Keywords

negotiation, adaptation, acceptance rate, SaaS, cloud computing

## 1 Introduction

Le taux d'attrition (la proportion de clients perdus sur une période donnée) est un des indicateurs les plus négatifs affectant les fournisseurs de « logiciel en tant que service » (SaaS). Une étude récente de la dynamique des clients a montré que la majorité des consommateurs aux Etats-Unis ont changé de fournisseurs de service car ils ne répondaient pas à leurs attentes [1]. Par ailleurs, une augmentation significative de la popularité des applications multimédia est prédite [8]. Ainsi, le marché de demain se formera autour d'une demande intensive et fluctuante, avec de fortes attentes client. Afin de palier cette rapide évolution, les fournisseurs de services applicatifs (ASP) migrent de plus en plus vers le *cloud* pour gérer leurs ressources de manière flexible et ainsi minimiser leurs coûts opérationnels. Ainsi, les ASP et les fournisseurs SaaS doivent minimiser le taux d'attrition tout en respectant leurs contraintes budgétaires. Dans le cadre du *cloud computing*, ce problème, objet de nombreuses études, est connu comme la gestion de l'élasticité ou l'*auto-scaling* [4, 15]. Cependant, la plupart de ces travaux adoptent une approche centralisée où l'ASP prend unilatéralement les décisions d'allocation de ressource. Par conséquent, les préférences utilisateurs sont négligées, et il est souvent présumé que leur seuil d'acceptabilité tolère le

meilleur service offert.

Les notions de taux d'acceptabilité et de taux d'attrition sont fortement liées à la perception subjective de l'utilisateur sur la qualité de service ou la qualité d'expérience (QoE) [12]. Bien que la plupart des études sur la QoE soient menées sur le plan conceptuel, la QoE est également un bon moyen de fournir une mesure pratique permettant de quantifier la satisfaction des utilisateurs et l'acceptation du service [31]. En particulier, la gestion de la QoE est apparue comme un processus visant à maximiser la QoE tout en optimisant les ressources utilisées [25]. Cependant, la plupart des travaux existants souffrent des deux limitations suivantes. Tout d'abord, la littérature sur la gestion de la QoE repose largement sur le *score d'opinion moyenne* (ou MOS) pour évaluer la satisfaction et l'acceptabilité du service [12]. Ce score a été utilisé à la fois pour les applications NGN (Next Generation Networks) et le *cloud computing*. Toutefois, comme il s'agit d'une moyenne d'opinions des utilisateurs, MOS masque des informations importantes sur la diversité des utilisateurs et leurs préférences personnelles, alors que cette différence d'opinions peut avoir un impact significatif pour le fournisseur [6]. Pour cette raison, d'autres mesures ont été proposées pour permettre au prestataire de comprendre la satisfaction de l'utilisateur final et d'estimer le taux d'attrition des clients. Par exemple, des quantiles et des percentiles ont été proposés comme des outils de mesure permettant au fournisseur de s'assurer que, par exemple, 95 % de ses utilisateurs trouvent le service acceptable ou mieux [7]. La seconde limitation est que, dans la majorité de ces travaux, la gestion de la QoE reste un processus unilatéral réalisé par le fournisseur sans intégrer les préférences de l'utilisateur final dans la boucle [25]. Par conséquent, malgré l'objectif déclaré d'établir une estimation subjective et centrée sur l'utilisateur de la qualité du service, les travaux existants sont principalement axés sur les fournisseurs.

Par définition, l'approche *agent* est liée à une perspective individuelle [32]. Les agents sont ainsi pertinents pour représenter la subjectivité des opinions des utilisateurs et l'acceptation d'un service donné [17]. En outre, les principes du comportement de négociation discutés dans [19] constituent un outil utile pour représenter les attentes des utilisateurs finaux. Ces derniers sont des déterminants clés de la satisfaction des utilisateurs et de l'acceptabilité des services [23, 33, 14]. De plus, les systèmes multi-agents fournissent une plate-forme distribuée capable d'entreprendre des tâches nécessitant une coopération telle que la gestion des ressources [28]. La négociation multi-agent « one-to-many » est une solution intéressante pour intégrer l'utilisateur final et son acceptabilité subjective dans la gestion de l'élasticité [9]. Néanmoins, la majorité des travaux existants dans la littérature de négociation « one-to-many » traite d'un scénario où le vendeur cherche à trouver un accord *atomique* conclu avec l'un des multiples acheteurs simultanés. En outre, la plupart de ces travaux supposent un ensemble *fermé* de participants dans lequel les adver-

saies de l'agent unique sont connus à l'avance avant le début du processus de négociation. De plus, dans la plupart des travaux existants, les offres sont envoyées et reçues de manière synchrone. Ces hypothèses ne tiennent pas dans l'écosystème *cloud* ouvert d'aujourd'hui, où des centaines d'utilisateurs peuvent entrer ou quitter le système chaque minute et où les sessions de négociation ne sont pas synchronisées.

Dans cet article, nous développons de nouveaux mécanismes de négociation et de coordination multilatéraux adaptatifs et ouverts permettant au fournisseur d'atteindre un taux d'acceptabilité de service ciblé tout en satisfaisant ses contraintes budgétaires. Les utilisateurs peuvent décider d'accepter ou de rejeter le service proposé en fonction de leurs attentes et d'une estimation subjective de la qualité du service [14]. En fonction de ses mesures de la portion d'utilisateurs qui considèrent que le service est inacceptable, le fournisseur ajuste sa stratégie de négociation afin de ramener le taux d'acceptabilité à ses objectifs prédéfinis. Ainsi, à l'aide du mécanisme proposé, le fournisseur SaaS peut (i) s'adapter à la nature dynamique et ouverte de l'écosystème *cloud*, où les utilisateurs peuvent entrer dans le système, entrer le système ou le quitter à volonté, et (ii) intégrer les préférences de l'utilisateur dans le processus de décision pour parvenir à des accords mutuellement acceptés en assurant un taux d'acceptabilité prédéfini précis et en atteignant ses objectifs commerciaux. Les mécanismes de négociation et de coordination proposés sont développés dans l'architecture EMan [13].

Le reste de l'article est structuré comme suit. La section 2 introduit brièvement l'architecture EMan dans laquelle vient s'intégrer notre proposition. La section 3 détaille la stratégie de négociation adaptative et l'approche d'apprentissage et de modélisation de l'adversaire proposée. La section 4 détaille le cadre expérimental et discute des résultats obtenus. La section 5 discute des travaux connexes, avant de conclure avec quelques perspectives en section 6.

## 2 L'architecture EMan

Le mécanisme proposé dans cet article est implémenté dans l'architecture EMan [16, 14, 13]. C'est une architecture multi-agent pour la gestion de l'élasticité de plate-formes SaaS. EMan suit le même schéma architectural partagé par la plupart des solutions « one-to-many » existantes dans la littérature. Dans la version initiale de EMan [14, 13], les utilisateurs sont assistés par des agents autonomes dont le but est de maximiser la QoE de leurs utilisateurs respectifs. Nous montrerons dans la section suivante comment nous y incluons le mécanisme d'adaptation que nous proposons.

### 2.1 Agents

L'architecture EMan représentée dans la figure 1 modélise la négociation qui a lieu entre un fournisseur SaaS et ses utilisateurs finaux ou utilisateurs de services (SU). La négociation entre les fournisseurs SaaS et les fournisseurs de



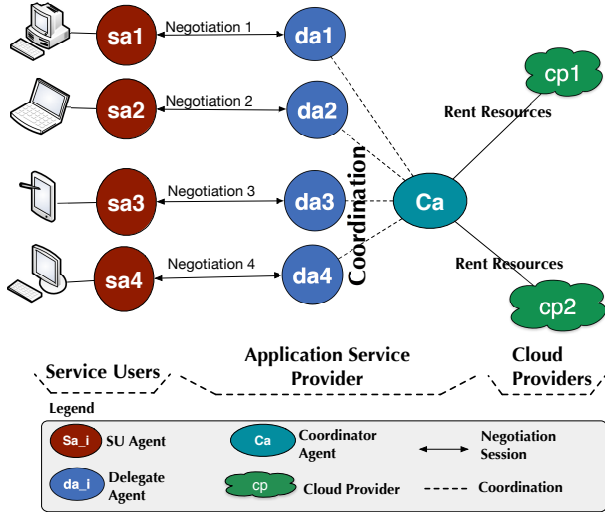


FIGURE 1 – L’architecture EMan déployée dans un écosystème cloud

cloud computing n’est pas décrite dans cet article. L’architecture EMan contient trois types d’agents : les agents utilisateur de service (notés  $sa_i$ ), les agents délégués (notés  $da_i$ ) et un seul coordinateur (noté  $ca$ ). Ces deux derniers types d’agents représentent le fournisseur.

**Agents utilisateurs (SA ou  $sa_i$ ).** Un agent utilisateur de service participe au processus de négociation pour le compte d’un utilisateur de service. Un  $sa_i$  a une fonction d’utilité  $M_{sa_i}$  qui encode ses préférences.  $M_{sa_i}$  est utilisée à chaque cycle  $t$  pour évaluer l’utilité des offres  $o_{da_i}^t$  provenant du délégué correspondant (voir section 2.2). Si le service comporte  $J$  attributs, alors  $M_{sa_i}$  est défini comme suit :

$$M_{sa_i}(o_{da_i}^t) = \sum_{j=1}^{j=J} w_{sa_i,j} \cdot \mu_{sa_i,at_j}(o_{da_i}^t[at_j]) \quad (1)$$

où  $w_{sa_i,at_j}$  est le poids associé à l’attribut  $at_j$  pour spécifier l’importance que cet utilisateur donne à cet attribut (ces pondérations doivent satisfaire  $\sum_{j=1}^J w_{sa_i,at_j} = 1$ ),  $o_{da_i}^t[at_j]$  est la valeur dans l’offre  $o_{da_i}^t$  du  $j^{\text{ème}}$  attribut (i.e.  $at_j$ ) et  $\mu_{sa_i,at_j}$  est la fonction d’utilité de cet attribut. Cette fonction est définie comme suit :

$$\mu_{sa_i,at_j}(o_{da_i}^t[at_j]) = \frac{rv_{sa_i,at_j} - o_{da_i}^t[at_j]}{rv_{sa_i,at_j} - pv_{sa_i,at_j}} \quad (2)$$

où  $\mu_{sa_i,at_j}(o_{da_i}^t[at_j]) \in [0, 1]$  et  $pv_{sa_i,at_j}$  et  $rv_{sa_i,at_j}$  sont respectivement la valeur préférée (i.e. la meilleure) et la valeur de réserve (i.e. la pire) de  $sa_i$  pour cet attribut<sup>1</sup>.

1. Comme discuté dans [14, 13],  $\mu_{sa_i,at_j}$  peut également être une fonction logarithmique dérivée de la loi Weber-Fechner Law [30] et de l’hypothèse logarithmique [21]. Cependant, les résultats du mécanisme adaptatif de cet article sont valides pour  $\mu_{sa_i,at_j}$  linéaire (Equation 2) ou logarithmique.

Notons que, comme l’ont confirmé des études empiriques, l’utilisateur est capable de choisir les valeurs de  $pv_{at_j}$ ,  $rv_{sa_i,at_j}$  et  $w_{sa_i,j}$  [23, 14, 33]. Par conséquent,  $sa_i$  peut les obtenir de l’utilisateur. Ces valeurs ne sont pas divulguées au fournisseur.

Pour prendre sa décision d’acceptation ou de rejet, un  $sa_i$  s’appuie sur sa fonction d’utilité et sur son *taux d’aspiration* (AR) courant.  $AR_{sa_i}^t \in [0, 1]$  indique la quantité d’utilité que  $sa_i$  s’attend à obtenir dans ce cycle de négociation  $t$ . Pour parvenir à des accords,  $sa_i$  fait des concessions en réduisant son AR. Tous les  $sa_i$  suivent une stratégie de concession temporelle (ou TBC) [5]. Cette hypothèse est assez courante dans la littérature, notamment dans les travaux dont l’objectif est de construire un modèle du comportement de l’adversaire [3]. Par conséquent,  $\Delta AR_{sa_i}^t$ , la concession faite par  $sa_i$  pour le cycle  $t$ , dépend du temps restant avant l’échéance. Il est calculé comme suit [27] :

$$\Delta AR_{sa_i}^t = AR_{sa_i}^{t-1} \cdot \left( \frac{t}{T_{sa_i}} \right)^{\lambda_{sa_i}} \quad (3)$$

où  $T_{sa_i}$  est la date limite de négociation, et  $\lambda_{sa_i}$  est un paramètre qui contrôle le degré de convexité de la courbe de concession de  $sa_i$ .  $\lambda_{sa_i}$  détermine le comportement de  $sa_i$  (conciliant, linéaire ou conservateur) [27].

**Agents délégués (DA ou  $da_i$ ).** Une fois qu’un nouvel agent utilisateur  $sa_i$  entre dans le système, le coordinateur crée un nouvel agent délégué  $da_i$  et initie une session de négociation bilatérale avec  $sa_i$ . Comme  $sa_i$ ,  $da_i$  a une fonction d’utilité  $M_{da_i} \in [0, 1]$ . Cependant, l’utilité d’une offre  $o_{sa_i}^t$  du point de vue du délégué  $da_i$  est déterminée par le coût requis pour servir l’offre.  $RC$  est une valeur clé qui détermine le comportement de négociation d’un délégué. Il représente le coût de réserve (maximum) que le délégué dépense sur les utilisateurs. Cette valeur, initialisée par le coordinateur lorsque le  $da_i$  est généré, est très importante puisqu’elle permet au fournisseur d’imposer sa contrainte budgétaire, i.e. le coût moyen dépensé pour un utilisateur ne doit pas dépasser  $RC$ . Si l’offre n’est pas acceptée par un  $da_i$ , il utilise sa stratégie de négociation pour générer une contre-offre. Pour conclure des ententes avec  $sa_i$ , le  $da_i$  utilise une stratégie de concession fondée sur le temps qui réduit son taux d’aspiration (AR). La concession est calculée comme suit :

$$\Delta AR_{da_i}^t = \frac{1}{T_{da_i}} \quad (4)$$

où  $T_{da_i}$  est l’échéance du délégué, initialisée par le  $ca$  en fonction de ses connaissances métier. Lorsque  $da_i$  atteint  $T_{da_i}$ , il ne quitte pas le processus de négociation, il cessera simplement de faire des concessions.

## 2.2 Négociation bilatérale entre SA et DA

Dans l’architecture EMan, les SA et DA n’ont pas accès aux préférences et aux stratégies de négociation des autres agents. Par conséquent, leur comportement de négociation suit la *fonction de décision de négociation* [5] et est déterminé par la fonction d’utilité et la stratégie de négociation.

Quant au coordinateur, il est capable de manipuler les stratégies de négociation des délégués si le mode d'adaptation est actif.

Le service négocié est caractérisé par un ensemble de  $J$  attributs,  $s = \langle at_1, at_2, \dots, at_J \rangle$ . L'objet  $o_i^t$  est l'offre ou contre-offre échangée pendant le processus de négociation à l'étape  $t$ , entre  $sa_i$  et  $da_i$ . Il attribue une valeur  $v_k^t$  à chacun des attributs décrivant le service négocié. La négociation entre un  $da_i$  et un  $sa_i$  est basée sur le *protocole d'offre alternative*. Lorsqu'un  $sa_i$  reçoit une offre, il peut (i) accepter l'offre, (ii) proposer une contre-offre, ou (iii) quitter le processus de négociation. Une fois qu'une offre est acceptée par un agent, elle ne peut pas être reniée [13].

Afin de tenir compte de la nature ouverte et dynamique de l'écosystème *cloud*, les sessions de négociation dans l'architecture EMan sont non synchrones, i.e. certaines sessions seront déjà terminées alors que d'autres sessions seront encore actives ou n'auront pas encore démarré. Pour plus d'informations sur le protocole de négociation et la stratégie d'acceptation, veuillez vous référer à [13, 14].

Cette section offrait un aperçu de l'architecture EMan. Nous savons que les protocoles de négociation et les stratégies présentés ci-dessus peuvent être exploitables. Toutefois, cette question dépasse le cadre de cet article. Nous considérons que cette question est d'une importance moindre pour la raison suivante : la négociation dans l'écosystème *cloud* a une forte composante *intégrative* puisque les ressources offertes par le *cloud* semblent être illimitées [29]. Par conséquent, au lieu d'exploiter les stratégies de négociation des utilisateurs et de les forcer à faire de lourdes concessions, le fournisseur est plus susceptible de rechercher des règlements gagnant-gagnant permettant de maximiser le taux d'acceptabilité tout en respectant ses contraintes budgétaires.

### 3 Stratégie de négociation « one-to-many » adaptative

Cette section détaille notre algorithme adaptatif. Afin d'estimer le taux d'acceptation à tout moment  $t$ , le fournisseur s'appuie sur un algorithme d'estimation de quantile présenté dans la section 3.1. Si le taux d'acceptation ciblé est violé, le coordinateur active le mode adaptatif. Avec ce mode actif, un délégué  $da_i$  doit analyser le comportement de son « adversaire »<sup>2</sup>  $sa_i$ , apprendre un modèle de sa stratégie de négociation afin d'estimer son délai de négociation  $T_{sa_i}^t$  et envoyer cette estimation au coordinateur (section 3.2). Ce dernier choisit les sessions de haute priorité (en fonction de leurs délais) et ajuste leurs stratégies de négociation tout en respectant les contraintes budgétaires (section 3.3).

2. Notons les guillemets autour de la notion d'adversaire, terme utilisé en négociation, bien qu'ici le délégué n'ait pas forcément de relation conflictuelle avec l'utilisateur.

#### 3.1 Déclenchement de l'adaptation

Les quantiles et les percentiles sont des valeurs qui partagent un ensemble fini de valeurs en  $q$  sous-ensembles de tailles (presque) égales. La littérature sur la QoE et la satisfaction de l'utilisateur recommande aux fournisseurs de s'appuyer sur les quantiles et les percentiles comme mesures plus précises (par rapport à MOS) pour évaluer l'acceptabilité du service [7].

Chaque fois que la session  $i$  est terminée, le coordinateur est informé par  $da_i$  des résultats de cette session. En utilisant ces données, le coordinateur exécute un algorithme d'estimation de quantile pour détecter le taux d'acceptabilité du service courant.

Soit  $Q$  la fonction d'estimation quantile ou percentile. Soit  $R$  le jeu de données contenant les résultats des sessions terminées.  $R$  peut contenir soit 0, pour les sessions échouées, soit 1 pour les sessions réussies. Si le coordinateur cherche à s'assurer que  $\beta$  pourcents des utilisateurs qui ont fait appel au service à ce jour ont eu une session de négociation réussie (acceptant la qualité de service proposée), le coordinateur doit vérifier que le  $(100 - \beta + 1)^{\text{ème}}$  percentile est égal à 1 :

$$Q(R, 100 - \beta + 1) = 1 \quad (5)$$

Tant que cette condition est maintenue, le coordinateur n'a pas besoin d'intervenir dans le processus de négociation. Dès que la condition de l'équation 5 est violée, le coordinateur déclenche le mécanisme d'adaptation en commandant à tous les délégués actifs d'activer leur mode adaptatif. De plus, lorsque le coordinateur engendre de nouveaux délégués, ils entreront également dans ce mode actif.

Notez que le coordinateur continue d'évaluer l'équation 5 même après l'activation du mode adaptatif.

La section suivante détaille le rôle des délégués après l'activation du mode adaptatif.

#### 3.2 Apprentissage du comportement de l'adversaire

**Estimation de la concession de l'adversaire.** Même lorsque le mode adaptatif n'est pas actif, lorsqu'un délégué  $da_i$  reçoit  $o_{sa_i}^t$  au cycle  $t$  du  $sa_i$  correspondant, il estime la concession faite par  $sa_i$  en comparant  $o_{sa_i}^t$  avec  $o_{sa_i}^{t-1}$  l'offre précédente faite par  $sa_i$ . Puisque  $da_i$  n'a pas accès aux préférences de  $sa_i$  ou à la fonction  $M_{sa_i}$ , il ne peut pas calculer la concession réelle de  $sa_i$ . Au lieu de cela, il s'appuie sur sa propre fonction d'utilité pour estimer la concession faite par  $sa_i$  en supposant qu'une concession faite par  $sa_i$  est synonyme d'un gain d'utilité pour  $da_i$ . Ainsi, l'estimation de la concession faite par  $sa_i$  au cycle de négociation  $t$ , est définie comme suit :

$$c_{sa_i}^t = M_{da_i}(o_{sa_i}^t) - M_{da_i}(o_{sa_i}^{t-1}) \quad (6)$$

Pour apprendre le comportement de concession de  $sa_i$ ,  $da_i$  effectue le suivi des concessions de  $sa_i$  pendant la session de négociation. La figure 2 compare  $\Delta AR_{sa_i}^t$ , les concessions réelles faites par  $sa_i$  (en bleu), avec  $c_{sa_i}^t$ , l'estimation par  $da_i$  de la même concession (en rouge). Dans cet

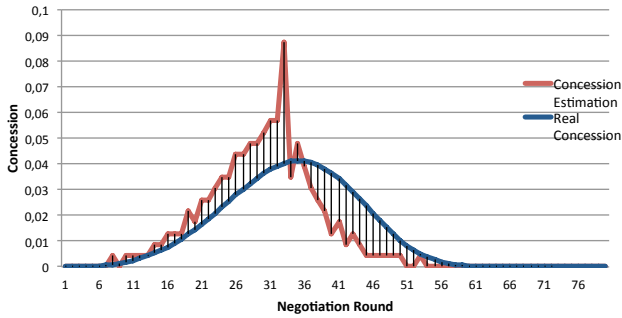


FIGURE 2 –  $\Delta AR_{sa_i}^t$ , la vraie concession faite par  $sa_i$  (en bleu) comparée à  $c_{sa_i}^t$ , l'estimation faite par  $da_i$  (en rouge).

exemple,  $T_{sa_i} = 80$  cycles et  $\lambda_{sa_i} = 3.0$ . Comme le montre la figure, bien que la courbe rouge soit bruitée (car elle est estimée à l'aide de la fonction d'utilité de  $da_i$ , puis normalisée), elle semble fournir une estimation adéquate de la concession de  $sa_i$  puisque l'objectif principal est d'apprendre  $T_{sa_i}$ , la date limite de négociation de  $sa_i$ .

#### Apprentissage du modèle de l'adversaire.

Si le mode d'adaptation est actif, les données de concessions recueillies dans le paragraphe précédent devraient être utilisées pour établir un modèle de comportement de concession de  $sa_i$  et inférer  $T_{sa_i}$ . Ceci peut être réalisé, comme cela a été montré dans la littérature [3], en utilisant une régression non linéaire supposant que tous les utilisateurs suivent une concession basée sur le temps définie par l'équation 3. Cependant, à la différence des travaux existants prédominants dans la littérature, où apprendre le modèle de négociation de l'adversaire est fait et corrigé à chaque cycle une fois qu'une nouvelle concession est faite, dans les arrangements de négociation « one-to-many » abordée dans cet article, une telle approche n'est pas pratique pour des raisons d'évolutivité. En effet, des centaines ou des milliers d'utilisateurs peuvent négocier avec le fournisseur simultanément. En outre, si le fournisseur exécute les algorithmes d'apprentissage des délégués sur les ressources louées sur le *cloud*, cela peut augmenter considérablement les coûts.

Pour cette raison, dans la solution proposée, un délégué peut exécuter l'algorithme de régression non linéaire une seule fois. Il s'agit donc maintenant de déterminer quand un délégué doit le faire. D'une part, si un délégué lance ce processus trop tôt dans la session de négociation, le processus de régression n'aura que quelques points d'entrée. D'autre part, si le délégué attend trop longtemps, l'utilisateur correspondant risque d'atteindre son échéance et de quitter le processus de négociation.

Tous les  $sa_i$  suivent une stratégie de concession fondée sur le temps dans laquelle le taux de concessions de  $sa_i$  ralentit quand  $sa_i$  a déjà fait la plupart des concessions qu'il pouvait faire. Par conséquent, quel que soit le type de  $sa_i$  (conservateur ou linéaire) et de son échéance réelle  $T_{sa_i}$ , lorsque le taux de changement (ou la dérivée) de l'estima-

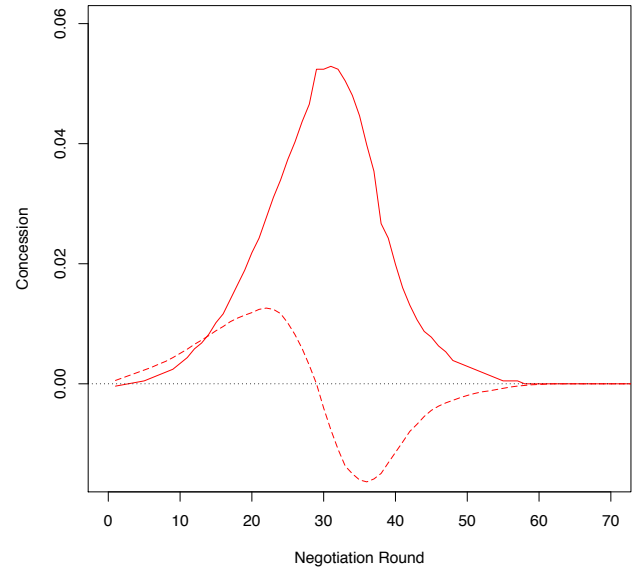


FIGURE 3 – Une version lissée de  $c_{sa_i}^t$  (ligne pleine) et sa dérivée (ligne pointillée).

tion par  $da_i$  de la concession de  $sa_i$  diminue significativement en valeurs négatives, cela signifie que  $sa_i$  a fait la plupart de ses concessions et que, pour le reste de la session de négociation avant que  $sa_i$  atteigne  $T_{sa_i}$ ,  $sa_i$  cessera de faire des concessions considérables. Ainsi, si  $da_i$  lance sa régression non linéaire à ce stade, il aura le plus de points de données utiles.

La figure 3 superpose  $c_{sa_i}^t$  avec sa première dérivée. Le  $sa_i$  de cette figure est exactement le même que celui de la figure 2, mais nous avons utilisé un filtre Savitzky-Golay pour lisser la courbe et filtrer le bruit [24]. Pour calculer le taux de variation des concessions, nous calculons la dérivée et nous lissons le résultat en utilisant une variante à du même filtre.

L'algorithme de régression non linéaire prend comme variables d'entrée le cycle de négociation ( $t$ ) et la concession estimée ( $c_{sa_i}^t$ ) et émet des valeurs estimées des paramètres  $T_{sa_i}$  et  $\lambda_{sa_i}$ , notée  $\tilde{T}_{sa_i}$  et  $\tilde{\lambda}_{sa_i}$  respectivement. Alors, d'après  $\tilde{T}_{sa_i}$ ,  $da_i$  calcule  $\tilde{r}_i = \tilde{T}_{sa_i} - t$ , le nombre estimé de cycles<sup>3</sup> restant dans la session  $i$ . Tant que la session  $i$  n'est pas terminée,  $da_i$  met à jour  $\tilde{r}_i$  à chaque cycle et l'envoie à plusieurs reprises au coordinateur.

### 3.3 Adaptation de la négociation

Comme nous l'avons expliqué dans la section précédente, le coordinateur reçoit  $\tilde{r}_i$  de chaque session de négociation  $i$  dont le délégué  $da_i$  considère que  $sa_i$  s'approche de la fin de son échéancier et mérite d'être priorisé. Ces estimations  $\tilde{r}_i$  sont stockées dans une *liste des priorités* qui est triée en continu par ordre croissant : une session dont la date limite est estimée plus tôt sera en haut de la liste.

3. Les termes *cycle* et *round* sont utilisés indifféremment ici.

Notons que puisque les sessions de négociation sont non synchronisées, le coordinateur doit répéter le tri chaque fois qu'il reçoit une nouvelle estimation d'un délégué  $da_i$ . En outre, chaque fois qu'une session est terminée avec succès ou non, le coordinateur supprime son enregistrement de la liste des priorités. Ainsi, dans la liste des priorités, les sessions sont triées de la plus haute à la plus basse des priorités. Maintenant le coordinateur doit décider combien de ces sessions prioritaires doivent être *priorisées*. Pour ce faire, le coordinateur entreprend ce processus comme suit :

1. Estimer le taux d'acceptabilité courant (en utilisant la fonction d'estimation de quantile comme décrit dans la section 3.1). Ensuite, calculer la différence entre le taux d'acceptabilité souhaité et sa valeur courante actuelle.
2. Estimer  $p$  le nombre de sessions réussies requises pour rétablir le taux désiré et choisir les premières  $p$  sessions de la liste de priorités et les *prioriser* en ajustant leurs stratégies de négociation pour les encourager à accepter le service.

Lorsque la stratégie de négociation d'un délégué  $da_i$  est ajustée,  $da_i$  utilise une stratégie de concession fondée sur le temps définie dans la section 2.1, affectées des modifications suivantes. Tout d'abord, le délai de négociation de  $da_i$  devient  $\tilde{r}_i$  au lieu de  $T_{da_i}$ . Deuxièmement, le coût de réserve de  $da_i$  est augmenté de la valeur notée  $RcPrio$ .  $RcPrio$  est calculé en divisant l'excédent disponible dans *Surplus* parmi toutes les sessions priorisées. *Surplus* est une variable dans laquelle le coordinateur accumule le surplus obtenu lors de sessions de négociation réussies. Plus précisément, lorsqu'une session est déclarée réussie, le coordinateur met à jour sa variable *Surplus* comme suit :

$$Surplus \leftarrow Surplus + surplus_i \quad (7)$$

où  $surplus_i$  est l'excédent obtenu de la session de négociation réussie  $i$ .  $surplus_i$  est calculé comme suit :

$$surplus_i = RC - Cost(\Pi(\hat{o}_i)) \quad (8)$$

Où  $RC$  est le coût de réserve des délégués,  $\hat{o}_i$  est l'offre acceptée dans la session  $i$  et  $Cost(\Pi(\hat{o}_i))$  est le coût à payer par le fournisseur pour satisfaire à cette offre. Ainsi, la différence entre le coût réel et le coût (maximum) de la réserve est considérée comme excédentaire. Notez que les délégués ne peuvent pas obtenir plus de  $RcPrio_{max} = RC/2$ , considéré comme la valeur maximale pour les sessions priorisées. Troisièmement, le coût préféré de  $da_i$  est changé pour le coût de la dernière offre qu'il a faite.  $AR_{da_i}^t$  est remis à 1.

Cette section a présenté la modélisation de l'adversaire et les processus d'adaptation de la négociation. La section suivante présente l'évaluation expérimentale de notre approche

## 4 Évaluation expérimentale

Pour évaluer le mécanisme adaptatif, nous l'implémentons dans l'architecture EMan [14, 13]. Cette dernière est

une architecture multi-agent pour la gestion de l'élasticité SaaS mise en œuvre à l'aide de Repast Symphony [18], un environnement de simulation multi-agent. L'évaluation est organisée en trois expériences. La première expérience (section 4.2) vise à évaluer l'algorithme d'adaptation. La deuxième expérience (section 4.3) examine l'impact de la charge de travail appliquée au fournisseur (c'est-à-dire le nombre d'utilisateurs entrant dans le système par minute) sur le taux d'acceptabilité. La troisième expérience (section 4.4) évalue la surcharge des mécanismes de négociation et coordination.

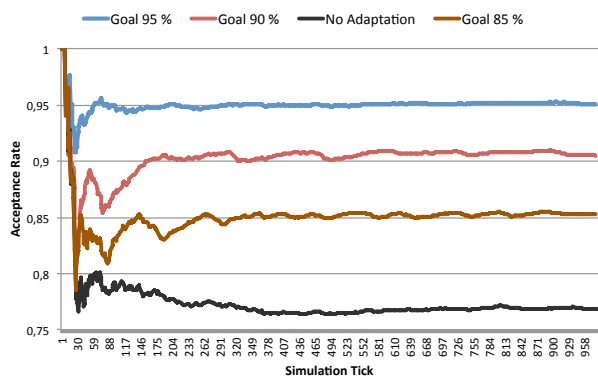
### 4.1 Paramètres expérimentaux

Le nombre total d'utilisateurs entrant dans le système est noté  $|SU|$ . Dans les expériences suivantes,  $|SU| = 10000$  utilisateurs. Les utilisateurs entrent dans la simulation suivant un processus aléatoire de Poisson dont la valeur moyenne est  $A$  par minute. Le service dans le scénario expérimental implique deux attributs : l'un est le délai de livraison du service tandis que l'autre représente la qualité du service. Les profils utilisateur (réserve, valeurs préférées et pondérations pour chaque attribut) sont générés au hasard. Les délais de négociation des  $sa_i$  sont générés au hasard :  $T_{sa_i} \in [40 : 120]$ , et  $\lambda_{sa_i} \in [1 : 8]$  (i.e. les utilisateurs peuvent être linéaires ou conservateurs). Le coût des services acceptables par les utilisateurs varie de 0.1 à 0.9.  $RC$ , le coût de réserve du délégué (le coût maximal alloué à un utilisateur non prioritaire) est fixé à 0.6. Ce paramètre représente les contraintes budgétaires du fournisseur. Par conséquent, sans le mode d'adaptation, environ un quart des utilisateurs n'accepteront pas le service puisque  $RC$  ne peut pas satisfaire leur service le moins acceptable,  $RcPrio_{max} = RC/2$ . Par conséquent, lorsqu'un utilisateur  $sa_i$  est prioritaire, le  $RC$  du délégué respectif  $da_i$  est au plus augmenté de  $RC/2$ . *Goal* est le pourcentage d'utilisateurs que le fournisseur cherche à satisfaire. Son impact est évalué dans la section suivante. Notez que les résultats décrits ci-dessous restent valides même si les valeurs des paramètres ci-dessus (par exemple,  $RC$ ,  $RcPrio_{max}$ ) sont modifiées. De plus, nous avons obtenu des résultats similaires lorsque la fonction d'utilité d'attribut d'utilisateur est une fonction logarithmique dérivée de l'hypothèse logarithmique [21, 14].

### 4.2 Évaluation du mécanisme d'adaptation

La figure 4 montre les résultats de cette expérience. La courbe bleue représente le taux d'acceptation lorsque  $Goal = 95\%$ . Comme le montre la figure, au début de la simulation, la valeur du taux d'acceptation a oscillé avant de se stabiliser sur la valeur d'objectif ( $Goal = 95\%$ ). Les courbes rouge et marron, qui tracent le taux d'acceptation lorsque  $Goal = 90\%$  et  $Goal = 85\%$ , montrent des résultats similaires. Pour comparer les résultats du mécanisme d'adaptation, la courbe noire trace le taux d'acceptation lorsque le mécanisme est désactivé.

Comme on peut le constater à partir de ces résultats, le mécanisme d'adaptation a réussi à rétablir le taux d'accep-

FIGURE 4 – Taux d'acceptation avec  $A = 160$ .TABLE 1 – Impact de la charge de travail  $A$  sur le taux d'acceptation

$A$	20	40	80	160	320	640	1280	2500
Avec adaptation	95%	95%	95%	95%	95%	94.9%	94.7%	93.3%
Sans adaptation	77%	77%	77%	77%	77%	77%	77%	77%

tation de l'objectif prédéfini  $Goal$  d'une manière précise. Ceci est expliqué par le fait que l'algorithme de prédiction discuté dans la section 3.2 a réussi à prédire le délai de négociation utilisateur  $T_{sa_i}$  avec une précision acceptable. Le taux d'erreur moyen entre la prédiction  $\tilde{T}_{sa_i}$  et la valeur réelle  $T_{sa_i}$  est d'environ 6 cycles, ce qui signifie qu'un délégué  $da_i$  surestime/sous-estime  $T_{sa_i}$  par trois cycles de négociation moyenne.

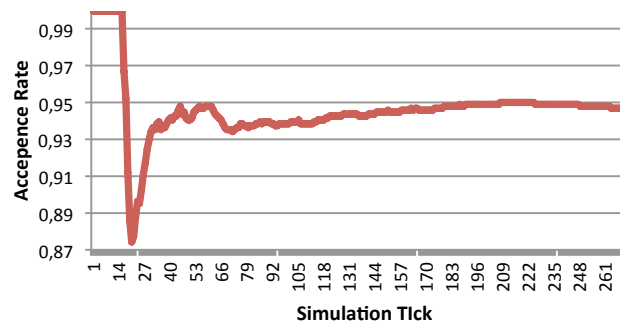
Intuitivement, le taux d'acceptabilité de service atteint par l'algorithme de prédiction est livré avec plus de coûts investis par utilisateur. Cependant, cette augmentation ne viole pas la contrainte budgétaire du fournisseur : le coût moyen par utilisateur était de 0.55, 0.52, 0.5 et 0.44 pour  $Goal = 95\%$ ,  $Goal = 90\%$ ,  $Goal = 85\%$  et non-adaptatif, respectivement. Ainsi, la contrainte budgétaire du fournisseur (i.e. le coût moyen par utilisateur ne dépasse pas  $RC = 0.6$ ) a été satisfaite parce que le fournisseur s'appuie sur le *Surplus* pour servir les utilisateurs priorités.

### 4.3 Impact de la charge $A$

Cette expérience étudie l'impact de  $A$ , la charge de travail appliquée au système, sur le taux d'acceptation afin d'évaluer l'élasticité du mécanisme d'adaptation.

La table 1 montre le taux d'acceptation, à la fin de la simulation, lorsque  $Goal = 95\%$ <sup>4</sup> et  $A$  augmentant de 20 à 2500 utilisateurs par minute. Comme le montre le tableau, le mécanisme s'est avéré très élastique. Lorsque  $A \in [20 : 320]$  le taux d'acceptation ne change pas à mesure que  $A$  augmente. Toutefois, lorsque  $A$  augmente à des valeurs plus élevées (640 à 2500), le taux d'ac-

4. Nous avons effectué la même expérience avec différentes valeurs de  $Goal \in \{92\%, 90\%, \dots\}$  et avons obtenu des résultats similaires.

FIGURE 5 – Taux d'acceptation ( $Goal=95\%$  et  $A = 1280$ ).

ceptation témoigne d'une légère diminution. Cependant, même avec un très haut  $A$ , les résultats obtenus par le mécanisme d'adaptation restent très proches de l'objectif. Cette diminution est expliquée comme suit : l'algorithme d'adaptation est réactif et le coordinateur ne peut évaluer le taux d'acceptation qu'à chaque fois qu'une session se termine avec succès ou sans succès. Ainsi, lorsque  $A$  est trop élevé, un nombre significatif de sessions peut échouer dans le même cycle de simulation et l'algorithme d'adaptation exécuté par le coordinateur souffre d'un léger retard et devient un peu court pour répondre à l'objectif prédéfini ( $Goal = 95\%$ ), comme on peut le voir dans le tableau. Pourtant, ce résultat est à relativiser pour la raison suivante :  $|SU|$ , le nombre total d'utilisateurs dans la simulation, influence le seuil d'élasticité du mécanisme d'adaptation. Par exemple, lorsque  $|SU| = 10000$  et  $A = 2500$ , les utilisateurs entrent dans le système en 4 vagues massives représentant environ un quart du nombre total d'utilisateurs. Ainsi, le coordinateur n'a pas assez de temps pour rétablir le taux d'acceptation à ( $Goal = 95\%$ ). La figure 5 montre comment le mécanisme d'adaptation réagit à une chute du taux d'acceptation lorsque  $A = 1280$ . Le coordinateur parvient à restaurer la valeur de l'objectif prédéfini ( $Goal = 95\%$ ). Pourtant, une nouvelle chute se produit avec la nouvelle vague d'utilisateurs et la simulation est terminée. Notez que dans un scénario réaliste, lorsque  $|SU|$  tend vers l'infini, le coordinateur sera toujours en mesure de restaurer le taux d'acceptation dans sa valeur d'objectif, même avec un  $A$  significatif.

### 4.4 Surcoût de la coordination et de la négociation

Pour évaluer les surcoûts induits par la coordination et les mécanismes de négociation, nous comptons le nombre de messages échangés pour entreprendre l'intervention du coordinateur et le nombre de messages échangés respectivement lors des sessions de négociation bilatérales.

Sans le mécanisme d'adaptation, le coordinateur est sollicité une seule fois par session pour obtenir le résultat de la session (succès ou échec). Avec le mécanisme adaptatif, le nombre de messages échangés entre les délégués et le coordinateur augmente d'environ 50%. Nous consi-

dérons que cette augmentation est insignifiante pour les raisons suivantes. Tout d'abord, même avec le mécanisme adaptatif, les tâches assumées par le coordinateur sont légères comme cela a été discuté dans les sections 3.1 et 3.3. Deuxièmement, étant donné qu'il est probable que le fournisseur SaaS gère les délégués et le coordinateur sur le même centre de données dans le *cloud*, le coût de communication entre eux est négligeable.

En ce qui concerne le nombre de messages de négociation, les résultats montrent une légère diminution ( $\approx 7\%$ ) lorsque le mécanisme adaptatif est actif, car ce dernier aide à conclure des accords plus rapidement.

## 5 Travaux connexes

En dépit de la littérature relativement riche sur les négociations « one-to-many », l'objectif de la plupart de ces travaux, modélisant des négociations entre un acheteur et de nombreux concurrents vendeurs, est de trouver un accord unique (i.e. *atomique*) qui maximise l'utilité de l'acheteur tandis que d'autres sessions, moins bénéfiques, sont avortées [10, 20]. En outre, les mécanismes de négociation et de coordination existants (par exemple [20, 10]) sont principalement *fermés* et *synchrones*. Par conséquent, ces solutions ne peuvent pas tenir compte de la nature agile de l'écosystème où des milliers d'utilisateurs peuvent surgir sur le portail de services.

Les travaux récents dans le domaine de la composition du service utilisent la négociation « one-to-many » pour atteindre des accords *composite* avec plus d'un vendeur de services atomiques afin de construire un service composite [22, 11]. Dans [22], les auteurs proposent un mécanisme de négociation « one-to-many » qui redistribue le surplus obtenu lors des sessions de négociation réussies aux sessions de négociation en cours afin d'augmenter leurs chances de succès. Néanmoins, ces travaux supposent l'ensemble de vendeurs atomiques *fermé* : ils sont tous connus avant le début de la négociation.

La négociation « one-to-many » a également été utilisée dans le domaine du *cloud computing*. Les auteurs de [2] développent une approche basée sur la négociation pour gérer l'allocation des ressources dans ce cadre. Cependant, un fournisseur accepte une offre si et seulement s'il peut obtenir un gain immédiat en acceptant l'offre. Pour cette raison, le taux d'acceptabilité des utilisateurs n'est pas pris en compte et aucun mécanisme d'adaptation n'est proposé. [26] développent un mécanisme de négociation simultanée de SLA dans les systèmes basés sur le *cloud*. Pourtant, la principale contribution du travail semble être axée sur le protocole. Par conséquent, il ne fournit pas un mécanisme pour représenter l'acceptabilité des utilisateurs ni offre une solution pour ajuster les comportements de négociation des délégués pour tenir compte des objectifs du fournisseur.

## 6 Conclusions

Cet article a présenté un mécanisme de négociation « one-to-many » adaptatif conçu pour améliorer le taux d'ac-

ceptabilité d'un fournisseur SaaS tout en respectant ses contraintes budgétaires. L'approche proposée confère au fournisseur un contrôle précis du taux d'acceptabilité souhaité. En outre, comme cela a été montré expérimentalement, avec la solution proposée le fournisseur de SaaS est capable de faire face aux pics de charge de manière adéquate. Notre travail de recherche futur visera à donner au fournisseur un contrôle plus fin sur le niveau de satisfaction des utilisateurs qu'il cherche à atteindre. En particulier, le fournisseur doit être en mesure de s'assurer qu'un pourcentage prédéfini d'utilisateurs considère le service comme étant « *bon* » ou « *meilleur* » [7]. Plusieurs rapports de l'ITU (Union internationale des télécommunications) et de l'ETSI (Institut européen des normes de télécommunications) recommandent d'aller dans cette direction [7].

## Références

- [1] Accenture. Accenture 2013 global consumer pulse survey global & u.s. key findings. report, Dublin, Ireland, 2013.
- [2] B. An, V. Lesser, D. Irwin, and M. Zink. Automated negotiation with decommitment for dynamic resource allocation in cloud computing. In *Proceedings of the 9th International Conference on Autonomous Agents and Multiagent Systems : volume 1-Volume 1*, pages 981–988, 2010.
- [3] T. Baarslag, M. J. Hendrikx, 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*, pages 1–50, 2015.
- [4] E. F. Coutinho, F. R. de Carvalho Sousa, P. A. L. Rego, D. G. Gomes, and J. N. de Souza. Elasticity in cloud computing : a survey. *annals of telecommunications-Annales des télécommunications*, 70(7-8) :289–309, 2015.
- [5] P. Faratin, C. Sierra, and N. R. Jennings. Negotiation decision functions for autonomous agents. *Robotics and Autonomous Systems*, 24(3) :159–182, 1998.
- [6] T. Hobfeld, R. Schatz, and S. Egger. Sos : The mos is not enough ! In *Quality of Multimedia Experience (QoMEX), 2011 Third International Workshop on*, pages 131–136. IEEE, 2011.
- [7] T. Hobfeld, P. E. Heegaard, M. Varela, and S. Möller. Qoe beyond the mos : an in-depth look at qoe via better metrics and their relation to mos. *Quality and User Experience*, 1(1) :2, 2016.
- [8] C. V. N. Index. Cisco vni forecast and methodology, 2015-2020. cisco white paper, june 1, 2016, 2016.
- [9] A. R. Lomuscio, M. Wooldridge, and N. R. Jennings. A classification scheme for negotiation in electronic commerce. *Group Decision and Negotiation*, 12(1) :31–56, 2003.
- [10] K. Mansour and R. Kowalczyk. A meta-strategy for coordinating of one-to-many negotiation over multiple issues. In *Foundations of Intelligent Systems*, pages 343–353. Springer, 2012.
- [11] K. Mansour and R. Kowalczyk. On dynamic negotiation strategy for concurrent negotiation over distinct objects. In *Novel Insights in Agent-based Complex Automated Negotiation*, pages 109–124. Springer, 2014.

- [12] S. Möller and A. Raake. *Quality of Experience*. Springer, 2014.
- [13] A. Najjar. *Multi-Agent Negotiation for QoE-Aware Cloud Elasticity Management*. PhD thesis, École nationale supérieure des mines de Saint-Étienne, 2015.
- [14] A. Najjar, C. Gravier, X. Serpaggi, and O. Boissier. Modeling user expectations satisfaction for saas applications using multi-agent negotiation. In *2016 IEEE/WIC/ACM International Conference on Web Intelligence (WI)*, pages 399–406, Oct 2016.
- [15] A. Najjar, X. Serpaggi, C. Gravier, and O. Boissier. Survey of elasticity management solutions in cloud computing. In *Continued Rise of the Cloud*, pages 235–263. Springer, 2014.
- [16] A. Najjar, X. Serpaggi, C. Gravier, and O. Boissier. Une négociation multi-partite pour une gestion élastique des applications hébergées dans un cloud. In *Journées Francophones Systèmes Multi-Agent*, Oct. 2014.
- [17] A. Najjar, X. Serpaggi, C. Gravier, and O. Boissier. Multi-agent systems for personalized qoe-management. In *Tetraffic Congress (ITC 28), 2016 28th International*, volume 3, pages 1–6. IEEE, 2016.
- [18] M. J. North, T. R. Howe, N. T. Collier, and J. Vos. The repast symphony runtime system. In *Agent 2005 Conference on Generative Social Processes, Models, and Mechanisms*, 2005.
- [19] D. G. Pruitt. *Negotiation behavior*. Academic Press, 2013.
- [20] I. Rahwan, R. Kowalczyk, and H. H. Pham. Intelligent agents for automated one-to-many e-commerce negotiation. In *Australian Computer Science Communications*, volume 24, pages 197–204. Australian Computer Society, Inc., 2002.
- [21] P. Reichl, S. Egger, R. Schatz, and A. D’Alconzo. The logarithmic nature of qoe and the role of the weber-fechner law in qoe assessment. In *Communications (ICC), 2010 IEEE International Conference on*, pages 1–5. IEEE, 2010.
- [22] J. Richter, M. Baruwal Chhetri, R. Kowalczyk, and Q. Bao Vo. Establishing composite slas through concurrent qos negotiation with surplus redistribution. *Concurrency and Computation : Practice and Experience*, 24(9) :938–955, 2012.
- [23] A. Sackl and R. Schatz. Evaluating the impact of expectations on end-user quality perception. In *Proceedings of International Workshop Perceptual Quality of Systems (PQS)*, pages 122–128, 2013.
- [24] A. Savitzky and M. J. Golay. Smoothing and differentiation of data by simplified least squares procedures. *Analytical chemistry*, 36(8) :1627–1639, 1964.
- [25] R. Schatz, M. Fiedler, and L. Skorin-Kapov. Qoe-based network and application management. In *Quality of Experience*, pages 411–426. Springer, 2014.
- [26] M. Siebenhaar, U. Lampe, D. Schuller, R. Steinmetz, et al. Concurrent negotiations in cloud-based systems. In *Economics of Grids, Clouds, Systems, and Services*, pages 17–31. Springer, 2012.
- [27] S. Son and K. M. Sim. A price-and-time-slot-negotiation mechanism for cloud service reservations. *IEEE Transactions on Systems, Man, and Cybernetics*, 42(3) :713–728, 2012.
- [28] D. Talia. Clouds meet agents : Toward intelligent cloud services. *IEEE Internet Computing*, (2) :78–81, 2012.
- [29] L. L. Thompson, J. Wang, and B. C. Gunia. Negotiation. *Annual review of psychology*, 61, 2010.
- [30] L. L. Thurstone. Psychophysical analysis. *The American journal of psychology*, 38(3) :368–389, 1927.
- [31] I. Wechsung and K. De Moor. Quality of experience versus user experience. In *Quality of Experience*, pages 35–54. Springer, 2014.
- [32] M. Wooldridge. *An introduction to multiagent systems*. John Wiley & Sons, 2009.
- [33] V. A. Zeithaml, L. L. Berry, and A. Parasuraman. The nature and determinants of customer expectations of service. *Journal of the academy of Marketing Science*, 21(1) :1–12, 1993.





# Détection de liens d'identité contextuels dans une base de connaissances

Joe Raad

Nathalie Pernelle

Fatiha Saïb

LRI, Université Paris-Sud, Bât. 650, F-91405 Orsay, France

prenom.nom@lri.fr

## Résumé

De nombreuses applications du Web de données exploitent des liens d'identités déclarés à l'aide du constructeur owl :sameAs. Cependant, différentes études ont montré qu'une utilisation abusive de ces liens peut conduire à des inférences erronées ou contradictoires. Dans ce papier nous proposons un nouveau lien d'identité contextuel : *identiConTo* qui peut remplacer owl :sameAs à lier des individus identiques dans un contexte spécifié. Pour détecter ces liens contextuels, nous avons défini un algorithme nommé *DECIDE*, et nous avons expérimenté cette approche sur des données scientifiques décrivant des processus de transformations.

## Mots Clef

Liage de données, Contextes, Ontologie, Base de connaissances.

## Abstract

Most of the Linked Data applications currently rely on the use of owl :sameAs for linking ontology instances. However, several studies have noticed multiple misuses of this identity link, which can lead to erroneous statements or inconsistencies. We propose in this paper a new contextual identity link : *identiConTo* that could serve as a replacement for owl :sameAs in linking identical instances in a specified context. To detect these contextual links, we have defined an algorithm named *DECIDE* that has been tested on scientific knowledge bases describing transformation processes.

## Keywords

Linked Data, Contexts, Ontology, Knowledge Base.

## 1 Introduction

Le Linked Open Data cloud est une initiative du W3C, qui définit un ensemble de bonnes pratiques pour publier et lier des données structurées sur le Web. En utilisant des technologies du Web sémantique, des applications peuvent partager, extraire, interroger ou raisonner sur les données publiées. Le Linked Open Data (LOD) a récemment pris une nouvelle dimension avec la publication de grandes quantités de données (le LOD est passé de 500 millions de triplets RDF en 2007 à 130 milliards de triplets en 2016). Ces

données sont encyclopédiques telles que DBpedia, Yago ou encore Google Knowledge Vault ou concernent différents domaines d'application comme les sciences du vivant, la culture ou encore l'économie. Toutefois, si ces données se retrouvent isolées, leur utilité reste très limitée. En effet, un des points angulaires du Web des données est le fait que les données soient liées entre elles par des liens sémantiques tels que les liens d'identité owl :sameAs qui expriment que deux ressources différentes réfèrent à la même entité (e.g., même personne, même article, même gène).

Ce type de liens est défini dans [11] avec une sémantique très stricte qui exprime le fait que déclarer un fait owl :sameAs entre deux objets indique que toutes les valeurs de propriétés déclarées pour l'un peuvent aussi être déclarées pour l'autre. Ainsi, si des faits owl :sameAs erronés sont déclarés dans les bases de connaissances cela peut conduire à inférer des informations erronées et même contradictoires. Dans [8], les auteurs ont évalué la qualité du résultat du liage de données obtenu entre des données DBLP et des données de DBpedia. Pour cela, ils ont mesuré la correction des nouveaux faits inférés en exploitant la sémantique logique des liens owl:sameAs. En choisissant arbitrairement 49 noms parmi les 491 796 auteurs disponibles dans DBLP 2006, ils ont montré que 92 % de ces 49 auteurs ont eu des publications affiliées incorrectement. Par ailleurs, l'absence d'autres types de liens alternatifs avec une sémantique claire, renforce l'utilisation abusive du owl:sameAs. Souvent, la relation d'identité entre objets est plus faible et dépend du contexte dans lequel on veut utiliser l'identité. Prenons l'exemple de deux éditions distinctes du même livre. Dans le cas où l'on s'intéresse à identifier s'il s'agit de la même édition de livre alors ces deux livres seront différents. Cependant, dans un contexte où l'on cherche à savoir s'il s'agit de la même oeuvre artistique alors ces deux livres seront identiques.

Dans cet article, nous proposons une approche de détection de liens d'identité contextuels dans des bases de connaissances RDF<sup>1</sup>. Un lien d'identité contextuel exprime une relation d'identité entre deux instances, valide dans un contexte défini par rapport à une ontologie de domaine. Pour cela, nous avons défini une notion de contexte global composé d'un ensemble de contextes locaux.

1. <https://www.w3.org/RDF/>

Nous avons développé un algorithme de détection de liens d'identité contextuels qui permet de déterminer pour chaque couple d'instances les contextes globaux les plus spécifiques dans lesquels ces instances peuvent être considérées comme étant identiques. Nous avons testé notre approche sur des données scientifiques issues d'un projet dans le domaine de l'agro-alimentaire.

Dans ce qui suit, nous présentons en section 2 un ensemble de travaux connexes. Puis, en section 3, nous fournissons les objectifs ainsi que les définitions utilisées dans notre approche. En section 4 nous présentons l'algorithme DECIDE qui calcule les liens d'identité contextuels. Enfin, nous présentons les premiers résultats d'expérimentation en section 5.

## 2 Travaux connexes

Les approches de découverte de liens d'identité permettent de détecter que deux descriptions se réfèrent au même objet du monde réel (e.g. même personne, même lieu, même gène). Avec l'initiative du Linked Open Data cloud (LOD) proposée par Tim-Berners Lee en 2007, un fort engouement a été constaté pour le développement d'approches de liage de données RDF, dans le domaine du Web sémantique (voir [4] pour un état de l'art). Pour représenter les liens d'identités générés par des approches (semi)-automatiques, le constructeur *owl:sameAs* défini dans le langage OWL2 [11] est utilisé, mais sa sémantique stricte exige que, si deux objets sont liés par un lien *owl:sameAs*, ces derniers doivent avoir les mêmes valeurs pour toutes les propriétés ( $((owl:sameAs(i_1, i_2) \wedge p(i_1, v) \Rightarrow p(i_2, v)))$ ). Certaines approches se sont focalisées sur la détection de liens erronés. Dans [2], les auteurs ont exploité l'hypothèse du nom unique pour détecter des liens d'identité erronés. Dans le même esprit, le travail de [3] propose une approche globale qui par analyse de la topologie du graphe des liens d'identité, détecte des liens d'identité invalides. Enfin, l'approche logique proposée dans [10] permet de détecter des liens d'identité invalides en exploitant un sous-graphe RDF construit en exploitant des axiomes de fonctionnalité et de complétude locale des propriétés.

D'autres approches ont caractérisé les différentes situations où le lien *owl:sameAs* serait utilisé à mauvais escient [5, 2, 3]. Halpin et al. [5] ont fait état de quatre cas où le lien *owl:sameAs* serait utilisé de façon inappropriée. Les auteurs citent en particulier le cas où un lien d'identité est déclaré entre deux objets dont les informations décrivent une entité dans deux contextes différents. En d'autres termes, il ne s'agit pas des mêmes informations qui sont renseignées dans tous les contextes (e.g. contexte social vs professionnel). Dans ce type de cas, il est nécessaire de pouvoir distinguer et représenter les contextes dans lesquels un lien d'identité serait valide. Il existe quelques propositions pour la représentation de liens d'identité faible tels que les prédicats SKOS [9] *skos:exactMatch*, *skos:closeMatch*, *skos:broadMatch* ou encore *skos:narrowMatch*. Ces prédicats ne permettent néanmoins pas de répondre au problème

de l'identité contextuelle. De plus, les prédicats SKOS ne peuvent être utilisés que pour des URI dont le type est un concept SKOS ce qui limite les cas d'utilisation possibles dans le LOD. Dans [5], les auteurs ont développé une ontologie de l'identité dans laquelle treize prédicats ont été hiérarchisés par la relation *rdfs:subPropertyOf* et caractérisés par les propriétés de réflexivité, de transitivité et de symétrie. Parmi ces prédicats, on trouve les prédicats SKOS cités précédemment mais également de nouveaux prédicats comme *id:claimsIdentical*, *id:matches* et *id:similar*. Le prédicat *owl:sameAs*, qui est réflexif, symétrique et transitif, est présenté au niveau le plus spécifique (i.e., le plus strict) et au niveau le plus général on trouve le prédicat *id:claimsSimilar* (réflexif, non-symétrique et non-transitif) et le prédicat *id:claimsRelated* (non-réflexif, non-symétrique et non-transitif). Les prédicats préfixés par le mot *claims* expriment une relation d'identité ou de similarité subjective dont la véracité dépend du contexte et/ou de l'interprétation du décideur humain. Bien que cette ontologie soit précise du point de vue structuration des relations d'identité et de similarité, elle ne permet néanmoins pas d'explicitier les contextes dans lesquels une relation d'identité serait valide. Dans le but d'utiliser des liens d'identité dans des contextes spécifiques, les auteurs de [6] ont proposé l'utilisation de graphes nommés pour représenter les contextes. Les différents contextes pertinents ne sont pas toujours faciles à expliciter, même pour un expert de domaine. Aussi, plus récemment, une approche développée par [1] permet de représenter l'ensemble des contextes possibles dans lesquels un lien d'identité pourrait être valide. Un contexte associé à un lien d'identité entre deux instances  $i_1$  et  $i_2$  correspond à un sous-ensemble de propriétés pour lesquelles  $i_1$  et  $i_2$  doivent avoir les mêmes valeurs. Cependant, dans cette approche, un contexte est un ensemble non structuré de propriétés qui ne tient pas compte de l'organisation multi-classes de données RDF associées à une ontologie. Par ailleurs, les auteurs ne présentent pas d'algorithme permettant de calculer les liens d'identité contextuels.

## 3 Une relation d'identité contextuelle guidée par l'ontologie et des connaissances expertes

L'objectif de notre approche est de découvrir des liens d'identité valides dans un contexte qui peut être défini comme une sous-partie d'une ontologie de domaine. Dans cette section, nous présentons tout d'abord le modèle de données sur lequel s'appuie notre approche. Puis, nous définissons les notions de contextes locaux et globaux utilisés pour représenter la relation d'identité proposée. Enfin, nous donnons la définition de la relation d'identité contextuelle que l'approche que nous avons développée permet de découvrir.

### 3.1 Modèle de données

Notre approche de détection de liens d'identité contextuelle s'appuie sur une base de connaissances où l'ontologie est représentée en OWL<sup>2</sup> et les données sont représentées en RDF. Une base de connaissances  $\mathcal{B}$  est définie par un couple  $(\mathcal{O}, \mathcal{F})$  où :

–  $\mathcal{O} = (\mathcal{C}, \mathcal{DP}, \mathcal{OP}, \mathcal{A})$  représente la partie conceptuelle de la base de connaissances définie par un ensemble de classes  $\mathcal{C}$ , un ensemble  $\mathcal{DP}$  de propriétés de type *owl:DataTypeProperty*, un ensemble  $\mathcal{OP}$  de propriétés de type *owl:ObjectProperty* et par un ensemble d'axiomes  $\mathcal{A}$  tels que la relation de subsomption entre classes, la disjonction entre classes ou encore la fonctionnalité des propriétés.

–  $\mathcal{F} = \{(s, p, o)\}$  est une collection de triplets (faits) de la forme (*sujet, propriété, objet*), exprimant des liens entre deux instances de classe ou une instance et une valeur littérale. On notera  $\mathcal{I}$  l'ensemble des instances de  $\mathcal{C}$  correspondant à l'ensemble des sujets  $s$  et des objets  $o$  tels que  $\exists(s, p, o) \in \mathcal{F}$ .

Une ontologie OWL peut être représentée par un graphe  $\mathcal{G} = (V, E)$  où l'ensemble de sommets  $V$  sont les classes et les types de données élémentaires (e.g. String, Date, Integer), et les arcs  $E$  sont les propriétés liant les classes entre elles ou liant les classes à des types de données élémentaires.

### 3.2 Présentation du problème de détection de liens d'identité contextuels

Le problème de détection de liens d'identité contextuels auquel nous nous intéressons peut être défini comme suit : étant donnée une base de connaissances  $\mathcal{B} = (\mathcal{O}, \mathcal{F})$  et l'ensemble  $\mathcal{I}^c$  des instances d'une classe cible  $c$  de l'ontologie  $\mathcal{O}$ , trouver pour l'ensemble des paires d'instances  $(i_1, i_2) \in (\mathcal{I}^c \times \mathcal{I}^c)$  les contextes les plus spécifiques pour lesquels  $(i_1, i_2)$  sont identiques.

Par exemple, on s'intéresse en Figure 1 à l'identité des individus de la classe cible *Jus* pour laquelle deux instances *jus1* et *jus2* sont représentées. L'un des contextes dans lequel les instances de jus de fruit *jus1* et *jus2* sont identiques est celui où l'on considère tous les produits composant le jus, et pour chaque produit une masse dont la description est réduite à l'unité de mesure, mais sans considérer leur quantité. Dans un deuxième contexte, la masse de Fraîse est considérée (avec sa valeur et son unité) et l'on observera la présence de banane mais on ne considérera pas la masse pour la banane.

Nous nous intéressons à la recherche de contextes communs correspondant à un sous-ensemble de classes et où pour chaque classe un sous-ensemble de propriétés de l'ontologie peut être pris en compte<sup>3</sup>. Enfin, certains contextes

2. <https://www.w3.org/OWL/>

3. Il ne s'agit pas ici de calculer le graphe le plus spécifique partagé par deux instances de Jus dans lequel les descriptions de classes pourraient varier selon les instances considérées (e.g. les propriétés prises en compte pour comparer les instances de la classe Masse ne doivent pas varier selon

sont clairement plus pertinents que d'autres (e.g. une valeur de masse sans unité n'a pas de sens). Nous considérons également une recherche de liens d'identités contextuels qui puisse être guidée par certaines connaissances expertes.

### 3.3 Contextes

L'approche de détection de liens d'identité contextuelle que nous avons développée utilise une notion de contexte global fondée sur la notion de contexte local qui peut être définie pour une classe quelconque de l'ontologie.

**Définition (Contexte Local).** Étant donnée une classe  $c$ , un contexte local peut être défini par  $\pi(c) = (c, DP^c, OP^c)$  où  $DP^c$  est un ensemble de propriétés de type *owl:DataTypeProperty* telles que  $c$  apparaît en domaine, ensemble auquel peuvent s'ajouter la propriété *rdf:type* ainsi que des propriétés d'annotations (e.g., *rdfs:label*).  $OP^c$  est un ensemble de propriétés de type *owl:ObjectProperty* telles que  $c$  apparaît en domaine ou en co-domaine (notée alors  $op^{-1}$ ).

Par exemple, un des contextes locaux de la classe *Masse* présenté en Figure 1 est :  $\pi(Masse) = (Masse, \{rdf:type, aValeur, aUnite\}, \{aMasse^{-1}\})$

**Définition (Relation d'ordre sur les contextes locaux).** Soient  $\pi_1(c)$  et  $\pi_2(c)$  deux contextes locaux pour la classe  $c$ . Le contexte  $\pi_1(c)$  est dit plus spécifique que  $\pi_2(c)$ , noté  $\pi_1(c) \leq \pi_2(c)$ , si et seulement si :  $DP_1^c \subseteq DP_2^c$  et  $OP_1^c \subseteq OP_2^c$ .

Par exemple, soient deux contextes locaux de la classe *Masse*  $\pi_1$  et  $\pi_2$  tels que :  $\pi_1(Masse) = (Masse, \{rdf:type, aValeur, aUnite\}, \{aMasse^{-1}\})$  et  $\pi_2(Masse) = (Masse, \{rdf:type, aValeur\}, \emptyset)$ , on a alors  $\pi_1(Masse) \leq \pi_2(Masse)$ .

L'ensemble des contextes locaux d'une classe  $c$  peut être représenté par un treillis de contextes locaux noté  $T(c)$  qui est composé de  $2^n$  contextes locaux, où  $n$  est le nombre de propriétés.

Pour réduire le nombre de contextes locaux à considérer (et donc le nombre de contextes globaux), nous prenons en compte des connaissances expertes, quand elles sont disponibles, concernant l'inutilité ou la nécessité de certaines propriétés ainsi que des informations sur l'importance de la cooccurrence de certaines d'entre elles. Plus précisément, pour chaque classe, un expert peut alimenter trois types de listes :

– *unwantedProps* : la liste des propriétés non pertinentes, qui n'interviendront pas dans le calcul des liens d'identité, soit parce qu'il s'agit d'une propriété non structurée (textuelle), ou parce que ses valeurs sont très hétérogènes ou encore parce que ses variations ne sont pas significatives compte tenu de l'utilisation des

que l'on compare la masse de la banane ou celle de la fraise)

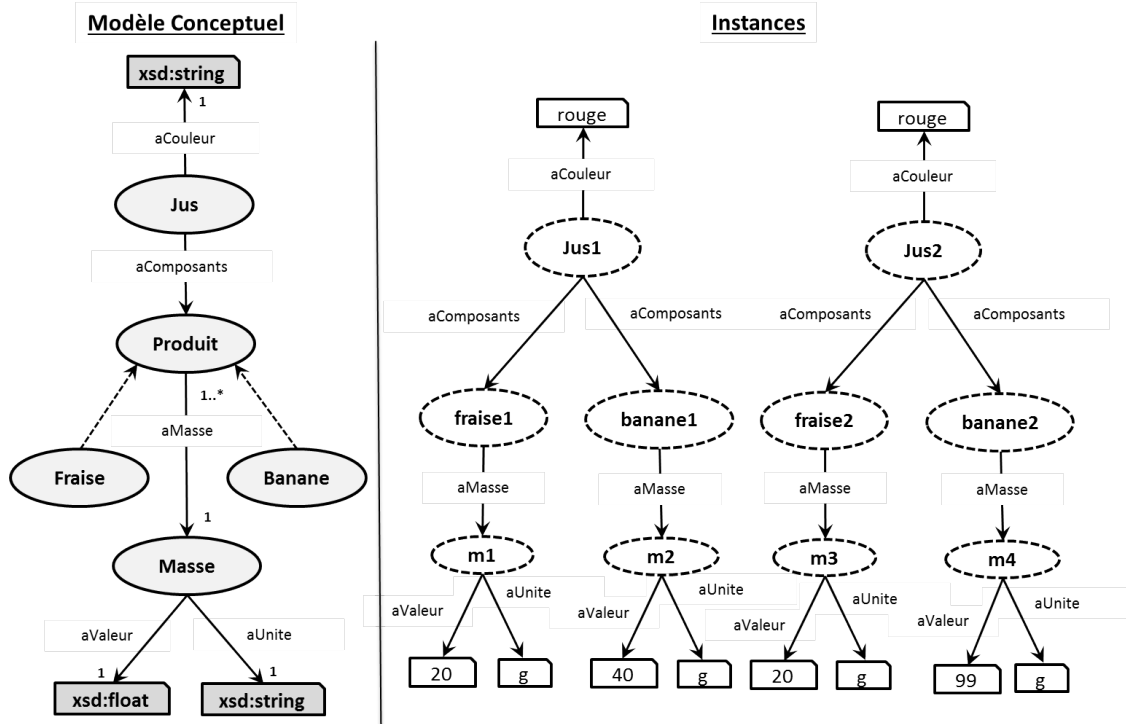


FIGURE 1 – Un extrait d’ontologie  $O$ , deux instances  $jus1$  et  $jus2$  de la classe cible  $Jus$ .

liens d’identité par l’expert (e.g. la couleur du jus de fruit). Ainsi, si l’on a ajouté la propriété  $rdf:type$  à  $unwantedProps$  avant la construction de  $T(Masse)$ , le contexte le plus spécifique sera  $\pi_k(Masse) = (Masse, \{aValeur, aUnite\}, \{aMasse^{-1}\})$  et  $T(Masse)$  contiendra seulement 8 contextes locaux.

–  $coProps$  : ensembles de propriétés devant être prises en compte ensemble, si elles sont considérées. Par exemple, une valeur de masse n’ayant pas de sens sans son unité de mesure, le couple  $\{aValeur, aUnite\}$  peut être ajouté par l’expert à  $coProps$ . Dans ce cas,  $T(Masse)$  se réduira à quatre contextes locaux possibles.

–  $necProps$  : ensemble des propriétés essentielles pour la comparaison des instances quel que soit le contexte considéré. Ainsi, si on ajoute  $\{rdf:type, aValeur, aUnite\}$  à  $necProps$ ,  $\pi_k(Masse) = (Masse, \{rdf:type, aValeur, aUnite\}, \{\emptyset\})$  sera le contexte le moins spécifique de  $T(Masse)$ .

**Définition (Contexte Global).** Etant donnée une classe cible  $cbl \in \mathcal{C}$ , un contexte global  $\Pi(cbl)$  est un sous-graphe connexe  $G$  de  $\mathcal{G}$  contenant  $cbl$ , tel que :  $\Pi(cbl) = \bigcup_{c_k \in C_G} \pi(c_k)$

Dans l’exemple présenté en figure 1, il existe différents contextes globaux dans lesquels les deux individus de la classe cible  $Jus$  sont identiques. On en citera deux :

–  $\Pi_a(Jus) = \{(Jus, \{rdf:type\}, \{aComposants\}), (Fraise, \{rdf:type\}, \{aComposants^{-1}\}), (Banane, \{rdf:type\}, \{aComposants^{-1}\})\}$

–  $\Pi_b(Jus) = \{(Jus, \{rdf:type, aCouleur\}, \{aComposants\}), (Fraise, \{rdf:type\}, \{aComposants^{-1}, aMasse\}), (Banane, \{rdf:type\}, \{aComposants^{-1}\}), (Masse, \{rdf:type, aValeur, aUnite\}, \emptyset)\}$

Les contextes globaux que nous considérons ne contiendront pas deux classes  $c_1$  et  $c_2$ , telles que  $c_1 \subseteq c_2$ . Cette contrainte permet d’éviter de construire des contextes globaux ne respectant pas les mécanismes d’héritage des propriétés. Pour cela nous sélectionnons les classes les plus générales pour lesquelles des instances ont été typées directement. Par exemple, s’il existe des instances directes de la classe  $Produit$ , les contextes globaux ne contiendront plus de contextes locaux définis pour les classes  $Fraise$  et  $Banane$  qui sont plus spécifiques. Les instances de ces classes seront représentées par le contexte local de la classe  $Produit$ . Le contexte global est alors plus abstrait et plus contraint puisqu’on ne pourra pas distinguer les contextes locaux de la classe  $Fraise$  et de la classe  $Banane$ , comme pour le contexte global  $\Pi_b$ .

**Définition (Relation d’ordre sur les contextes globaux).** Soient  $\Pi_1(cbl)$  et  $\Pi_2(cbl)$  deux contextes globaux pour la classe  $cbl$ . Le contexte global  $\Pi_1(cbl)$  est dit plus spécifique que  $\Pi_2(cbl)$ , noté  $\Pi_1(cbl) \leq \Pi_2(cbl)$ , si et seulement si :  $\forall \pi_i(c) \in \Pi_2(cbl), \exists \pi_j(c) \in \Pi_1(cbl)$  tel que  $\pi_j(c) \leq \pi_i(c)$ .

Dans l’exemple ci-dessus,  $\Pi_b(jus) \leq \Pi_a(jus)$ , puisque

$\pi_b(Jus) \leq \pi_a(Jus)$ ,  $\pi_b(Fraise) \leq \pi_a(Fraise)$  et  $\pi_b(Banane) \leq \pi_a(Banane)$ .

### 3.4 Relation d'identité contextuelle

Une relation d'identité contextuelle est valide dans un contexte global si toutes les valeurs littérales (à une équivalence près) et toutes les instances appartenant à ce contexte sont identiques. Cette relation peut être exprimée de la façon suivante :

**Définition (Relation d'identité contextuelle).** Soient  $(i_1, i_2)$  deux instances d'une classe cible  $cbl$ . Soit  $\Pi(cbl) = \{\pi_1(c_1), \dots, \pi_n(c_n)\}$  un contexte global pour la classe  $cbl$ . Deux instances  $(i_1, i_2)$  sont liées par une relation d'identité dans le contexte global  $\Pi(cbl)$ , notée  $identiConTo_{<\Pi(cbl)>}(i_1, i_2)$ , si et seulement si les sous-graphes RDF décrivant  $i_1$  et  $i_2$  et obtenus en utilisant la partie de l'ontologie représentée dans le contexte global, sont identiques, au renommage d'URI près et à une réécriture de valeurs littérales près.

Les relations d'identités ne seront représentées que pour les contextes les plus spécifiques et pourront être générées pour des contextes plus généraux si besoin. Soient  $\Pi_1(cbl)$  et  $\Pi_2(cbl)$  deux contextes globaux d'une classe cible  $cbl$ . Si  $\Pi_1(cbl) \leq \Pi_2(cbl)$  alors  $identiConTo_{<\Pi_1(cbl)>}(i_1, i_2) \Rightarrow identiConTo_{<\Pi_2(cbl)>}(i_1, i_2)$

## 4 DECIDE - Un algorithme de détection de liens d'identité contextuels

Le but de l'algorithme *DECIDE* (DEtection of Contextual IDENTITY) est de déterminer pour chaque couple d'instances  $(i_1, i_2) \in I \times I$  d'une classe cible  $cbl$  fixée par l'utilisateur, l'ensemble des contextes globaux les plus spécifiques pour lesquels la relation  $identiConTo$  est vraie. *DECIDE* prend en paramètres la base de connaissances  $\mathcal{B}$ , la classe cible  $cbl$ , et les trois listes de contraintes des experts *unwantedProps*, *coProps* et *necProps* si elles existent. La construction des contextes globaux se déroulent en trois étapes :

- Construction de la liste *Cdep* des classes les plus générales du graphe connexe maximal de  $cbl$  qui comportent des instances directement typées par ces classes.
- Pour chaque classe  $c \in Cdep$ , construction des treillis de contextes locaux  $T(c)$  pertinents en tenant compte des listes *unwantedProps*, *coProps* et *necProps* définies par les experts.
- Pour chaque couple d'instances  $(i_1, i_2)$  de  $cbl$ , appel de la fonction *IdentiConMax* détaillée dans l'algorithme 1 qui calcule l'ensemble des contextes globaux les plus spécifiques (*CGSet*).

Pour chaque couple d'instances  $(i_1, i_2)$ , *IdentiConMax* retourne l'ensemble des contextes globaux les plus spécifiques, dans lesquels les instances du couple  $(i_1, i_2)$  sont identiques. Cette fonction effectue un parcours en

profondeur d'abord des propriétés décrivant les instances à comparer et construit au fur et à mesure les contextes globaux les plus spécifiques.

---

#### Algorithme 1 : identiConToMax

---

##### Inputs :

- $cbl$  : la classe cible
- $Cdep$  : l'ensemble des classes du graphe connexe maximal de  $cbl$
- $(i_1, i_2)$  : une paire d'instances de la classe  $c$

**Output :** *CGSet* : l'ensemble des contextes globaux les plus spécifiques concernant le couple  $(i_1, i_2)$

$CGSet \leftarrow \emptyset$ ;  $\Pi(cbl) \leftarrow \emptyset$ ;  $fileCP.add(\Pi(cbl))$

**while** *fileCP* is not empty **do**

$\Pi(cbl)^{cour} \leftarrow fileCP.getNextElement()$   
 $dejaVu \leftarrow \emptyset$ ;  $propSrc \leftarrow nil$ ;  $cSrc \leftarrow nil$   
 $\Pi(cbl)^{cour} = compareCouple(i_1, i_2, dejaVu, propSrc, cSrc, \Pi(cbl)^{cour})$   
 $CGSet \leftarrow CGSet \cup \Pi(cbl)^{cour}$

---

Plus précisément, l'algorithme commence avec un contexte global courant vide  $\Pi(cbl)^{cour}$  qui sera enrichi par les contextes locaux au fur et à mesure de l'exploration. Pour ce contexte, on appelle la fonction *compareCouple* qui permet de comparer un couple d'instances d'une même classe  $c$ . Une file de contextes globaux partiels, notée *fileCP*, contient les contextes globaux alternatifs qui seront détectés lors de l'exploration et qui ne sont pas encore traités.

La fonction *compareCouple*, décrite dans l'algorithme 2, permet d'enrichir un contexte courant pour trouver un contexte plus spécifique dans lequel les instances du couple  $(i_1, i_2)$  sont identiques. Les données étant représentées sous la forme d'un graphe pouvant comporter des cycles, une liste de paires d'instances déjà explorées nommée *dejaVu* est maintenue afin d'éviter de recalculer les contextes identiques pour ces paires. La fonction *comparerCouple* prend en paramètre la liste *dejaVu*, la propriété objet source *propSrc* qui a permis d'atteindre le couple  $(i_1, i_2)$  à comparer, la classe source *cSrc* et le contexte global en cours  $\Pi(cbl)^{cour}$  (*dejaVu*, *propSrc* et *cSrc* seront réinitialisés à nil pour chaque nouveau contexte courant exploré). Cette fonction détermine les propriétés de type *DataTypeProperty* et les propriétés d'annotation dont les valeurs sont identiques (noté *DPeg*)<sup>4</sup>. Elle détermine également l'ensemble des propriétés objets *OPcom* instanciées pour  $i_1$ , et  $i_2$ . La fonction *possibleContext* va ensuite rechercher dans le treillis  $T(c)$  si ce contexte local  $\pi(c)^{cour}$  est possible (compte tenu des connaissances expert). S'il n'existe pas un tel contexte, le contexte local de  $T(c)$  le plus spécifique généralisant  $\pi(c)^{cour}$  est retourné.

4. Une mesure de similarité peut être utilisée et différer suivant chaque propriété

Si ce contexte retourné est le contexte vide, la fonction s'arrête et retourne le contexte global en cours  $\Pi(cbl)^{cour}$ . Si le contexte retourné de la fonction *possibleContext* n'est pas vide, 4 cas sont possibles :

1 – Il n'existe pas de contexte local déjà défini pour  $c$  dans  $\Pi(cbl)^{cour}$ , on ajoute alors  $\pi(c)^{cour}$  à  $\Pi(cbl)^{cour}$ . Ensuite, la fonction *compareObjProperties*, est appelée pour comparer tous les ressources reliées au couple  $(i_1, i_2)$ . Par exemple, si  $i_1=Jus1$  et  $i_2=Jus2$ , la fonction *compareObjProperties* va comparer les couples  $(fraise1, fraise2)$  puis  $(banane1, banane2)$ .

2 – Il existe un contexte local identique  $\pi(c)^{existent}$  déjà défini pour  $c$ . On appelle alors directement *compareObjProperties*.

3 – Il existe un contexte local pour  $c$  dans  $\Pi(cbl)^{cour}$  qui est plus général que  $\pi(c)^{cour}$ . Dans ce cas, on crée un nouveau contexte global  $\Pi(cbl)^{nouv}$  contenant le contexte local  $\pi(c)^{cour}$ , et on l'ajoute à la file *fileCP* qu'il restera à explorer. Ensuite, comme  $\pi(c)^{cour} \leq \pi(c)^{existent}$ , on doit vérifier si les instances ont bien les mêmes valeurs pour les *objectProperties* de  $\pi(c)^{existent}$  en appelant la fonction *compareObjProperties*.

4 – Dans tous les autres cas, on crée un nouveau contexte global  $\Pi(cbl)^{nouv}$  contenant le contexte local  $\pi(c)^{cour}$ , et on l'ajoute à la file *fileCP* pour qu'il soit exploré dans une prochaine itération de l'algorithme 1. Dans le contexte courant, comme on est sûr que le couple n'est pas identique en considérant  $\pi(c)^{existent}$ , le contexte de la classe source  $\pi(src)^{existent}$  est remplacé par un nouveau contexte local  $\pi(src)^{nouv}$  ne contenant pas la propriété *propSrc* dans  $\Pi(cbl)^{cour}$ .

La fonction *compareObjProperties* appelée dans *identiConToMax* permet de comparer l'ensemble des instances reliées au couple  $(i_1, i_2)$  pour chacune des propriétés  $p$  du contexte local. Les listes des instances liées à  $i_1$  et  $i_2$  sont examinées classe par classe et les propriétés sont supposées être localement complètes pour chacune des classes (e.g. si l'on connaît certains fruits composants un jus, nous supposons qu'ils sont tous représentés). Pour chacune des classes, afin de limiter le nombre de paires à examiner en cas de multivaluation, on vérifie tout d'abord que le nombre de valeurs de  $p$  est identique, puis on utilise une heuristique, qui permet de ne former que les couples  $(i'_1, i'_2)$  qui partagent le plus de valeurs de *data properties*. Après avoir formé la liste des meilleures paires, la fonction rappelle la fonction *compareCouple* pour chacune de ces paires afin de trouver les contextes les plus spécifiques dans lesquels ces paires sont identiques.

L'algorithme *DECIDE* appliqué sur l'exemple de la figure 1 permet de trouver les contextes globaux les plus spécifiques suivants pour *Jus1* et *Jus2* :

$$\Pi_a(Jus) = \{(Jus, \{rdf : type, aCouleur\}, \{aComposants\}), (Fraise, \{rdf : type\}, \{aComposants^{-1}, aMasse\}),$$


---

**Algorithme 2 : compareCouple**


---

**Inputs :**

- $(i_1, i_2)$  : une paire d'instances
- *dejaVu* : la liste des couples d'instances déjà traités
- *propSrc* : l'object property source
- *cSrc* : la classe source (i.e., le domaine ou le co-domaine de *propSrc*)
- $\Pi(cbl)^{cour}$  : le contexte global courant de la classe cible.

**Output :**  $\Pi(cbl)^{cour}$  : le contexte global courant de la classe cible, mis à jour.

*dejaVu*  $\leftarrow$  *dejaVu*  $\cup (i_1, i_2)$ ;

$c \leftarrow$  la classe commune de  $i_1$  et  $i_2$  tel que  $c \in Cdep$ ;

*OPcom*  $\leftarrow$  les propriétés objets communes de  $i_1$  et  $i_2$ ;

*DPeg*  $\leftarrow$  les data properties, rdftype et propriétés d'annotation ayant mêmes valeurs pour  $i_1$  et  $i_2$ ;

$\pi(c)^{cour} \leftarrow possibleContext(c, DPeg, OPcom)$

**if**  $\pi(c)^{cour} \neq \emptyset$  **then**

**if** (Il n'existe pas de contexte local de  $c$  dans

$\Pi(cbl)^{cour}$ ) **then**

$\Pi(cbl)^{cour}.add(\pi(c)^{cour})$ ;

*compareObjProperties*( $i_1, i_2, c, dejaVu,$

*propSrc, cSrc,  $\Pi(cbl)^{cour}$* );

**else**

$\pi(c)^{existent} \leftarrow getLocalContext(c, \Pi(cbl)^{cour})$ ;

**if**  $\pi(c)^{existent} == \pi(c)^{cour}$  **then**

*compareObjProperties*( $i_1, i_2, c, dejaVu,$   
*propSrc, cSrc,  $\Pi(cbl)^{cour}$* );

**else**

$\Pi(cbl)^{nouv} \leftarrow \emptyset$ ;

$\Pi(cbl)^{nouv}.add(\pi(c)^{cour})$ ;

**if**  $\Pi(cbl)^{nouv}$  n'existe pas dans *fileCP* **then**

*fileCP*  $\leftarrow fileCP.add(\Pi(cbl)^{nouv})$ ;

**if**  $\pi(c)^{cour} \leq \pi(c)^{existent}$  **then**

*compareObjProperties*( $i_1, i_2, c,$   
*dejaVu, propSrc, cSrc,  $\Pi(cbl)^{cour}$* );

**else**

$\pi(cSrc) \leftarrow$

*getLocalContext*(*cSrc,  $\Pi(cbl)^{cour}$* );

$\pi(cSrc)^{nouv} \leftarrow possibleContext(cSrc,$

$\pi(cSrc).DPeg, \pi(cSrc).OPcom -$   
*propSrc*);

$\Pi(cbl)^{cour} \leftarrow replaceLocalContext( $cSrc, \Pi(cbl)^{cour}, \pi(cSrc)^{nouv}$ );$

**return**  $\Pi(cbl)^{cour}$ ;

---

TABLE 1 – Taille des Datasets

	<i>Echantillon</i>	<i>D. Réelles</i>
# Faits	24 458	1 269 624
# Classes	4 635	4 743
# Instances	876	272 724
# ObjectProperties	50	50
# DataProperties	11	11
# Annotation Properties	25	25

TABLE 2 – Résultats

	<i>Echantillon</i>	<i>D. Réelles</i>
# Instances (cible)	22	220
# Paires d'instances (cible)	231	24 090
# Classes (graphe)	128	306
# Instances (graphe)	606	90 384
# Contextes Globaux	8	30
# Liens d'identité	252	25 189
Temps d'exécution	9 s	372 s

(Banane, {rdf : type}, {aComposants<sup>-1</sup>}),  
 (Masse, {rdf : type, aValeur, aUnité}, {aMasse<sup>-1</sup>}) et  
 $\Pi_b(\text{Jus}) = \{(\text{Jus}, \{\text{rdf} : \text{type}, \text{aCouleur}\}, \{\text{aComposants}\})\}$ ,  
 (Fraise, {rdf : type}, {aComposants<sup>-1</sup>, aMasse}),  
 (Banane, {rdf : type}, {aComposants<sup>-1</sup>, aMasse}),  
 (Masse, {rdf : type, aUnité}, {aMasse<sup>-1</sup>})

## 5 Expérimentations

Notre approche a été évaluée sur des données scientifiques relatives au domaine de l'agro-alimentaire. Nous avons exploité la version 1.4 de l'ontologie  $PO^2$  [7] qui décrit des processus de transformations<sup>5</sup>. Dans le cadre du projet INRA CellExtraDry,  $PO^2$  a été enrichie par une partie d'Agrovoc et la version de  $PO^2$  a été peuplée par 10 processus de transformation de micro-organismes (levures) suivant 20 itinéraires où chaque itinéraire représente des séquences d'étapes de transformation (séchage, chauffage ...). Les données réelles ne pouvant être mises à disposition, pour des raisons de confidentialité, un échantillon de ces données, décrivant un seul processus de stabilisation et dont certaines valeurs ont été modifiées, est accessible avec le code en *Java* de l'algorithme et ses résultats sur <https://github.com/raadjoe/DECIDE>. La taille des données réelles du projet et celle de l'échantillon sont décrites dans la Table 1. Cette première expérimentation a eu pour but d'observer le nombre de contextes distincts et le nombre de liens d'identité contextuels pouvant être générés dans une ontologie décrivant des données scientifiques. Dans cette expérimentation, une égalité stricte des valeurs littérales a été utilisée (à une normalisation près). L'algorithme *DECIDE* a été appliqué pour chacun de ces datasets afin de découvrir les contextes les plus spécifiques pour lesquels les individus de la classe cible *Mixture* sont identiques. Une instance de cette classe représente un ensemble de produits sur lequel est appliqué une étape d'un processus de transformation. La taille de la classe cible

<sup>5</sup> L'ontologie core de  $PO^2$  est accessible sur <http://agroportal.lirmm.fr/ontologies/PO2>

pour chaque dataset, la taille du graphe connexe maximal (i.e. classes et individus pouvant être atteints depuis la classe cible), et les résultats de l'algorithme sont présentés dans la Table 1. L'algorithme a été exécuté sur une machine de 8GB de RAM, ayant un processeur Intel Core 4 × 2.6GHz (Windows 10).

Les 220 instances de la classe cible *Mixture* des données réelles, forment 24 090 couples d'individus candidats. 30 contextes globaux générés par *DECIDE* suffisent pour représenter les contextes les plus spécifiques dans lesquels ces couples sont identiques (8 pour l'échantillon). Le nombre de liens d'identité (25 189), plus élevé que le nombre de couples d'individus, montre que certains couples sont identiques dans plusieurs contextes globaux non comparables. La hiérarchie des 8 contextes globaux créés dans l'échantillon est représentée dans la Figure 2. Le nombre qui apparaît à droite de chaque contexte global dans cette figure représente le nombre de couples qui sont identiques dans ce contexte (après inférence). Le contexte  $\Pi(\text{mix})_1 = \{(Mixture, \{\text{rdf} : \text{type}\})\}$ , qui est le contexte global le moins spécifique de tous les contextes générés, est un contexte dans lequel tous les couples sont identiques (231 pour l'échantillon). Le contexte le plus spécifique qui a été généré ( $\Pi(\text{mix})_8$ ), contient 11 contextes locaux, et peut garantir une identité plus forte. Un couple de mixtures identiques dans ce contexte signifie qu'elles sont de même nature (e.g humides) et partagent les mêmes quantités et unités de mesure de leurs composants : eau, silicose, levure, glucose, et cystéine.

En ré-appliquant l'algorithme *DECIDE*, et en prenant compte cette fois d'une seule connaissance experte qui indique qu'une valeur de masse n'a pas de sens sans son unité de mesure et vice-versa, les contextes locaux de 17 contextes globaux parmi 30 des données réelles, et 5 parmi 8 contextes globaux de l'échantillon étaient affectés. Ce qui indique que l'ajout d'une seule connaissance au début de l'algorithme a permis d'améliorer la sémantiques d'environ 60% des contextes globaux générés, sans affectant le nombre des contextes pour les deux datasets.

## 6 Conclusion

Nous proposons une approche de détection de liens d'identité contextuels permettant de générer pour chaque couple instance d'une classe cible, l'ensemble des contextes les plus spécifiques dans lesquels ces instances sont identiques. Une première expérimentation a été réalisée pour des données réelles décrivant des expérimentations scientifiques sur la levure, menées à l'INRA. Cette expérimentation a montré que différents contextes à différents niveaux d'abstraction pouvaient être découverts en un temps raisonnable.

Nous souhaitons maintenant montrer les résultats obtenus aux experts de domaine pour connaître l'intérêt des différents contextes et recueillir le maximum de contraintes permettant d'élaguer l'espace de recherche. Cette collecte pourrait être facilitée par la création d'un outil permettant

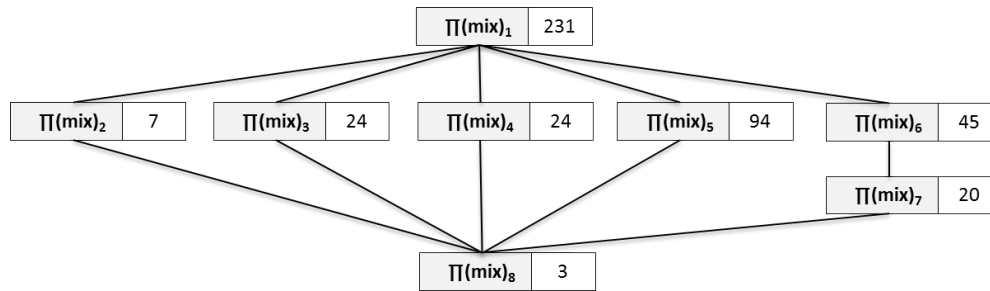


FIGURE 2 – La hiérarchie des 8 contextes globaux les plus spécifiques de la classe Mixture.

d’interagir avec les experts du domaine. Nous souhaitons ensuite utiliser ces liens pour faire prédire des valeurs manquantes à différents niveaux de confiance selon le niveau de spécificité du contexte d’identité. Il nous envisageons également de découvrir relations de causalité entre certains faits décrivant les expériences et ceux décrivant les résultats des observations en considérant l’identité contextuelle des autres facteurs expérimentaux.

## Remerciements

Ce travail est soutenu par le Center for Data Science, financé par IDEX Paris-Saclay, ANR-11-IDEX-0003-02.

## Références

- [1] W. Beek, S. Schlobach, and F. Harmelen. A contextualised semantics for owl : Sameas. In *Proceedings of the 13th International Conference on The Semantic Web. Latest Advances and New Domains - Volume 9678*, pages 405–419, New York, NY, USA, 2016. Springer-Verlag New York, Inc.
- [2] G. de Melo. Not quite the same : Identity constraints for the web of linked data. In M. desJardins and M. L. Littman, editors, *AAAI*. AAAI Press, 2013.
- [3] L. Ding, T. Finin, J. Shinavier, and D. L. McGuinness. owl :sameAs and linked data : An empirical study. In *In The Semantic Web - ISWC*, page 145–160, Apr. 2010.
- [4] A. Ferrara, A. Nikolov, and F. Scharffe. Data linking for the semantic web. *Int. J. Semantic Web Inf. Syst.*, 7(3):46–76, 2011.
- [5] H. Halpin, P. J. Hayes, J. P. McCusker, D. L. McGuinness, and H. S. Thompson. When owl : sameas isn’t the same : An analysis of identity in linked data. In *International Semantic Web Conference*, pages 305–320. Springer, 2010.
- [6] H. Halpin, P. J. Hayes, and H. S. Thompson. When owl : sameas isn’t the same redux : towards a theory of identity, context, and inference on the semantic web. In *International and Interdisciplinary Conference on Modeling and Using Context*, pages 47–60. Springer, 2015.
- [7] L. Ibanescu, J. Dibia, S. Dervaux, E. Guichard, and J. Raad. Po<sup>2</sup>-a process and observation ontology in food science. application to dairy gels. In *Metadata and Semantics Research : 10th International Conference, MTSR 2016, Göttingen, Germany, November 22-25, 2016, Proceedings*, pages 155–165. Springer, 2016.
- [8] A. Jaffri, H. Glaser, and I. Millard. Uri disambiguation in the context of linked data. In C. Bizer, T. Heath, K. Idehen, and T. Berners-Lee, editors, *Linked Data on the Web - LDOW*, volume 369 of *CEUR Workshop Proceedings*. CEUR-WS.org, 2008.
- [9] A. Miles and S. Bechhofer. Skos simple knowledge organization system reference. w3c recommendation 18 august 2009., 2009.
- [10] L. Papaleo, N. Pernelle, F. Saïs, and C. Dumont. Logical detection of invalid sameas statements in RDF data. In *Knowledge Engineering and Knowledge Management - 19th International Conference, EKAW 2014, Linköping, Sweden, November 24-28, 2014. Proceedings*, pages 373–384, 2014.
- [11] P. F. Patel-Schneider, P. Hayes, and I. Horrocks. OWL Web Ontology Language Semantics and Abstract Syntax Section 5. RDF-Compatible Model-Theoretic Semantics. Technical report, W3C, Dec. 2004.



## Communication multi-niveaux pour des IoT-a. Interactions autour d'un mur d'écrans connectés.

Alexandre SCHMITT    Florent CARLIER    Valérie RENAULT

Université du Maine  
Laboratoire Centre de Recherche en Éducation de Nantes (CREN)  
Avenue O. Messiaen, 72085 Le Mans, France  
prénom.nom@univ-lemans.fr

### Résumé

*A l'heure où beaucoup d'objets du quotidien deviennent connectés, où des masses considérables d'information sont diffusées sur les réseaux, des questions de méthodologie, de formalisation et de normalisation se posent. Comment pouvons-nous faire cohabiter, de façon cohérente, le monde des objets connectés (Internet of Things : IoT) et du Big Data, le monde des objets réels avec les masses d'informations immatérielles ? Comment visualiser et manipuler, de façon intuitive, ces masses de données hétérogènes issues de capteurs diffus dans notre environnement ? Nous proposons d'introduire de l'intelligence au plus près des composants électroniques de ces objets connectés. L'objectif est de considérer ces objets physiques comme des "Internet of Things - Agents", que nous nommons IoT-a. La plateforme Triskell3S a ainsi été expérimentée sur un ensemble hétérogène d'objets connectés formant un mur d'écrans interactif et modulable. Ce mur d'écrans est constitué de briques autonomes, de IoT-a, pouvant être composées d'un dispositif visuel et d'une carte embarquée. L'architecture sous-jacente repose sur l'adaptation de modèles multi-agents au plus près du hardware des objets physiques réels. Dans un souci d'interopérabilité, cette architecture se focalise sur l'adéquation de protocoles de communication : le protocole MQTT, dédié aux communications inter-objets et le protocole D-Bus, dédié aux échanges intra-objets. Ces deux protocoles sont ainsi intégrés à notre architecture tout en respectant la norme FIPA-ACL, définie pour les systèmes multi-agents.*

### Mots Clef

Objets connectés, Internet of Things, Systèmes multi-agents embarqués, Mur d'écrans, Table surface, Systèmes embarqués.

### Abstract

*When several day-to-day objects become connected, or a considerable amount of information is released throughout networks, one can wonder about matters of methodologies, formalization and normalization. How can one coherently make IoT and Big Data live together, this world of ac-*

*tual objects and tons of immaterial information ? How can one visualize and manipulate, intuitively, this important amount of heterogenous data sent by sensors spread in our environment ? We propose the introduction of intelligence as close to those connected appliances as can be. The main goal is to consider those physical objects as Internet of Things – Agents, which we name IoT-a. The Triskell3S platform was thus tested out on a heterogenous system of connected objects composing a multi-screen wall display which is interactive and adjustable. This multi-screen wall display is composed of autonomous bricks, which are IoT-a and which can be composed of a visual device and an embedded system. The resulting architecture rests on the adaptation of multi-agents models located as close as can be to the real physical objects. As a means of allowing interoperability, this architecture focuses on matching with communication protocols : protocol MQTT, dedicated to inter-objects exchange, and protocol D-Bus, dedicated to intra-objects exchanges. These two protocols are thus integrated to our architecture in full respect of the FIPA-ACL norm, defined for multi-agents systems.*

### Keywords

Internet of Things, Multi-Agents Systems, Wall display, Surface table, Embedded system.

## 1 Introduction

Le domaine des objets connectés (Internet of Things / IoT) connaît un essor important et est de plus en plus présent dans notre environnement. De nombreuses entreprises souhaitent disposer de plus amples informations concernant leurs activités afin de chiffrer et mieux évaluer leur rendement. Les objets connectés leur apportent une solution concernant cette problématique. Les IoT sont des Systèmes Embarqués (SE) légers en interaction avec l'environnement réel. Leur nombre croît à une vitesse importante et ces objets feront partie intégrante de notre vie quotidienne dans quelques années. Les entreprises se tournent aujourd'hui vers ces solutions pour avoir un retour d'informations précis de leurs activités ou de leurs environnements de travail.

Dans ce contexte, nous pensons que l'intégration de Systèmes Multi-Agents (SMA) dans les IoT doit permettre de traiter des données hétérogènes, issues d'un environnement complexe et distribué. Nous rajoutons ainsi de l'intelligence au sein même de ces objets connectés. En intégrant l'intelligence au plus près des composants électroniques, nous nommons IoT-a les objets physiques du type "Internet of Things - agents". Ainsi l'intelligence reste distribuée dans les objets et non centralisée et déportée dans le "cloud" [5].

Dans cet article, nous introduisons le contexte industriel dans lequel est menée cette recherche afin d'aboutir au positionnement des problématiques. Nous proposons ensuite une revue détaillée des différents paradigmes développés dans le domaine des IoT et les liens établis dans la littérature entre les IoT et les systèmes multi-agents. Une présentation des applications envisagées sera détaillée dans la section 5. En particulier, nous montrerons comment les paradigmes et normes des domaines agents et des IoT peuvent cohabiter en mettant en oeuvre les protocoles MQTT et D-Bus tout en respectant les standards FIPA-ACL. Enfin, nous discuterons des différentes perspectives à envisager autour de ce travail selon qu'il est abordé d'un point de vue IoT-a, données ou usages.

## 2 Interactions de données hétérogènes via des supports interactifs

L'entreprise SARP souhaite bénéficier d'un nouvel outil pour ses équipes d'organisation. Chaque agence est composée d'un ou plusieurs centres opérationnels. Ces centres organisent des chantiers au quotidien. Ils sont constitués de deux à trois personnes appelées assistants d'exploitation. Un chantier nécessite la manipulation de trois types de données contraintes : les ressources humaines, matérielles et la planification de chantiers.

Les ressources humaines concernent les équipes de chantiers. Elles sont composées d'opérateurs ayant des qualifications propres en rapport avec différents types de travaux (travail en hauteur, permis poids lourds, amiante, etc.). Les véhicules se composent de matériel spécifique et ne sont compatibles que pour certains chantiers et opérateurs.

Concernant la planification, des aléas à différentes échelles (la détérioration de matériel, l'absence de personnels qualifiés, l'accès impossible à un chantier dû au retard d'un autre corps de métier) perturbent les plannings établis. Un assistant d'exploitation reçoit en continue des demandes client. Il planifie avec eux les interventions puis élabore les équipes d'opérateurs et de matériels compatibles pour les chantiers tout en ayant en tête les aléas pouvant perturber l'intervention.

Lors de cette organisation, il est important que les équipes de gestion aient un oeil sur la rentabilité des activités. Un outil interne calcule des indicateurs cohérents à l'activité quotidienne. Il est donc nécessaire de bénéficier d'un affichage interactif prenant en compte toutes ces données pour un meilleur contrôle des coûts et de la rentabilité.

Une large surface d'affichage et un espace collaboratif de manipulation des données, permettra aux opérateurs du centre de contrôle de visualiser et d'interagir de façon collective avec toutes ces données. Les assistants d'exploitation pourront ainsi visualiser différentes échelles et points de vue de ces données hétérogènes et distribuées. Plusieurs possibilités d'agrégats d'informations sont envisageables. Comment visualiser simplement l'antériorité des chantiers : leurs localisation, retour d'informations antérieures, spécificités techniques et organisationnelles pour en ressortir une information utile à l'assistant d'exploitation durant une organisation de journée ? Comment connaître les ressources matérielles et humaines disponibles et en capacité d'intervenir, sur un secteur d'activité, pour répondre à une urgence client plus efficacement ?

L'ensemble du projet reposera sur l'utilisation d'une plateforme multi-agents. Cette dernière sera la clé de voûte des échanges entre les différents éléments hétérogènes portant les informations représentatives des activités de l'entreprise. Ce système se voudra inter-opérable quel que soit le type de matériel utilisé (PC, Tablette, Smartphone, Table surface, etc.)

## 3 Problématiques

Nos travaux reposent sur deux problématiques principales conjointes.

La première concerne les aspects matériel et logiciel dans le contexte des SMA appliqués aux objets connectés. Par exemple, comment intégrer les notions de SMA au sein d'objets connectés, pour en optimiser la communication et les échanges ?

La seconde prend en compte les contraintes de données complexes. Comment manipuler et visualiser à différentes échelles des données hétérogènes et distribuées reposant sur des IoT-a ?

Ces recherches devront être directement expérimentées et appliquées afin de réaliser un réseau de communication entre différentes données réparties sur un territoire géographique et un centre de contrôle. L'ensemble du système sera visualisé sur un mur d'écrans interactifs manipulable via une table de surface.

## 4 Etat de l'art

Différents travaux ont montré l'intérêt des SMA dans les systèmes embarqués [12]. Des plateformes spécifiques telles que Spade [10] ou JADE [2] permettent d'expérimenter des modèles SMA. JADE, par exemple, "permet aux développeurs d'implémenter et de déployer des systèmes multi-agents, y compris des agents tournant sur des réseaux sans-fil ou sur des dispositifs ayant des ressources limitées" [3]. Le plus souvent ces solutions adoptent l'utilisation d'une couche d'abstraction logicielle [1]. Elle permet de s'affranchir de l'aspect matériel des systèmes sur lesquels ils peuvent être exécutés. En l'occurrence JADE repose sur le langage JAVA utilisant un moteur virtuel pour le rendre interopérable.

Tableau 1 – Comparatif des protocoles de communication

	AMQP	HTTP	MQTT	CoAP
Taille de messages	<b>Faible</b> Flux d'octets adaptés à l'information. Type + Objet	<b>Importante</b> Entête volumineuse Informations sous forme de messages texte	<b>Très faible</b> Flux d'octets adaptés à l'information. Information brute	<b>Moyenne</b> Informations sous forme de messages texte
Qualité de service	Même pour tous les messages En développement	Même pour tous les messages	Trois types de qualité de service	3 types de confirmations d'envoi
Type de distribution de messages	1 vers 1 et 1 vers N	1 vers 1	1 vers 0, 1 vers 1 et 1 vers N	1 vers 1 et 1 vers N
Architecture	Publish/Subscribe and Request/Response	Request/Response	Publish/Subscribe	Request/Response and Publish/Subscribe
Bridge	non	non	oui	non
Contexte d'usage	Bancaire	Page WEB	Objets connectés	Objets connectés

Ces solutions ne sont pas optimales pour un contexte embarqué.

Notre approche repose sur l'intégration d'agents au plus près de la couche matérielle des IoT [11].

En effet, un IoT est d'abord un système embarqué intégrant du matériel et un logiciel adapté à son activité. Ce matériel dispose très souvent de ressources (CPU, mémoire, énergie) assez limitées. Ce choix permet d'avoir une consommation la plus faible possible. Le logiciel embarqué est conçu pour répondre au mieux aux exigences du matériel. Il est alors intéressant de proposer des solutions, apportant un minimum d'intelligence, adaptées à une intégration matérielle exigeante telle que celle des IoT.

Tout IoT repose sur un système embarqué plus ou moins performant en fonction de la nature des informations dont il dispose. Il devient donc possible de rendre connecté tout ce qui est mesurable et/ou interopérable. Cette diversité et hétérogénéité des objets laisse apparaître un problème de taille concernant l'analyse cohérente des informations collectées.

Notre postulat est donc d'intégrer, au sein même des IoT, une couche agent autorisant une communication entre les différents objets réseaux. Nous proposons le terme de IoT-a comme acronyme de "Internet of Things-Agents" [4], pour représenter ces entités autonomes diffusées dans l'environnement et interconnectées. Dans cette vision, chaque IoT est un agent ayant des caractéristiques communes avec tous les autres IoT. Cet IoT-a a, entre autre, la capacité d'interagir avec d'autres agents IoT-a et il peut ainsi être doté de nouveaux comportements hétérogènes en fonction de son rôle dans l'environnement. Ces comportements le différencient alors d'un simple "capteur connecté". Nos travaux vont se centrer ainsi sur la communication entre IoT-a afin de mettre en adéquation les standards agent et les standards du domaine des IoT.

De nombreux réseaux industriels permettent aujourd'hui d'interconnecter des IoT : les réseaux SigFox ou LoRa [7] sont des supports d'informations fiables à faible coût, pour des échanges légers. Afin d'organiser l'information de tous ces objets connectés différents protocoles ouverts sont disponibles [9]. Le tableau 1 dresse un comparatif non exhaustif de quatre des principaux protocoles d'échanges d'informations et notamment les protocoles MQTT et CoAP adaptés aux communications pour objets connectés. CoAP est très proche du protocole HTTP et intègre d'ailleurs de nombreuses fonctionnalités. Il a été conçu pour bénéficier d'une consommation plus faible par échange par rapport au HTML. CoAP repose néanmoins sur une logique d'échange de message texte ce qui peut s'avérer particulièrement lourd lors d'échange de données numérique et peut être adapté aux contraintes de notre projet. Il est montré dans [18] que MQTT apporte moins de délais de transmission et moins de paquets perdus pour une grande quantité d'échange d'information. Le protocole MQTT ressort comme étant le plus adapté à notre contexte de recherche. Ce protocole est ouvert "Machine-to-Machine", il permet des échanges de données bas niveau portés par un réseau TCP/IP. Il est encapsulé dans des trames TCP et dispose de plusieurs types de qualité de service. Il s'avère à la fois très léger et particulièrement robuste. Ces caractéristiques sont nécessaires car la communication avec les IoT est bien souvent effectuée sur des canaux de transmission assez bruités, car éloignés des bases, et à faible bande passante pour bénéficier d'une consommation la plus faible possible.

Afin d'optimiser la communication des IoT-a, il sera aussi nécessaire de définir les messages échangés et donc l'ontologie spécifique adaptée aux interactions inter-agents et agent-humains [13].

Notre expérimentation portera sur un mur d'écrans [17] permettant la visualisation des données métiers et sur son interaction avec une table surface permettant la manipulation collaborative de ces données.

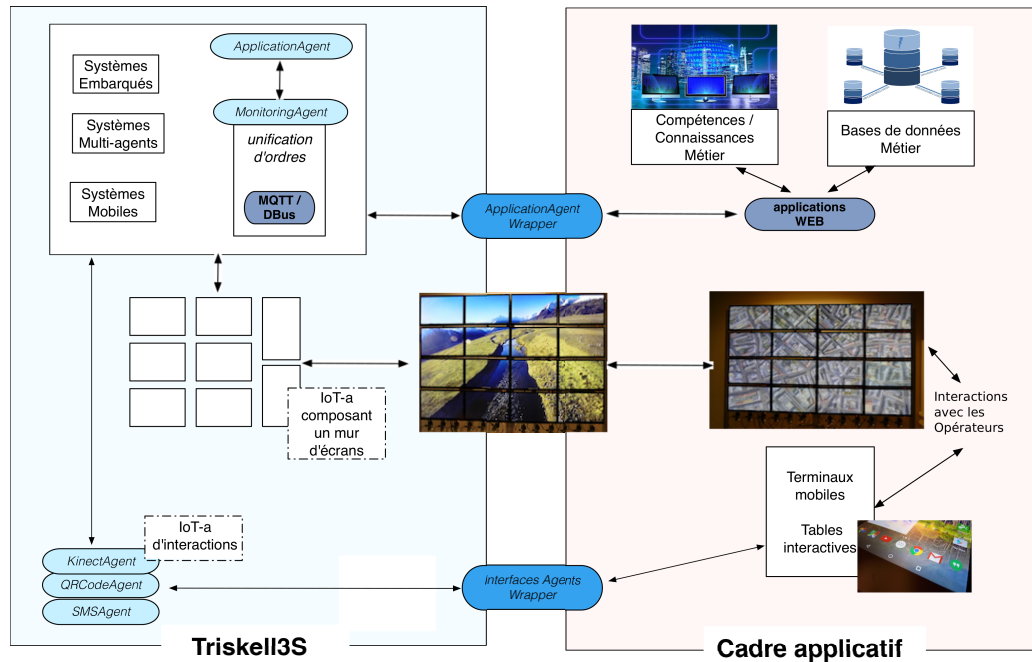


FIGURE 1 – Plateforme Triskell3S et ses interactions

Chaque brique écran du mur, ainsi que la table surface, étant considérée comme un IoT-a comme résumé sur la figure 1.

Dans notre cas l'ensemble de systèmes embarqués IoT-a est un mur écrans couplé d'une table surface tactile. Les interactions entre IoT-a sont au centre de ce sujet de recherche. Comment les optimiser pour les rendre compatible avec les interactions qu'impose la norme FIPA ? Quelles sont les redondances de communication à éviter pour un meilleur gain en performance ?

## 5 Communication et interactions multi-niveaux

### 5.1 Communication entre agents

Afin de faire évoluer l'ensemble des IoT-a de notre système dans un environnement multi-agents l'implémentation d'une plateforme dédiée aux systèmes embarqués a été nécessaire. Les développements ont été effectués de façon à respecter le standard FIPA. La plateforme Triskell3S [4] développée en C++ fait cohabiter les différents agents s'occupant de la gestion de contenus vidéo à afficher sur le mur d'écrans [5]. Les communications inter-agents/intracarte se font via le protocole D-Bus permettant des accès bas niveau et à haute vitesse avec le système. Il est possible d'interagir entre différents agents situés à des niveaux différents (de la couche matérielle bas niveau vers la couche logicielle utilisateur). Les communications inter-carte sont effectuées via le réseau ethernet avec le protocole MQTT. Avant de spécifier l'ontologie, la première étape de nos recherches repose sur l'optimisation de l'architecture des

protocoles des messages entre IoT-a.

Le MQTT repose sur le mécanisme de *publisher/subscriber* [6]. Un client peut souscrire à un fil de messages nommé *topic* et/ou publier sur un fil particulier. Cette architecture permet des communications synchrones de 1 vers 1 ou 1 vers N clients ou encore asynchrone de 1 vers 0 hôte (il est possible de publier un message sur un topic souscrits par aucun client). Un serveur d'échange de messages MQTT est appelé *broker*. Plusieurs projets proposent des *brokers*, comme par exemple, IBM WebSphere MQ, HiveMQ, Mosquitto, etc. Nous utilisons dans le cadre de nos recherches, le projet open source Mosquitto. Il propose un *broker* et une librairie pour la création de clients légers développés en langage C. Cette solution est la plus pertinente pour notre environnement embarqué. La version actuelle du *broker* Mosquitto permet un code exécutable de 120Ko consommant 3Mo de mémoire RAM pour 1000 clients connectés [16]. Son fonctionnement sur plateformes x86 et ARM (architecture de nos systèmes embarqués) a été validé.

Le MQTT présente de nombreux avantages parmi lesquels la notion de *bridge*. Cette option permet une interconnexion privilégiée entre les *brokers*. Plusieurs systèmes disposant chacun d'un *broker*, peuvent se partager mutuellement et simultanément leurs informations via un *bridge*. Cette solution ouvre de nombreuses possibilités et nous offre principalement le moyen de communiquer de façon décentralisée sur tout élément agent de notre réseau.

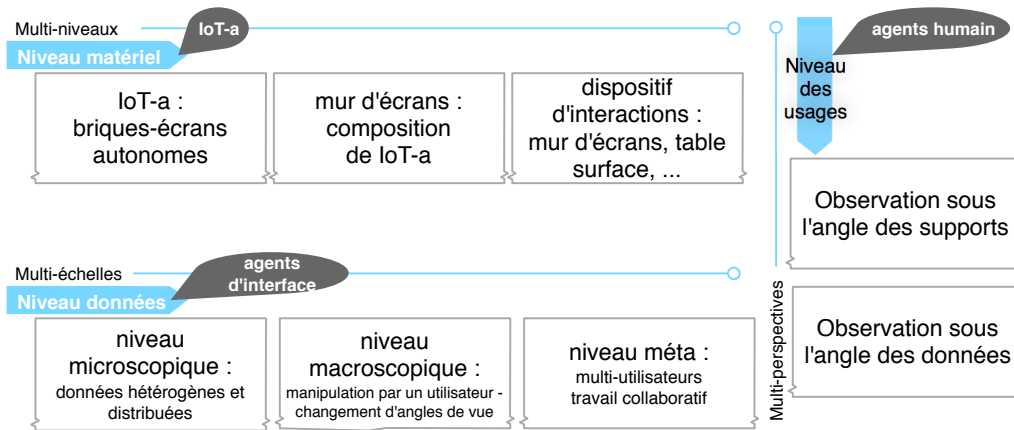


FIGURE 2 – Niveaux croisés de compétences

Différentes expérimentations devront maintenant être menées afin de mesurer les avantages et les limites de cette solution. Ces mécanismes de communication, via les bridges, permettent d'envisager des interactions multi-niveaux selon les besoins d'affichage et d'interactions des opérateurs humains.

## 5.2 Multi-échelles, multi-niveaux et multi-perspectives

Comme résumé dans la figure 2, ces travaux peuvent être abordés sous différentes perspectives selon que l'on s'intéresse aux aspects matériels, aux données ou aux usages. L'intégration des IoT-a au niveau des briques-écrans, du mur d'écrans et des différents outils d'interactions mis en oeuvre pour manipuler les données implique une analyse *multi-niveaux* de la plateforme Triskell3S et de la gestion de modèle agent de communication. Les agents présents au sein de la plateforme Triskell3S peuvent ainsi être appréhendés à des niveaux d'abstractions différents et regroupés en fonction de leurs capacités. Une vue microscopique de l'environnement permet ainsi de comprendre les interactions au sein d'une même carte embarquée. La vue macroscopique, quant à elle, permet d'appréhender les interactions entre des agents de plus haut niveau exécutant des services plus complexes.

Du point de vue des données, la granularité des informations métier transitant dans notre système suggère une approche *multi-échelles*. Nos travaux se rapprochent alors de ceux de C. Duchêne [8] pour son approche multi-échelles appliquée à l'agrégation d'agents dans le domaine de la cartographie. Son objectif est d'agréger un jeu de données géographiques vectorielles, dans un but de généralisation cartographique. Ces études montrent la difficulté du passage à l'échelle et le rôle organisationnel que peuvent avoir les agents dans ce contexte. Dans notre cas, la distribution des données métier et les différents niveaux de compétences des utilisateurs nous permettront de confronter cette approche multi-échelles dans un environnement embarqué

contraint en ressources de calcul. Cette approche pourra donc avoir des conséquences sur les protocoles de communication sous-jacents entre nos IoT-a.

Le troisième axe de la figure 2 concerne les usagers du dispositif, et en particulier les opérateurs qui devront manipuler de façon collaborative les données métier. Là encore, plusieurs perspectives sont à co-construire selon que les travaux se placent sur les moyens matériels d'interactions supportés par le dispositif ou sur les données à manipuler. Ces différentes granularités dans les approches seront abordées dans les travaux avec une priorité sur l'analyse multi-niveaux des agents et sur leur processus de communication. Par exemple, comment interagir de façon cohérente à des niveaux différents, pour répondre à des problèmes globaux au niveau des interfaces ? Est-il pertinent de regrouper certains domaines de compétences des agents pour faciliter la gestion de contenus hétérogènes ?

Différents travaux ont été menés mettant en perspective le rôle des systèmes multi-agents dans des environnements interactifs tels que des tables interactives [13]. Ces travaux soulignent que la conception de dispositifs interactifs complexes doivent être menés de front avec la conception de la plateforme sous-jacente.

## 6 Conclusion et Perspectives

L'objectif de cette recherche est d'étudier et d'optimiser ces protocoles de communication multi-niveaux dans le cadre des IoT-a, tout en tenant compte des problèmes de sécurité inhérents aux IoT. Il sera nécessaire de tester la réactivité et l'efficacité de ces protocoles lors d'interactions plus importantes avec les usagers. Cette étape devrait permettre de développer de nouveaux moyens d'interaction et d'usage du mur d'écrans en proposant de nouveaux IoT-a d'interactions complémentaires aux agents d'IHM.

Cette étape devrait aussi conduire à s'interroger sur la typologie et l'optimisation des messages à mettre en place au travers des protocoles prédéfinis MQTT et D-Bus.

Kubicki et al. [14] présentent un état des lieux des expériences utilisateurs et des critères d'évaluation des tables interactives, notamment dans des contextes de tâches collaboratives. La multi-modalité et les caractères multi-utilisateurs peuvent modifier la collaboration traditionnelle des opérateurs. Des agents d'interface devront donc être introduits progressivement dans le système afin d'aider à l'acceptation du dispositif. Des travaux complémentaires dans le domaine des Interfaces Homme-Machine seront donc nécessaires. Ils permettront aussi d'étendre et d'expérimenter le dispositif mur d'écrans et table de surface à d'autres contextes [15], tels que le contexte éducatif.

## Remerciements

Les travaux menés dans cet article bénéficient d'un financement de thèse CIFRE en partenariat entre le laboratoire Centre de Recherche en Éducation de Nantes (CREN) et l'entreprise SARP - SOA Le Mans avec la collaboration de Frédéric Le Bouguenec, Directeur Technique et Innovation.

## Bibliographie

- [1] Luigi Atzori, Antonio Iera, and Giacomo Morabito. The internet of things : A survey. *Computer Networks*, 54(15) :2787 – 2805, 2010.
- [2] Fabio Luigi Bellifemine, Giovanni Caire, and Dominic Greenwood. *Developing Multi-Agent Systems with JADE (Wiley Series in Agent Technology)*. John Wiley & Sons, 2007.
- [3] Federico Bergenti, Giovanni Caire, and Danilo Gotta. Agents on the move : Jade for android devices. In *Proceedings of the XV Workshop "Dagli Oggetti agli Agenti" (WOA 2014)*, volume 1260 of *CEUR-WS*, pages 1613–073. Corrado Santoro (ed.), 2014.
- [4] F. Carlier and V. Renault. Iot-a, embedded agents for smart internet of things : Application on a display wall. In *2016 IEEE/WIC/ACM International Conference on Web Intelligence, The First International Workshop on the Internet of Agents (IoA)*, pages 80–83, Omaha, Nebraska, USA, 13-16 Oct. 2016 2016. IEEE Computer Society.
- [5] F. Carlier and V. Renault. Systeme d'affichage multi-agents et methode d'affichage associee. Brevet, Notice, 10 Mars 2017.
- [6] MQTT community. MQTT homepage. <http://mqtt.org/>. Page d'accueil organisation mqtt.
- [7] Walid Dabbous. Quels réseaux pour les Big Data ? In Mokrane Bouzeghoub and Rémy Mosseri, editors, *Les Big Data à découvert*. CNRS Editions, 2017.
- [8] Cécile Duchêne. *Automatisation de la généralisation cartographique - Relations et interactions, orchestration et approches multi-agents*. Habilitation à diriger des recherches, spécialité sciences et technologies de l'information géographique, Université Paris-Est, jun 2016.
- [9] Simon Duquennoy, Gilles Grimaud, and Jean-Jacques Vandewalle. The Web of Things : interconnecting devices with high usability and performance. In *ICESS 2009*, HangZhou, Chile, May 2009.
- [10] Miguel Escrivá Gregori, Javier Palanca Cámara, and Gustavo Aranda Bada. A jabber-based multi-agent system platform. In *Proceedings of the Fifth International Joint Conference on Autonomous Agents and Multiagent Systems, AAMAS '06*, pages 1282–1284, New York, NY, USA, 2006. ACM.
- [11] Jean-Paul Jamont and Michel Ocelllo. Designing embedded collective systems : The DIAMOND multiagent method. In *IEEE International Conference on Tools with Artificial Intelligence - ICTAI 07*, pages 91–94, Patras, Greece, 2007. IEEE Computer Society.
- [12] Jean-Paul Jamont and Michel Ocelllo. A multiagent method to design open embedded complex systems. In P. Fritzsche, editor, *Tools in Artificial Intelligence*, pages 205–222. In-Teh, 2008.
- [13] Céline Joiron, Frédéric Fürst, Gilles Kassel, Alistair Jones, Jean-Paul Barthès, Claude Moulin, and Dominique Lenne. An Ontology for Designing a Collaborative Platform Involving a Multi-touch and Multi-user Interactive Table. In Jan Dietz, editor, *KEOD*, pages 1–12, Vilamoura, Portugal, September 2013.
- [14] Sébastien Kubicki, Katarzyna Borgiel, Sophie Lepreux, Marion Wolff, and Christophe Kolski. Réflexions autour des tables interactives : expérience utilisateur, utilisabilité, évaluation. *Le travail humain*, pages 229–252, December 2012.
- [15] Sébastien Kubicki, Denis Pasco, and Ingrid Arnaud. Utilisation en classe d'un jeu sérieux sur table interactive avec objets tangibles pour favoriser l'activité des élèves : une évaluation comparative en cours préparatoire. *Sciences et Technologies de l'Information et de la Communication pour l'Éducation et la Formation (STICEF)*, 21 :1, October 2014.
- [16] Roger Light. Project mosquito. <https://www.mosquitto.org>. Page d'accueil du projet Mosquitto.
- [17] V. Renault, F. Carlier, and A. Schmitt. Framework sma pour visualisation multi-écrans. In *Journée Interaction Homme-Machine et Intelligence Artificielle*, Université Pierre et Marie-Curie, Paris, France, 17 Mars 2017.
- [18] D. Thangavel, X. Ma, A. Valera, H. X. Tan, and C. K. Y. Tan. Performance evaluation of MQTT and CoAP via a common middleware. In *2014 IEEE Ninth International Conference on Intelligent Sensors, Sensor Networks and Information Processing (ISSNIP)*, pages 1–6, April 2014.

# Formalisation et résolution d'un problème en Santé Animale avec le cadre Leader-Follower MDP

A.-F. Viet<sup>1</sup>C. Belloc<sup>1</sup>R. Sabbadin<sup>2</sup><sup>1</sup> BioEpAR, INRA, Oniris, Nantes<sup>2</sup> MIAT, INRA, Toulouse

anne-france.viet@oniris-nantes.fr ; catherine.belloc@oniris-nantes.fr ; regis.sabbadin@inra.fr

## Résumé

*Le contrôle des maladies transmissibles est une préoccupation majeure du secteur des productions animales. Pour les maladies non réglementées, chaque éleveur décide la mise en place éventuelle d'actions de maîtrise. Pour aider à la coordination entre éleveurs, des organisations proposent des approches collectives de maîtrise, s'appuyant sur des incitations. Nous souhaitons concevoir un outil d'aide à la coordination, basé sur le cadre des Leader-Follower Markov Decision Processes (LF-MDP). La résolution exacte étant difficile, nous avons proposé des méthodes de résolution exacte et approchée (basée sur l'agrégation d'états). Nous avons appliqué ces méthodes sur le cas d'étude du virus du Syndrome Dysgénésique Respiratoire Porcin. La résolution exacte a pu être effectuée jusqu'à  $n = 20$  suiveurs. La résolution approchée a pu être effectuée jusqu'à  $n = 100$  suiveurs et les solutions exacte et approchée sont proches pour  $n \leq 20$ .*

## Mots Clef

Epidémiologie animale, Processus Décisionnel de Markov, Théorie des jeux

## Abstract

*Transmissible disease control is a major concern for the livestock sector. For unregulated diseases, each farmer chooses to implement control actions. To help coordination between farmers, organizations propose collective control approaches using incentives. Our objective is to design a tool, based on the Leader-Follower Markov Decision Processes framework (LF-MDP), to help coordination. Exact solution is hard, however we have proposed an exact solution method and an approximate (using state aggregation) one. We have applied these methods to a case study concerning the Porcine Reproductive and Respiratory Syndrome (PRRS) virus. Exact solution is possible for up to  $n = 20$  followers. Approximate solution could be performed for up to  $n = 100$  followers and we observed similar exact and approximate solutions for  $n \leq 20$ .*

## Keywords

Animal epidemiology, Markov Decision Process, Game theory

## 1 Introduction

Le contrôle des maladies transmissibles est une préoccupation majeure pour le secteur des productions animales. On sépare classiquement les maladies épidémiques des maladies endémiques. Les maladies épidémiques ont une propagation souvent très rapide et sont majoritairement réglementées (mise en place obligatoire de mesures de maîtrise). En revanche, les maladies endémiques sont souvent non réglementées et sont présentes à divers niveaux de prévalence. Leur présence induit une diminution de compétitivité des élevages liée à des pertes économiques [7, 2]. Pour ces maladies, chaque éleveur choisit de mettre en place ou non des actions de maîtrise en se basant sur différents critères dont sa perception du risque. Les critères et la perception diffèrent d'un éleveur à un autre. Or, pour les maladies transmissibles, le choix d'une stratégie de maîtrise dans une exploitation va influencer la propagation de l'agent pathogène dans cette exploitation, et donc le risque de transmission aux autres exploitations (par voisinage ou achat d'animaux), créant une externalité positive [12]. D'autres élevages peuvent choisir de ne rien faire du fait des actions de leurs voisins, adoptant alors un comportement de "free rider" [11]. Pour limiter ces comportements, il est nécessaire de coordonner les décisions des élevages. Pour aider à la coordination entre élevages, des organisations professionnelles regroupant un ensemble d'élevages proposent parfois des approches collectives de maîtrise d'une maladie pouvant s'appuyer sur des incitations financières. Comme le collectif n'impose pas d'actions de maîtrise à ses membres, il doit tenir compte du comportement possible des éleveurs. Un outil permettant de définir ces stratégies collectives serait utile. L'enjeu est donc de proposer des approches d'aide à la décision collective tenant compte des prises de décisions des éleveurs. Les décisions sont prises régulièrement (séquentielles) et réajustées en tenant compte de la situation épidémiologique courante (adapta-

tion). On s'intéresse donc aux approches d'optimisation pour la prise de décisions séquentielles en environnement incertain. Les méthodes de l'Intelligence Artificielle sont d'intérêt pour répondre à nos questions. En santé animale, peu de travaux utilisant de telles approches ont été proposés et ils sont souvent limités à un seul décideur.

Du point de vue de l'Intelligence Artificielle, notre question en santé animale peut être vue comme une forme de problème décisionnel multi-agent et séquentiel. Tous les agents (éleveurs) maximisent leurs objectifs propres lorsqu'ils décident de mettre en place ou non une mesure de maîtrise. En parallèle, un pseudo-agent du type leader (l'organisation) a pour objectif de minimiser l'impact de la maladie dans l'organisation en donnant des incitations aux éleveurs pour que leurs comportements se rapprochent de l'optimum collectif. Du fait des interactions très spécifiques entre l'organisation et les éleveurs, nous sommes face à un problème de type *Leader-Follower Markov Decision Process (LF-MDP)* [23]. La résolution des LF-MDP est la plupart du temps réalisée par simulation, par des approches de type *Apprentissage par Renforcement* [23, 10]. Pour envisager une résolution exacte, nous nous sommes intéressés à un sous-problème plus simple, dont les hypothèses étaient acceptables pour répondre à nos questions en santé animale. Nous avons supposé que le nombre de statuts des suiveurs vis à vis de la maladie était petit et que les suiveurs ne sont pas individuellement identifiés par l'organisation, excepté par leur statut vis à vis de la maladie. Cette dernière hypothèse signifie que 2 éleveurs ayant le même statut vont utiliser la même politique (qui peut être stochastique). Nous avons pu résoudre de manière exacte ce problème pour des petits nombres de suiveurs ( $n \leq 20$ ). Pour de plus grandes valeurs, nous avons proposé une méthode de résolution approchée, par agrégation d'états.

L'objectif de ce papier est de présenter la formulation par un LF-MDP du problème de coordination de décision individuelle par une organisation ainsi que les approches développées pour sa résolution exacte et approchée. Après un état de l'art, nous présentons notre contribution à la résolution de LF-MDP et ensuite un cas d'étude permettant d'illustrer notre approche sur une question en santé animale.

## 2 Etat de l'art

En santé animale, il existe des travaux s'intéressant aux problématiques de coordination de décisions. Le plus souvent, ceux-ci s'intéressent à la modélisation et à la simulation de la propagation d'une épidémie dans une région. Ils se limitent alors à évaluer, par simulation, plusieurs scénarios représentant différentes actions imposées par un collectif et supposées mises en place dans tous les élevages [15, 24]. Il existe toutefois quelques approches pour la conception par optimisation de stratégies de contrôle collectives [9, 13], ou individuelles [25, 19, 20].

Le non suivi des actions proposées par une organisation par une partie des éleveurs est un facteur impactant la propaga-

tion [27]. De rares travaux prennent en compte le non-suivi des actions proposées, dans le cadre de l'optimisation de stratégies collectives [26].

Quelques travaux, en économie [23] et en Intelligence Artificielle / Recherche Opérationnelle [28], se sont intéressés à la coordination d'un leader et de plusieurs suiveurs, dans des problèmes de décision. Néanmoins, ces travaux se limitent à des comparaisons par simulation de politiques de coordination prédéfinies ou à l'étude de systèmes ne comprenant que quelques individus ou un seul pas de temps. En négligeant le rôle particulier de l'organisation, des approches de type *Competitive Markov Decision Process* aussi appelées *Stochastic Game* [8] ou de type *Mean Field Game* [14] seraient envisageables pour traiter le problème qui nous intéresse. Du fait des interactions très spécifiques entre l'organisation et les éleveurs, nous sommes plutôt dans les problèmes de type *Leader-Follower Markov Decision Process (LF-MDP)* [23] et *Dynamic Principal-Agent Problem* [18]. Pour résoudre ces problèmes, il est nécessaire de résoudre un grand nombre de jeux à  $n$  joueurs, à chaque pas de temps, pour construire la politique des  $n$  éleveurs. La complexité (exponentielle) de cette étape se répercute sur la résolution des LF-MDP. C'est pourquoi la plupart des approches existantes opèrent par simulation, en utilisant des méthodes d'apprentissage par renforcement [23, 10]. Mais ces méthodes ne permettent pas de traiter facilement le contrôle des maladies non réglementées en santé animale, où il faut considérer : (1) des horizons de plusieurs pas de temps et (2) plusieurs centaines d'éleveurs en interaction. Nous avons, dans [21, 22], proposé une approche générique de résolution pour ces problèmes, entrelaçant programmation dynamique et calculs d'équilibres de Nash. Cette approche est basée sur une hypothèse simplificatrice de *substituabilité* des suiveurs dans les LF-MDP. Pour le cas d'étude qui nous occupe ici, elle permet de résoudre, de manière exacte, des problèmes avec 20 suiveurs. En utilisant une approximation basée sur l'agrégation des suiveurs, nous pouvons résoudre, de manière approchée, des problèmes avec au moins 100 suiveurs.

## 3 Formalisation et résolution

Nous considérons 2 types d'agents : un leader (l'organisation) et des suiveurs (les éleveurs). Le leader cherche à maximiser sa propre fonction d'utilité qui est dépendante de la quantité d'incitations versées (fonction des décisions des suiveurs) et de l'état du système (niveau d'infection dans l'organisation). Les suiveurs maximisent leur propre fonction d'utilité dépendant de leur statut individuel (pertes dues à la maladie) et des actions de maîtrise si elles sont mises en place (coût des actions diminué des incitations du leader si il en distribue). Dans notre situation, le leader influence les fonctions de récompense des suiveurs via les incitations. Les décisions individuelles, de maîtrise ou non, des suiveurs influencent la dynamique du système (propagation de la maladie) et donc indirectement la fonction de récompense du leader.



Nous présentons d'abord le cadre LF-MDP (Section 3.1) proposé par [23]. Puis, nous décrivons brièvement un algorithme de type *Backward Induction* "naïf", de complexité exponentielle en le nombre de suiveurs, permettant de résoudre un LF-MDP (Section 3.2). Enfin, nous rappelons les principes et la complexité des algorithmes de résolution, exacte et approchée, que nous avons proposés dans [22] (Section 3.3).

### 3.1 Formalisation LF-MDP

**Etats, actions, transitions et récompenses.** Un LF-MDP [23] modélise un processus de décision séquentielle dans l'incertain impliquant un *leader* et  $n \geq 1$  *suiveurs* (*followers*). Il se définit, en horizon fini<sup>1</sup>, par :  $\mathcal{M} = \langle n, \Sigma, A^L, \{A_i^F\}_{i=1..n}, T, r^L, \{r_i^F\}_{i=1..n}, H \rangle$ , où :

- $n$  est le nombre de suiveurs.
- $\Sigma$  est l'espace d'états joint du leader et des suiveurs. Dans le cas général, il peut être factorisé :  $\Sigma = S^L \times S_1^F \times \dots \times S_n^F$ .
- $A^L = \{1, \dots, m\}$  est l'ensemble fini des actions du leader.
- $A_i^F = \{1, \dots, p_i\}$  est l'ensemble fini des actions du suiveur  $i$ . Pour simplifier l'exposé, nous supposons que ces ensembles sont identiques pour tous les suiveurs :  $A^F = \{1, \dots, p\}$ .
- $T : \Sigma \times (A^F)^n \times \Sigma \rightarrow [0, 1]$  est la fonction de transition jointe.  $T(\sigma' | \sigma, \{a_i^F\}_{i=1..n})$  est la probabilité de passer de l'état  $\sigma$  à l'état  $\sigma'$ , lorsque les actions des suiveurs sont  $a^F = \{a_i^F\}_{i=1..n}$ . Notons que l'action du leader n'influence pas directement ces probabilités.
- $r^L : \Sigma \times A^L \times (A^F)^n \rightarrow \mathbb{R}$  est la fonction de récompense instantanée du leader.
- $r_i^F : \Sigma \times A^L \times A^F \rightarrow \mathbb{R}$  est la fonction de récompense instantanée du suiveur  $i$ .
- $H$  est le nombre d'étapes de décision.

**Stratégies du leader et des suiveurs.** Nous supposons, comme classiquement dans les Processus Décisionnels de Markov (MDP, pour *Markov Decision Processes*) à horizon fini, que les agents choisissent leurs actions à l'étape  $t$  en suivant des stratégies *non-stationnaires*,  $\delta_t^L, \{\delta_{t,i}^F\}_{i=1..n}$ . Nous nous restreindrons à des stratégies *Markoviennes*, *stochastiques*, car les stratégies d'équilibre d'un LF-MDP vérifient ces hypothèses [23].

$\delta_t^L(a^L | \sigma)$  est la probabilité que  $a^L \in A^L$  soit choisie par le leader à l'étape  $t$ , connaissant l'état courant  $\sigma \in \Sigma$ .  $\delta_{t,i}^F(a_i^F | \sigma, a^L)$  est la probabilité que  $a_i^F \in A^F$  soit choisie par le suiveur  $i$  à l'étape  $t$ , connaissant l'état courant  $\sigma$  et l'action  $a^L$  choisie par le leader<sup>2</sup>.

Dans un LF-MDP, la stratégie optimale du leader est *déterministe* [23, 22] :  $\delta_t^L(a^L) \in \{0, 1\}$ . Aussi, nous écrivons :  $a^L = \delta_t^L(\sigma)$ . Dans le cas (non général) où les stratégies

des suiveurs sont déterministes, nous écrirons également :  $a_i^F = \delta_{t,i}^F(\sigma, a^L)$ .

#### Valeurs d'une stratégie jointe, stratégies d'équilibre.

Soit  $\Delta = \{\delta_t^L, \{\delta_{t,i}^F\}_{i=1..n}\}_{t=1..H}$  une stratégie jointe fixée du leader et des suiveurs. Les *valeurs*  $Q_\Delta^L$  et  $Q_\Delta^{F,i}$  de cette stratégie jointe pour le leader et pour les suiveurs sont définies, à chaque pas de temps et dans chaque état joint, par :

$$Q_\Delta^L(\sigma, t) = E \left[ \sum_{t'=t}^H r_{t'}^L \mid \Delta, \sigma \right], \quad (1)$$

$$Q_\Delta^{F,i}(\sigma, t) = E \left[ \sum_{t'=t}^H r_{t',i}^F \mid \Delta, \sigma \right]. \quad (2)$$

Résoudre un LF-MDP consiste à trouver une stratégie d'équilibre,  $\Delta^* = \{\delta_t^{L*}, \{\delta_{t,i}^{F*}\}_{i=1..n}\}_{t=1..H}$ , pour le leader et les suiveurs.

#### Définition 1 (stratégie d'équilibre d'un LF-MDP)

$\Delta^* = \{\delta_t^{L*}, \{\delta_{t,i}^{F*}\}_{i=1..n}\}_{t=1..H}$  est une stratégie d'équilibre si et seulement si elle vérifie,  $\forall t, \delta_t^L, \{\delta_{t,i}^F\}, \sigma \in \Sigma$  :

$$Q_{\Delta^*}^L(\sigma, t) \geq Q_{\Delta^* \downarrow \delta^L}^L(\sigma, t), \forall \delta^L, \quad (3)$$

$$Q_{\Delta^*}^{F,i}(\sigma, t) \geq Q_{\Delta^* \downarrow \delta_i^F}^{F,i}(\sigma, t), \forall i, \delta_i^F. \quad (4)$$

$\Delta^* \downarrow \delta^L$  (resp.  $\Delta^* \downarrow \delta_i^F$ ) est une stratégie jointe où les  $\delta_t^{L*}$  (resp.  $\delta_{t,i}^{F*}$ ) ont été remplacées par des stratégies arbitraires  $\delta_t^L$  (resp.  $\delta_{t,i}^F$ ),  $\forall t (\forall i)$ .

Dans [22], nous avons exploité les résultats de [23, 8] montrant l'existence d'une stratégie d'équilibre dans laquelle les stratégies du leader sont déterministes, pour définir un algorithme exact, par programmation dynamique arrière, de calcul de stratégies jointes d'équilibre d'un LF-MDP.

### 3.2 Résolution

Soit  $\mathcal{M}$ , un LF-MDP. Une stratégie d'équilibre  $\Delta^*$  peut être calculée par l'algorithme suivant [21] :

**Pas de temps final.** Au pas de temps final,  $H$ , tout suiveur  $i$  appliquant l'action  $a_i^F \in A^F$  alors que l'état joint est  $\sigma$  et que l'action du leader est  $a^L$ , reçoit une récompense immédiate  $r_i^F(\sigma, a^L, a_i^F)$ , indépendante des actions des autres suiveurs. On peut toujours trouver une stratégie  $\delta_{H,i}^{F*}$  déterministe<sup>3</sup> :

$$\delta_{H,i}^{F*}(\sigma, a^L) \in \arg \max_{a_i^F \in A^F} r_i^F(\sigma, a^L, a_i^F) \text{ et}$$

$$Q_{\Delta^*, a^L}^{F,i}(\sigma, H) = \max_{a_i^F \in A^F} r_i^F(\sigma, a^L, a_i^F), \forall (\sigma, a^L). \quad (5)$$

Nous définissons la récompense espérée du leader à tout pas de temps  $t \in \{1, \dots, H\}$ , pour une stratégie jointe

1. Nous considérerons, pour raisons de simplicité des notations, des fonctions de transition et de récompense stationnaires, mais cette limitation est triviale à lever.

2. La caractéristique principale d'un LF-MDP est que le leader communique son action aux suiveurs, afin d'influencer leurs choix d'action.

3. Au cas où plusieurs actions généreraient la même récompense maximale, il suffit de choisir l'une d'elles, arbitrairement.

stochastique des suiveurs,  $\delta_t^F = \{\delta_{t,i}^F\}_{i=1..n}$  :

$$r_{\delta_t^F}^L(\sigma, a^L) = \sum_{a^F} \left( \prod_{i=1}^n \delta_{t,i}^F(a_i^F | \sigma, a^L) \right) r^L(\sigma, a^L, a^F). \quad (6)$$

Donc, pour le leader au pas de temps  $H$  :

$$\begin{aligned} \delta_H^{L*}(\sigma) &\in \arg \max_{a^L \in \mathcal{A}^L} r_{\delta_H^{F*}}^L(\sigma, a^L), \\ Q_{\Delta^*}^L(\sigma, H) &= \max_{a^L \in \mathcal{A}^L} r_{\delta_H^{F*}}^L(\sigma, a^L), \forall \sigma. \end{aligned} \quad (7)$$

**Etapes d'induction.** La stratégie jointe (stochastique) d'équilibre des suiveurs à l'étape  $t < H$  est calculée inductivement, en fonction des stratégies d'équilibre des pas de temps suivants<sup>4</sup>, par la résolution de jeux sous forme normale à  $n$  joueurs, pour chaque couple  $(\sigma, a^L)$ .

La valeur pour le suiveur  $i$  de l'action jointe des suiveurs,  $a^F$ , dans l'état  $\sigma$  au pas de temps  $t$ , après une action du leader,  $a^L$ , et en supposant qu'une stratégie jointe d'équilibre est appliquée par la suite, est :

$$\begin{aligned} G_{\sigma, a^L, \Delta^*}^t(i, a^F) &= r_i^F(\sigma, a^L, a_i^F) \\ &+ \sum_{\sigma'} T(\sigma' | \sigma, a^F) Q_{\Delta^*}^{F,i}(\sigma', t+1). \end{aligned} \quad (8)$$

Soit  $\{\alpha_1^*, \dots, \alpha_n^*\}$ , une solution du jeu  $G_{\sigma, a^L, \Delta^*}^t \cdot \alpha_i^*(a_i^F)$  est la probabilité que  $i$  "joue"  $a_i^F \in A^F$ . Une politique d'équilibre des suiveurs est obtenue par :  $\delta_{t,i}^{F*}(a_i^F | \sigma, a^L) = \alpha_i^*(a_i^F), \forall i$  et

$$Q_{\Delta^*, a^L}^{F,i}(\sigma, t) = \sum_{a^F} \left( \prod_{j=1}^n \alpha_j^*(a_j^F) \right) G_{\sigma, a^L, \Delta^*}^t(i, a^F). \quad (9)$$

Puisqu'une action  $a^L$  du leader détermine un équilibre de Nash pour les suiveurs, à travers l'équation (9), la stratégie optimale du leader peut être calculée en résolvant un simple MDP non-stationnaire  $\langle \Sigma, \mathcal{A}^L, \{T_{\delta_t^{F*}}\}, \{r_{\delta_t^{F*}}^L\}_{t=1..H}, H \rangle$ , dans lequel les  $\{r_{\delta_t^{F*}}^L\}_{t=1..H}$  ont été définis plus haut et

$$T_{\delta_t^{F*}}(\sigma' | \sigma, a^L) = \sum_{a^F} \prod_{j=1}^n \delta_{t,j}^{F*}(a_j^F | \sigma, a^L) T(\sigma' | \sigma, a^F). \quad (10)$$

Les fonctions  $T_{\delta_t^{F*}}$  et  $r_{\delta_t^{F*}}^L$  sont calculées au fur et à mesure qu'elles sont utilisées dans l'algorithme d'induction arrière. Cet algorithme calcule des stratégies  $\delta_t^{L*}$  et des fonctions de valeur  $Q_{\Delta^*}^{L*}$  optimales, par induction arrière :

$$\begin{aligned} \delta_t^{L*}(\sigma) &\in \arg \max_{a^L \in \mathcal{A}^L} \left\{ r_{\delta_t^{F*}}^L(\sigma, a^L) \right. \\ &\left. + \sum_{\sigma' \in \Sigma} T_{\delta_t^{F*}}(\sigma' | \sigma, a^L) Q_{\Delta^*}^{L*}(\sigma', t+1) \right\}, \\ Q_{\Delta^*}^L(\sigma, t) &= \max_{a^L \in \mathcal{A}^L} \left\{ r_{\delta_t^{F*}}^L(\sigma, a^L) \right. \\ &\left. + \sum_{\sigma' \in \Sigma} T_{\delta_t^{F*}}(\sigma' | \sigma, a^L) Q_{\Delta^*}^{L*}(\sigma', t+1) \right\}. \end{aligned} \quad (11)$$

4. Cette stratégie jointe, découlant de la recherche d'équilibres de Nash dans des jeux sous forme normale, n'est pas unique.

La complexité de cet algorithme est calculée dans [22]. Dans cet article, on montre également comment la complexité algorithmique peut être réduite grâce à certaines hypothèses sur la structure du problème. Dans la Section suivante nous décrivons brièvement ces résultats. Pour une description plus complète, le lecteur se référera à [22].

### 3.3 Considérations de complexité algorithmique

Les différentes étapes de l'algorithme LF-MDP générique ont des complexités variées.

**Etape 1 : Génération de jeux en forme normale.** Afin de calculer les stratégies d'équilibre des suiveurs, il est nécessaire de construire des jeux en forme normale,  $G_{\sigma, a^L, \Delta^*}^t$  (Equation 8). Chaque jeu comprend  $O(n \times |A^F|^n)$  éléments, et il y a  $|\Sigma| \times |A^L|$  tels jeux. La complexité temporelle de la construction de ces jeux est donc en  $O(n \times |A^F|^n \times |\Sigma| \times |A^L|)$ . Néanmoins, il n'est nécessaire de stocker qu'un jeu à la fois dans l'algorithme.

**Etape 2 : Calcul et stockage des stratégies jointes d'équilibre des suiveurs.** Les stratégies jointes d'équilibre des suiveurs sont obtenues à partir de la résolution des jeux précédemment définis. Leur stockage nécessite un espace en  $O(n \times |\Sigma| \times |A^F| \times |A^L|)$ . De plus, leur calcul nécessite de résoudre de nombreux jeux, chaque résolution étant elle-même difficile<sup>5</sup>.

**Etape 3 : Calcul des fonctions de transition et de récompense du leader.** Les fonctions de transition  $T_{\delta_t^{F*}}$  sont calculées grâce à l'équation 10. Ce calcul nécessite un espace en  $O(|\Sigma|^2 \times |A^L|)$  et un temps en  $O(n \times |A^F|^n \times |\Sigma|^2 \times |A^L|)$ . Le calcul des fonctions de récompense  $r_{\delta_t^{F*}}^L$  nécessite un espace en  $O(|\Sigma| \times |A^L|)$  et un temps en  $O(n \times |A^F|^n \times |\Sigma| \times |A^L|)$ , en utilisant l'équation 6.

**Etape 4 : Calcul de la stratégie optimale du leader.** Le calcul de  $\delta_t^{L*}$  nécessite un espace en  $O(|\Sigma|)$  et un temps en  $O(|\Sigma|^2 \times |A^L|)$ , en utilisant l'équation 11.

**Réduction de la complexité.** Dans [22], nous avons montré que la propriété de *substituabilité* permettait de simplifier la résolution d'un LF-MDP. Les suiveurs sont substituables dans un LF-MDP, si, pour chacun d'entre eux, seul son état propre et les *nombre*s de suiveurs dans chaque état déterminent sa fonction de transition et de récompense. Sous cette hypothèse, un algorithme de résolution de LF-MDP de complexité réduite peut être défini, dont l'espace d'états  $\Sigma^L$  est formé de l'ensemble des vecteurs d'entiers positifs ou nuls,  $c = (c_1, \dots, c_k)$ , dont la somme est  $n$ . Nous obtenons alors  $|\Sigma^L| = O(n^k)$  et, plus généralement, la taille des divers objets du LF-MDP réduit devient polynomiale en  $n$ . En conséquence, la résolution devient polynomiale<sup>6</sup>.

5. Chaque résolution est un problème PPAD-complet, où PPAD est une classe de complexité supposée inclure *strictement* la classe P.

6. Sauf la construction de  $\bar{T}$ , qui reste exponentielle, mais peut être "approchée" en temps polynomial.

TABLE 1 – Complexité de différents éléments utiles dans la résolution d’un LF-MDP dans le cadre d’une résolution naïve, ou exploitant la substituabilité seule, ou incluant également une agrégation des nombres de suiveurs dans chaque état.

	Résolution naïve	Substituabilité	+ Agrégation
Taille de l’espace d’états	$ \Sigma  = O(k^n)$	$ \Sigma^L  = O(n^k)$	$ \overline{\Sigma^L}  = O(K^k)$
Taille de l’espace d’actions des suiveurs	$ A^F ^n$	$ A^F ^k$	$ A^F ^k$
Taille des matrices de transition	$ T  = O(k^{2n} A^F ^n)$	$ \overline{T}  = O(n^{2k} A^F ^k)$	$ \hat{T}  = O(K^{2k} A^F ^k)$
Taille d’un jeu	$ G_{\sigma, a^L, \Delta^*}^t  = O(n A^F ^n)$	$ G_{c, a^L, \Delta^*}^t  = O(k A^F ^k)$	$ G_{\kappa, a^L, \Delta^*}^t  = O(k A^F ^k)$
Nb jeux par étape	$nb =  A^F ^n  A^L $	$nb =  A^F ^k  A^L $	$nb = K^k  A^L $

Néanmoins, une complexité en  $O(n^k)$  peut être problématique, si  $n \geq 100\dots$  Aussi, nous avons proposé une méthode de résolution approchée des LF-MDP substituables, consistant à *agréger* les vecteurs  $c$ . Pour se faire, nous avons proposé de considérer une partition de l’ensemble  $\{0, \dots, n\}$  en  $K + 2$  intervalles non-homogènes, où  $K$  est un entier divisant  $n : I_0 = \{0\}, I_{K+1} = \{n\}$  et  $I_i = \{\frac{(i-1)n}{K} + 1, \dots, \frac{in}{K}\}, \forall i = 1, \dots, K$ . Nous avons défini les *états agrégés* comme des k-uplets d’entiers  $(\kappa_1, \dots, \kappa_k) \in \{0, \dots, K\}^k$ . L’ensemble des états agrégés correspond à l’ensemble des combinaisons d’ensembles  $I_{\kappa_1} \times \dots \times I_{\kappa_k}$ , telles qu’il existe un état  $c \in \Sigma^L, c_h \in I_{\kappa_h}, \forall h = 1, \dots, k$ . L’ensemble  $\overline{\Sigma^L}$  de ces états agrégés vérifie  $|\overline{\Sigma^L}| = O(K^k)$ . Sa taille est donc indépendante de  $n$ . Plus généralement, tous les objets d’un LF-MDP agrégé ont une taille indépendante de  $n$ , ce qui permet une résolution en temps quasiment indépendant de  $n$ . Seule, encore une fois, la construction de la matrice de transition approchée reste exponentielle en  $n$ , mais peut être approchée par simulation. La table 1 regroupe des éléments de complexité des différents sous-modèles.

## 4 Cas d’étude

### 4.1 Modèle

Notre cas d’étude concerne la coordination d’un ensemble d’éleveurs afin de limiter la propagation du virus du Syndrome Dysgénésique Respiratoire Porcin (SDRP) [17]. SDRP est une maladie endémique non réglementée. Elle impacte la santé et le bien-être des animaux ainsi que l’économie de l’exploitation. Dans certaines zones (par exemple : [16, 5]), des actions sont proposées par des organisations, dont des incitations financières permettant de limiter le coût des actions de maîtrise pour les éleveurs. Pour les autres zones souhaitant mettre en place une coordination, un outil d’aide à la décision serait d’intérêt. Pour définir notre modèle LF-MDP, nous présentons d’abord les statuts et les transitions entre statuts au niveau individuel en fonction des décisions individuelles. Pour les transitions, par souci de lisibilité, on donne un diagramme par action, mais des éleveurs peuvent retenir des actions différentes au même instant s’ils ont des statuts différents ou même si leurs statuts sont identiques, si la stratégie ob-

tenue lors du calcul de l’équilibre de Nash est mixte.

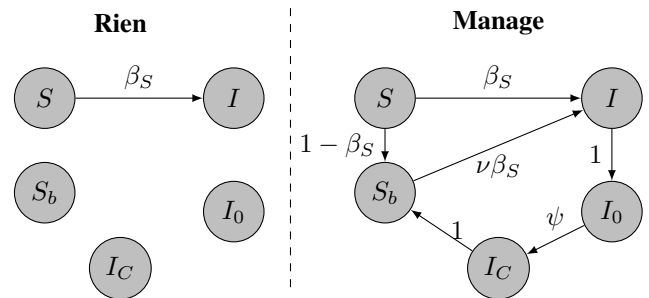


FIGURE 1 – Transitions individuelles pour les suiveurs, en fonction des actions retenues. Les auto-transitions ne sont pas représentées pour simplifier le graphique.

Pour SDRP, nous avons défini, pour les suiveurs, 5 statuts vis à vis de la maladie et 2 actions possibles "Rien" et "Manage" avec les transitions suivantes (Fig.1) :

- $S$  (non infecté) : un suiveur  $S$  devient  $I$  (infecté) avec une probabilité  $\beta_S$ . Sinon, soit il reste  $S$  (si  $a_S^F = 0 \iff$  Rien) soit il devient  $S_b$  (si  $a_S^F = 1 \iff$  Manage, correspondant à prendre des mesures de biosécurité).
- $S_b$  (non infecté avec management) : seul  $a_{S_b}^F = 1 \iff$  Manage (correspondant à de la biosécurité) est possible (pas d’arrêt de la maîtrise). En  $S_b$  le risque d’infection et donc de passage en  $I$  est réduit comparé à  $S$ . La probabilité est  $\nu\beta_S$ , avec  $\nu < 1$  qui dépend de la qualité de la biosécurité.
- $I$  (infecté) : soit le suiveur reste  $I$  (si  $a_I^F = 0 \iff$  Rien) soit il devient  $I_0$  (si  $a_I^F = 1 \iff$  Manage, correspondant à de la vaccination).
- $I_0$  (infecté en cours de contrôle) : Seul  $a_{I_0}^F = 1 \iff$  Manage (correspondant à la vaccination) est possible (pas d’arrêt de la maîtrise). La transition vers  $I_C$  est définie avec une probabilité  $\psi$ , modélisant un temps de séjour en  $I_0$  stochastique, représentant un délai avant un contrôle effectif.
- $I_C$  (infecté contrôlé) : soit le suiveur reste  $I_C$  si  $a_{I_C}^F = 0 \iff$  Rien (maintien de la vaccination) soit il devient  $S_b$  si  $a_{I_C}^F = 1 \iff$  Manage pour revenir non infecté (dépeuplement).

Pour la transmission, on suppose que les suiveurs sont tous en contact les uns avec les autres de par leur situation géographique et les achats/ventes d'animaux. En conséquence, seul la proportion totale d'infectés est à prendre en compte pour le taux de transmission. Cela nous permet de faire une hypothèse de *substituabilité* des suiveurs. Le paramètre  $\beta_S$  a une forme fréquence-dépendant [1] :

$$\beta_S(c) = \frac{1}{n} \sum_{h=1}^k \beta(h)c_h + \beta_{out} \quad (12)$$

où  $\beta(h)$  est le taux de transmission de la maladie par les suiveurs du statut  $h$  et  $\beta_{out}$  le taux de transmission par l'extérieur. Ces paramètres ont été fixés par des experts à  $\beta(I) = 0.08, \beta(I_0) = 0.06, \beta(I_C) = 0.01$  et  $\beta_{out} = 0.005$ .

Les états du leader correspondent à la répartition des suiveurs dans les différents statuts :  $\Sigma^L = \{c = (n_S; n_{S_b}; n_I; n_{I_0}; n_{I_C})$  avec  $n_{S_b} + n_I + n_{I_0} + n_{I_C} = n\}$  où  $n_h$  le nombre de suiveurs dans le statut  $h$ . La taille de l'espace d'états est  $|\Sigma^L| = C_{n+4}^4$ . Deux actions sont possibles :  $a^L = 0 \Leftrightarrow Rien$  et  $a^L = 1 \Leftrightarrow Incitation$ . Dans le cas  $a^L = 1$  une proportion ( $perc \leq 1$ ) des coûts de l'action Manage des suiveurs est prise en charge par le leader. Les transitions entre états sont calculées à partir des transitions entre statuts, de manière similaire à [26].

Les fonctions de récompense des suiveurs ( $r_F$ ) et du leader ( $r_L$ ) sont les suivantes, pour  $\sigma = (s_1, \dots, s_n)$  :

$$r^L(\sigma, a^L, a^F) = -c^L(a^L) - \sum_{i=1}^n c^F(s_i)q^L(a^L, a_i^F) - L^L(s_i),$$

$$r_i^F(\sigma, a^L, a_i^F) = -E_{s_i} [L^F(s_i')] - c^F(s_i)q^F(a^L, a_i^F),$$

avec

- $c^L$ , le coût pour le leader de la mise en place des incitations,
- $L^F(s_i)$  les pertes dans le statut  $s_i$  pour le suiveur,
- $L^L(s_i) = red \times L^F(s_i)$  les pertes dans le statut  $s_i$  pour le leader correspondant à une proportion ( $red \leq 1$ ) de celles des suiveurs,
- $c^F(s_i)$  le coût de l'action Manage pour les suiveurs,
- $q^L(a^L, a_i^F)$  la proportion prise en charge par le leader pour les suiveurs retenant l'action Manage :  $q^L(a^L, a_i^F) = perc \leq 1$  si  $a^L = 1$  et  $a_i^F = 1$  et 0 sinon,
- $q^F(a^L, a_i^F)$  la proportion restant à la charge des suiveurs retenant l'action Manage ( $a_i^F = 1$ ) :  $q^F(a^L, a_i^F) = 1 - perc$  si  $a^L = 1$  et 1 si  $a^L = 0$ .

Pour la résolution, seuls 3 statuts nécessitent des décisions individuelles ( $S, I, I_C$ ). Par conséquent, la taille de l'espace d'actions des suiveurs est  $2^3$  et la taille d'un jeu  $3 \times 2^3$ .

## 4.2 Evaluation de l'approximation

Pour valider l'approximation présentée dans la section 3.3, nous avons réalisé, pour  $n \in \{12, 15, 20\}$ , une résolution exacte et des résolutions approchées avec différentes valeurs de  $K$  compatibles avec  $n$ . Pour chaque combinaison

de  $K$  et  $n$ , nous avons calculé les politiques exacte ( $\delta^*$ ) et approchée ( $\delta^K$ ). Puis, nous avons comparé le comportement du modèle avec ces différentes politiques.

Pour le paramétrage du modèle, nous avons considéré plusieurs jeux de paramètres générés en faisant varier les valeurs des paramètres dans des intervalles de possibles donnés par les experts pour évaluer les comportements de notre modèle. Nous avons considéré deux distributions initiales (à  $t = 0$ ) (i)  $\Gamma_U$  uniforme sur tous les états et (ii)  $\Gamma_E$  uniforme seulement sur les états représentant une situation endémique (état avec au moins 40% dans les statuts  $S + S_b$  et au moins 40% dans le statut  $I_C$ ).

Nous avons comparé les politiques et le comportement du modèle selon 6 indicateurs

- *#Diff* Le nombre de pas de temps où la politique approchée diffère de la politique exacte,
- *Max\_Gap* la valeur maximale de la proportion d'états par pas de temps où il y a une différence d'actions entre les politiques,
- *DB<sub>x</sub>* la distance de Battacharyya [3] entre les distributions sur les états du leader au pas de temps final, en partant d'une distribution initiale  $\Gamma_x$  avec  $x = U$  ou  $E$ ,
- *RMSE<sub>x</sub>* l'écart entre les valeurs espérées calculées sur tous les états si  $x = U$  et uniquement les états endémique si  $x = E$  en utilisant la formule :

$$RMSE_x = \sqrt{\sum_{c \in \Sigma^L} ((V^K(c) - V^*(c))^2 \times \Gamma_x(c))},$$

avec  $x = U$  ou  $E$  et  $V^K(\cdot)$  la fonction de valeur de la politique  $\delta^K$ .

## 4.3 Résultats

Plusieurs jeux de paramètres testés conduisent à très peu d'incitations du leader. Dans certains jeux de paramètres, la politique consiste à ne rien faire sauf à un seul pas de temps où l'incitation est retenue pour quelques états. Pour ces jeux de paramètres, les différences liées à l'approximation sont faibles. Pour les illustrations de ce papier, nous avons sélectionné 2 jeux de paramètres avec des incitations réparties sur plusieurs pas de temps (Table 2). Pour ces deux jeux, les profils d'incitations sont conservés : les instants où il y a au moins un état avec une incitation sont les mêmes quelle que soit l'approximation. Néanmoins, les états avec incitations changent selon  $K$ .

Il y a quelques légères différences en termes de politique et de RMSE, mais elles restent faibles (Fig. 2). Il n'y a pas de relation claire entre la valeur de  $K$  et les écarts à la résolution exacte.

Avec notre approximation, le problème peut être résolu pour  $n = 100$  avec  $K = 5$ . Mais du fait de la taille du modèle ( $|\Sigma^L| = 4598251$ ), nous n'avons pas pu calculer la fonction de valeur sur l'espace complet. Si on considère uniquement les états endémiques pour l'application au SDRP, il est néanmoins possible de calculer par simulation les fonctions de valeurs pour ces états.

TABLE 2 – Valeurs des paramètres de transition des suiveurs et des récompenses (leader et suiveurs) dans 2 jeux

Jeu	$\nu$	$\psi$	$L^F(s_i)$ (pertes)	$c^F(s_i)$ (coûts des suiveurs)	$c^L(a_L)$ (coût leader)	$perc$	$red$
A	0.5	0.5	(0,0,6,5,4)	(4,1,4,2,101)	(0,3)	0.5	0.7
B	0.7	0.5	(0,0,4,8,5,6,2,8)	(7.84,1.4,11.76,2.8,101.4)	(0,4,2)	0.7	0.7

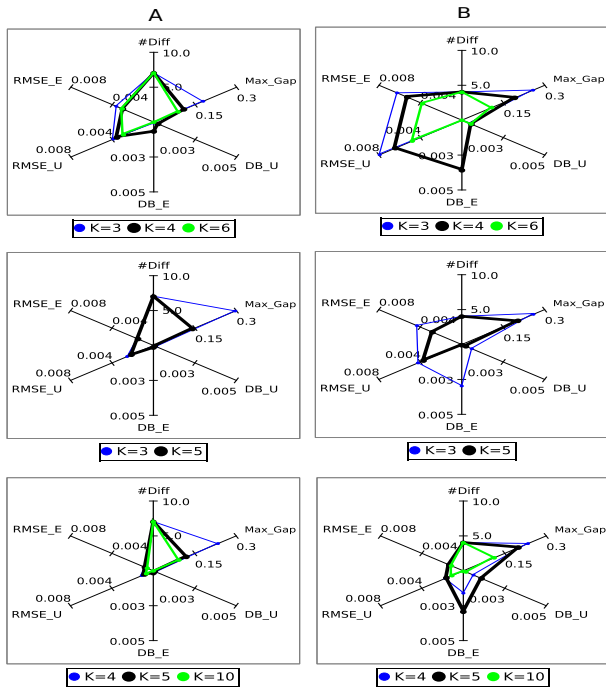


FIGURE 2 – Représentation des différences entre la résolution exacte et approchée, sur différents indicateurs et pour différentes approximations ( $K$ ) pour  $N = 12$  (ligne du haut),  $N = 15$  (ligne du milieu) et  $N = 20$  (ligne du bas)

## 5 Conclusion

Nous avons proposé une approche pour la gestion collective de maladies animales transmissibles non réglementées via la coordination de décisions individuelles, dans le cadre LF-MDP. Nous avons développé un algorithme générique permettant la résolution exacte ( $n \leq 20$ ) ou approchée ( $n \leq 100$ ) d'un problème de contrôle collectif du SDRP. Dans notre cas d'étude, nous avons montré par comparaison, la proximité des solutions exacte et approchée pour  $n \leq 20$ . Les écarts sont d'autant plus acceptables que notre approche est principalement un outil de réflexion pour les décideurs. Pour illustrer l'intérêt de la politique obtenue, qui est adaptative, l'impact de celle-ci devra être comparé à celui des stratégies classiquement considérées qui sont non-adaptatives, comme dans [19, 20].

Ces travaux peuvent être poursuivis selon différents axes. Tout d'abord, bien que la résolution approchée soit possible avec  $n = 100$ , le calcul exact des matrices de transitions est très long pour  $n > 100$  (plusieurs heures/jours). Ces matrices peuvent être approchées par des simulations

non exhaustives des transitions, ce qui permettrait un important gain de temps. Ensuite, nous avons supposé que l'horizon de la prise de décision individuelle était le même que celui de l'organisation. Or les éleveurs raisonnent souvent à un horizon plus court, parfois d'un seul pas de temps. La prise en compte d'horizons différents pour le leader et les suiveurs est possible dans le cadre LF-MDP et il serait intéressant d'étudier l'impact sur la politique optimale du leader de la considération d'un horizon court pour les suiveurs. Enfin, les éleveurs peuvent présenter une *aversion au risque* plus ou moins importante. L'effet de l'attitude vis à vis du risque de gestionnaires forestiers sur leurs politiques de gestion a déjà été étudié dans le cadre MDP [4, 6]. L'extension de ces travaux au cadre LF-MDP est un problème ouvert, important en épidémiologie animale.

Enfin, le cadre LF-MDP est générique et pourrait être adopté dans d'autres domaines que la santé animale. Notre approche peut être utilisée dans de nombreux problèmes où des décideurs maximisant leur profit individuel sont sous l'influence d'un coordinateur (via des actions impactant les jeux entre suiveurs) visant à atteindre son propre objectif. De tels problèmes peuvent se rencontrer en épidémiologie humaine ou végétale (dans ce dernier cas, on se place au niveau de parcelles agricoles gérées par des décideurs indépendant). Toujours dans le cadre de la gestion Environnementale, la taxation des émissions de carbone est un autre domaine d'application potentiel.

## Remerciements

Ce travail a été soutenu par l'ANR, projets ANR-10-BINF-07 (MIHMES) et ANR-13-AGRO-0001-04 (AgroBioSE) et par l'Europe (fond FEDER, Pays-de-la-Loire).

## Références

- [1] M. Begon, M. Bennett, R.G. Bower, N.P. French, S.M. Hazel, J. Turner, A clarification of transmission terms in host-microparasite models : numbers, densities and areas. *Epidemiology and Infection*, 129, pp. 147-153, 2002.
- [2] R. Bennett, The 'direct costs' of livestock disease : the development of a system of models for the analysis of 30 endemic livestock diseases in Great Britain. *Journal of Agricultural Economic* 54 pp. 55-71, 2003
- [3] A. Bhattacharyya, On a measure of divergence between two statistical populations defined by their probability distributions. *Bulletin of the Calcutta Mathematical Society*, 35, pp. 99-109, 1943.
- [4] M. Brunette, S. Couture, J. Laye, Optimising forest management under storm risk with a Markov decision

- process model. *Journal of Environmental Economics and Policy*, 4, pp. 141-163, 2015.
- [5] C.A. Corzo, E. Mondaca, S. Wayne, M. Torremorell, S. Dee, P. Davies, R.B. Morrison, Control and elimination of porcine reproductive and respiratory syndrome virus. *Virus research*, 154, pp. 185-192, 2010.
- [6] S. Couture, M.J. Cros, R. Sabbadin, Risk aversion and optimal management of an uneven-aged forest under risk of windthrow : A Markov decision process approach. *Journal of Forest Economics*, 25, pp. 94-114, 2016.
- [7] K. Ekboir, The role of the public sector in the development and implementation of animal health policies. *Preventive veterinary Medicine* 40 pp. 101-115, 1999.
- [8] J. Filar, K. Vrieze, *Competitive Markov Decision Processes*, Springer, 1996.
- [9] L. Ge, M. Mourits, A.R. Kristensen, R. Huirne, A modelling approach to support dynamic decision-making in the control of FMD epidemics, *Preventive veterinary medicine*, 95, pp. 167-174, 2010.
- [10] J. Hu, M. P. Wellman, Nash Q-learning for general-sum stochastic games, *Journal of Machine-Learning Research*, 4, pp. 1039-1069, 2003.
- [11] Y. Ibuka, M. Li, J. Vietri, G.B. Chapman, A.P. Galvani, Free-riding behavior in vaccination decisions : An experimental study. *PloS one* 9 pp. e87164, 2014.
- [12] E. Klein, R. Laxminaryan, D.L. Smith, C.A. Gilligan, Economic incentives and mathematical models of disease. *Environment and Development Economics* 12 pp. 707-732, 2007.
- [13] M. Kobayashi, T.E. Carpenter, B.F. Dickey, R.E. Howitt, A dynamic, optimal disease control model for foot-and-mouth disease : I. Model description. *Preventive Veterinary Medicine* 79 pp. 257-273, 2007.
- [14] L. Laguzet, G. Turinici, Individual Vaccination as Nash Equilibrium in a SIR Model with Application to the 2009-2010 Influenza A (H1N1), *Bulletin of Mathematical Biology*, 77, pp. 1955-1984, 2015.
- [15] A. Le Menach, E. Vergu, R.F. Grais, D.L. Smith, A. Flahault, Key strategies for reducing spread of avian influenza among commercial poultry holdings : lessons for transmission to humans. *Proceedings of the Royal Society B, Biological Science* 273, pp. 2467-2475, 2006.
- [16] M-F. Le Potier, P. Blanquefort, E. Morvan, E. Albina, Results of a control programme for the porcine reproductive and respiratory syndrome in the French 'Pays de la Loire' region. *Veterinary microbiology*, 55 pp. 355-360, 1997.
- [17] G. Nodelijk, Porcine Reproductive and Respiratory Syndrome (PRRS) with special reference to clinical aspects and diagnosis : A review. *Veterinary Quarterly*, 24, pp. 95-100, 2002.
- [18] E.L. Plambeck, S.A. Zenios, Performance-based incentives in a dynamic principal-agent model. *Manufacturing & service operations management* 2, pp. 240-263, 2000.
- [19] O. Rat-Aspert, C. Fourichon, Modelling collective effectiveness of voluntary vaccination with and without incentives, *Preventive Veterinary Medicine*, 93, pp. 265-275, 2010.
- [20] O. Rat-Aspert, S. Krebs, Individual and collective management of endemic animal diseases : an economic approach, In : *2012 Conference of the International Association of Agricultural Economists*, August 18-24, 2012, Foz do Iguacu, Brazil
- [21] R. Sabbadin, A.F. Viet, A Tractable Leader-Follower MDP Model for Animal Disease Management, In : *27th AAAI Conference on Artificial Intelligence*, pp. 1320-1326, 2013.
- [22] R. Sabbadin, A.F. Viet, Leader-Follower MDP model with factored state space and many followers - followers abstraction, structured dynamics and state aggregation, In : *22nd European Conference on Artificial Intelligence (ECAI 2016)*, pp. 116-124, 2016.
- [23] K. Tharakunnel, S. Bhattacharyya, Single-leader-multiple-follower games with boundedly rational agents. *Journal of Economic Dynamics and Control*, 33, pp. 1593-1603, 2009.
- [24] M.J. Tildesley, N.J. Savill, D.J. Shaw, R. Deardon, S.P. Brooks, M.E.J. Woolhouse, B.T. Grenfell, M.J. Keeling, Optimal reactive vaccination strategies for a foot-and-mouth outbreak in the UK. *Nature* 440, pp. 83-86, 2006.
- [25] N. Toft, A.R. Kristensen, E. Jørgensen, A framework for decision support related to infectious diseases in slaughter pig fattening units. *Agricultural Systems* 85, pp. 120-137, 2005.
- [26] A.-F. Viet, L. Jeanpierre, M. Bouzid, A.-I. Mouadib, Using Markov Decision Process to define an adaptive strategy to control the spread of an animal disease. *Computer and Electronics in Agriculture*, 80, pp. 71-79, 2012.
- [27] A. Vonk Noordegraaf, J.A.A.M. Buijtelts, A.A. Dijkhuizen, P. Franken, J.A. Stegeman, J. Verhoeff, An epidemiological and economic simulation model to evaluate the spread and control of infectious bovine rhinotracheitis in the Netherlands. *Preventive Veterinary Medicine* 36 pp. 219-238, 1998.
- [28] C. Wernz, A. Deshmukh, Unifying temporal and organizational scales in multiscale decision-making. *European Journal of Operational Research* 223, pp. 739-751, 2012.

# **Activité AFIA**

1<sup>er</sup> août 2015 – 31 juillet 2017

Editeurs : Conseil d'Administration de l'AFIA – Année 2017





## Table des matières

### Activité de l'AFIA entre Août 2015 et Juillet 2017

Flyer AFIA. <b>A propos de l'AFIA</b> .....	i
Membres du Conseil d'Administration de l'AFIA. <b>CA AFIA</b> .....	iii
Collèges AFIA. <b>Collèges SIC et SMAA</b> .....	v
Flyer AFIA. <b>Prix de Thèse IA</b> .....	ix
Flyer AFIA. <b>Bulletins Trimestriels</b> .....	xi
Flyer AFIA. <b>Journées propres</b> .....	xiii
Flyer AFIA. <b>Compétitions</b> .....	xxi
Flyer AFIA. <b>Journées communes</b> .....	xxv
Plates-formes AFIA. <b>RFIA 2016, PFIA 2017</b> .....	xliii



# 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 d'en assurer la visibilité.

L'AFIA anime la communauté par l'organisation de grands rendez-vous annuels. L'AFIA organise ainsi chaque année une semaine de l'IA, la « Plate-forme IA » ([PFIA 2013](#) Lille, [RFIA 2014](#) Rouen, [PFIA 2015](#) Rennes, [RFIA 2016](#) Clermont-Ferrand, [PFIA 2017](#) Caen) au sein de laquelle se tiennent les « Rencontres des Jeunes Chercheurs en IA » ([RJCIA](#)) et la « Conférence sur les Applications Pratiques de l'IA » ([APIA](#)) ainsi que des conférences thématiques qui peuvent évoluer d'une année à l'autre sans récurrence obligée.

C'est ainsi que du 3 au 7 juillet 2017 à Caen, lors de son édition 2017, la Plate-Forme IA a hébergé, outre la 15<sup>ème</sup> [RJCIA](#) et la 3<sup>ème</sup> [APIA](#), les 11<sup>èmes</sup> Journées [IAE](#), la 28<sup>ème</sup> Conférence [IC](#), les 12<sup>ème</sup> Journées [JFPDA](#), et les 25<sup>ème</sup> Journées [JFSMA](#). Y était également programmé par l'AFIA une toute nouvelle compétition « [IA des Jeux Vidéos](#) » ainsi que plusieurs journées communes avec d'autres associations.

Fort de soutien de ses 401 adhérents actuellement à jour de leur cotisation, l'AFIA assure :

- Le maintien d'un [Site Web](#) dédié à l'IA, reproduisant également les Brèves de l'IA,
- Une journée recherche annuelle sur les Perspectives et Défis en IA ([PDIA 2016](#)),
- Une journée industrielle annuelle ou Forum Industriel en IA ([FIIA 2017](#)),
- La remise annuelle d'un [Prix de Thèse](#) de Doctorat en IA,
- Le soutien à plusieurs [Collèges](#), actuellement au nombre de 4, ayant leur propre activité :
  - Collège Industriel (depuis janvier 2016),
  - Collège Science de l'Ingénierie des Connaissances (depuis avril 2016),
  - Collège Systèmes Multi-Agents et Agents Artificiels (depuis octobre 2016),
  - Collège Représentation et Raisonnement (depuis avril 2017)
- La parution trimestrielle des [Bulletins](#) de l'AFIA, en accès libre à tous depuis le site web,
- Un lien entre adhérents sur les réseaux sociaux [LinkedIn](#) et [Facebook](#), et [Twitter](#),
- Le [parrainage](#), scientifique et financier de conférences et d'écoles d'été en IA,
- La diffusion mensuelle de Brèves sur les actualités de l'IA en France (mail [breves@afia.asso.fr](mailto:breves@afia.asso.fr)),
- La réponse aux consultations officielles ou officieuses (OPECST, MENESR, MEIN, ANR, ...),
- La réponse aux questions de la presse écrite et de la presse orale, également sur internet.

L'AFIA organise aussi mensuellement des Journées communes avec d'autres Associations (en 2017 : [EF&IA](#) avec le MEDEF, [PS&IA](#) avec la SPS, [IHM&IA](#) avec l'AFIHM, [EIAH&IA](#) avec l'ATIEF, [RO&IA](#) avec la ROADEF, RI&IA avec CORIA) et avec des structures du CNRS (en 2017 : [JI&IA](#) avec le pré-GdR AFAIA, [E&IA](#) avec le COMETS, MACS&IA avec le GdR MACS).

Enfin, l'AFIA contribue à la participation de ses membres aux grands événements de l'IA. Ainsi, les membres de l'AFIA, pour leur inscription à PFIA2017, ont-ils bénéficié d'une réduction équivalente à deux fois le coût de leur adhésion à l'AFIA, leur permettant d'assister à PFIA 2017 su 5 jour au [tarif de 95,00 €](#) !

Rejoignez-vous 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 2016

<b>Président</b>	Yves DEMAZEAU	GT Communication, RFIA 2016
<b>V.-Président</b>	Pierre ZWEIGENBAUM	IM&IA 2016, Recherche
<b>Trésorier</b>	Catherine FARON-ZUCKER	GT Trésorerie, RFIA 2016
<b>Secrétaire</b>	Olivier BOISSIER	Parrainages, Prix de Thèse, RFIA 2016, PDIA 2016
<b>Webmestre</b>	Patrick REIGNIER	Web & Mail, Cloud, Enseignement, Votes
<b>Rédacteur</b>	Florence BANNAY	GT Bulletins, Nuit Info 2016
<b>Membres</b>	Carole ADAM	GT Compétitions, Brèves
	Patrick ALBERT	Communication, PDIA 2016
	Olivier AMI	Bulletins, Applications, Réseaux sociaux
	Audrey BANEYX	GT Web & Mail, Dossier n°92
	Sandra BRINGAY	GT Enseignement, Web
	Cédric BUCHE	RV&IA 2016
	Thomas GUYET	Trésorerie, Recherche, Web
	Frédéric MARIS	Archivage HAL, RFIA 2016, Dossier n°94
	Nicolas MAUDET	Bulletins, RO&IA 2016
	Engelbert MEPHU NGUIFO	GT Journées, EGC&IA 2016
	Davy MONTICOLO	Adhésions, Trésorerie
	Philippe MORIGNOT	IA / Robots 2016, Bulletins
	Philippe MULLER	GT Recherche, TAL&IA 2016
	Bruno PATIN	GT Applications, FIIA 2016, Dossier n°93

## Conseil d'Administration 2017

<b>Président</b>	Yves DEMAZEAU	GT Collèges, Communication, Conférences, PFIA 2017
<b>V.-Président</b>	Pierre ZWEIGENBAUM	Parrainages, Fiches, Collèges, Livre Blanc, PDIA 2017
<b>Trésorier</b>	Catherine FARON-ZUCKER	GT Trésorerie, Livre Blanc, PFIA 2017
<b>Secrétaire</b>	Olivier BOISSIER	GT Livre Blanc, Prix de Thèse, PFIA 2017, E&IA 2017
<b>Webmestre</b>	Audrey BANEYX	GT Web & Mail, Site Web, Cloud
<b>Rédacteur</b>	Florence BANNAY	GT Bulletins, Nuit Info 2017, IA Dossier n°97
<b>Membres</b>	Carole ADAM	GT Compétitions, Brèves, Votes
	Emmanuel ADAM	Fiches, Web & Mail, MACS&IA 2017
	Patrick ALBERT	Communication, Livre Blanc
	Olivier AMI	Bulletins, Dossier n°96, Réseaux sociaux
	Sandra BRINGAY	GT Cartographie, Web & Mail
	Frédéric MARIS	Archivage HAL, PFIA 2017
	Arnaud MARTIN	Postmaster, Site Web, Web & Mail
	Engelbert MEPHU NGUIFO	GT Journées, Collèges, PFIA 2017, Dossier n°98
	Davy MONTICOLO	Web & Mail, EIAH&IA 2017
	Philippe MORIGNOT	Bulletins, Compétitions, Collèges
	Jean-Denis MULLER	Parrainages, Fiches, Collèges, FIIA 2017
	Philippe MULLER	Cartographie, Collèges, Livre Blanc, Dossier n°95
	Bruno PATIN	Collège Industriel, Collèges, FIIA 2017
	Serena VILLATA	Adhésions, Trésorerie, Livre Blanc, FR@IJCAI 2017





# Collège Science de l'Ingénierie des Connaissances

## Objectif du Collège

L'objectif du collège est d'animer la communauté de recherche française en Ingénierie des Connaissances. Les thématiques de recherche de cette communauté sont relatives aux méthodes et outils de l'Intelligence Artificielle dédiées au partage et la gestion des données et des connaissances au sein des organisations ou des communautés, via leur explicitation, représentation, mise en relation, diffusion et maintenance. Cette communauté de recherche est par essence pluridisciplinaire, se rapproche des sciences cognitives d'autres domaines de l'IA comme la représentation des connaissances ou le traitement automatique du langage naturel. Depuis quelques années, sa problématique est bousculée par la disponibilité de grandes quantités de données et de documents sur le web, et par l'objectif ambitieux de faire du web un web sémantique et un web de données liées. Ces recherches portent sur :

- La définition de modèles, méthodes, algorithmes et outils permettant l'acquisition et l'intégration de connaissances ainsi que les raisonnements automatiques sur ces connaissances dans des environnements informatiques, et sur le web,
- La construction, utilisation et maintenance d'ontologies permettant de représenter des connaissances généralistes ou spécifiques, de typer des données et de faciliter leur échange ou leur mise en relation,
- La construction des systèmes à base de connaissances faisant appel au traitement automatique des langues, à l'apprentissage automatique, à la fouille de données, à la recherche d'information, à la recherche opérationnelle, à l'analyse de données,
- La modélisation des environnements informatiques incluant la modalisation de ses acteurs humains ou agents logiciels et de leurs interactions,
- Le traitement de la masse de données et de documents disponibles, leur caractérisation sémantique par l'analyse des données et des contenus pour les partager, diffuser, visualiser, interroger, contextualiser, expliquer, valider, recommander, classer, prédire, approximer, sécuriser, fiabiliser, valider, pour pallier à leur imprécision voire incohérence.

Les défis actuels dans le domaine de l'Ingénierie des Connaissances portent sur l'extrême variabilité des sources de connaissances, les dimensions temporelles et spatiales, les passages à l'échelle avec la massification du Web et l'émergence de dispositifs associant la construction de connaissances nouvelles dans des cadres ontologiques dynamiques.

## Programme de Travail

Le collège Ingénierie des Connaissances a d'abord un rôle de diffusion de l'information. Il gère une liste de diffusion sur l'ingénierie des connaissances (qui compte près de 800 inscrits), contribue à la rubrique « Collèges » du Bulletin de l'AFIA, et maintiendra prochainement une page web du collège IC sur le nouveau 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), ainsi que de la conférence Terminologie et Intelligence Artificielle (TIA) et des Journées Francophones des Ontologies (JFO). Ses membres contribuent également à la fabrication de certaines « Journées Communes » de l'AFIA avec d'autres sociétés savantes, notamment les journées « Informatique médical et Intelligence Artificielle » (IM&IA), « Systèmes d'information et Intelligence Artificielle » (SI&IA), et à l'organisation d'ateliers sur la Plateforme Française d'Intelligence Artificielle (PFIA) et d'ateliers co-localisés avec la Conférence Nationale d'Intelligence Artificielle (CNIA).

### **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 Terminologie et Intelligence Artificielle (TIA) et/ou de la conférence Journées Francophones des Ontologies (JFO), et la plupart 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
- 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 et de TIA  
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
- Pascale Kuntz, membre du comité de programme de IC  
Université de Nantes, LINA
- Alain Mille, membre des comités de programme de IC et de JFO  
Université Claude Bernard Lyon 1, LIRIS
- 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**

Liste de diffusion Ingénierie des connaissances : [info-ic@listes.irisa.fr](mailto:info-ic@listes.irisa.fr)

Comité de pilotage du collège Ingénierie des connaissances : [college-ic-bureau@irit.fr](mailto:college-ic-bureau@irit.fr)

Deux membres du comité de pilotage du collège sont membres du conseil d'administration de l'AFIA :

- Catherine Faron Zucker, responsable du collège IC vis-à-vis de l'AFIA, [faron@unice.fr](mailto:faron@unice.fr)
- Sandra Bringay, [sandra.bringay@lirmm.fr](mailto:sandra.bringay@lirmm.fr)





# Collège Systèmes Multi-Agents et Agents Autonomes

## Objectif du Collège

Le collège SMAA a pour mission de contribuer aux activités menées par l'AFIA, par des actions relatives au domaine de l'IA, émanant de la communauté des Agents Autonomes et des Systèmes Multi-Agents (SMAA). Les thématiques de recherche couvertes par ce collège sont celles relevant des Journées Francophones en Systèmes Multi-Agents (JFSMA) et du Workshop Affects, Compagnon Artificiel, Interactions (WACAI). Elles portent sur la modélisation à base d'agents logiciels mis en interaction au sein d'environnements partagés pour :

- La résolution collective de problème (modèles de coordination ; apprentissage, adaptation et planification distribuées ; décisions sous incertitudes, etc.),
- La simulation de phénomènes complexes où la modélisation multi-agents apporte un cadre conceptuel permettant la représentation et la simulation de systèmes faisant intervenir différentes entités en interaction
- Le développement de systèmes informatiques décentralisés où l'approche SMA permet l'intégration flexible et la coopération de logiciels et de services autonome
- Le développement de systèmes médiatisés dans lesquels utilisateurs humains et agents artificiels interagissent directement ou indirectement, dans le cadre d'activités collectives de type éducatif, culturel ou social
- L'étude, la modélisation, le développement et l'évaluation de l'interaction de systèmes interactifs avec leurs partenaires (humains ou artefacts).

Les recherches menées sur ces thèmes adoptent fréquemment une démarche pluridisciplinaire (biologie, linguistique, psychologie, sociologie...) dans la modélisation de sociétés d'agents artificiels ou pour la prise en compte de l'activité et des interactions avec des utilisateurs humains au sein de communautés mixtes.

## Programme de Travail

Les missions du collège SMAA concernent le soutien à l'organisation de manifestations scientifiques (conférences, ateliers), l'animation de groupes de travail, l'édition de dossiers techniques ou de numéros spéciaux de journaux sur des thématiques d'intérêt pour la communauté et la diffusion et la communication autour des recherches des communautés françaises sur les disciplines ciblées.

Le collège SMAA accompagne notamment l'organisation régulière des JFSMA et de WACAI en sollicitant un parrainage de l'AFIA sous la forme de remise de prix, réduction des tarifs d'inscription, annonces d'événements sur ses relais de communication... Il propose également un soutien similaire à d'autres événements, ponctuels ou récurrents, relevant de son périmètre scientifique.

Il contribue aux actions initiées par l'AFIA sur la mise en place de journées bilatérales, notamment les journées communes « Réalité Virtuelle & Intelligence Artificielle » (RV & IA), « Interaction Homme-Machine & Intelligence Artificielle » (IHM & IA) ainsi que « Modélisation, Analyse et Conduite des

Systèmes dynamiques & Intelligence Artificielle » (MACS & IA) ; par une participation aux actions du collège industriel de l'AFIA ; par des contributions à la rubrique « Collèges » du bulletin de l'AFIA.

Les actualités du collège SMAA et plus largement des communautés françaises de recherche associées aux conférences JFSMA et WACAI sont publiées régulièrement sur son site web : <http://smaa.lip6.fr/>

### **Comité de Pilotage**

Le comité de pilotage du collège SMAA est constitué pour la période 2016-2018 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
- Pierre Chevaillier  
Ecole Nationale d'Ingénieurs de Brest, Lab-STICC
- Rémy Courdier  
Université de la Réunion, LIM
- Zahia Guessoum  
Université de Reims, CReSTIC/LIP6
- Salima Hassas  
Université Claude Bernard Lyon 1, LIRIS
- Fabien Michel  
Université de Montpellier, LIRMM
- Frédéric Migeon  
Université Paul Sabatier, IRIT
- Maxime Morge  
Université de Lille 1, CRISAL
- Nicolas Sabouret  
Université Paris-Sud, LIMSI
- Laurent Vercouter  
INSA Rouen Normandie, LITIS

### **Contacts**

Liste de diffusion de la communauté francophone SMA : [sma@loria.fr](mailto:sma@loria.fr)

Liste de diffusion de la communauté ACAI : [acai@poleia.lip6.fr](mailto:acai@poleia.lip6.fr)

Responsable du collège SMAA vis-à-vis de l'AFIA : [laurent.vercouter@insa-rouen.fr](mailto:laurent.vercouter@insa-rouen.fr)

Un membre du comité de pilotage du collège est membre du conseil d'administration de l'AFIA :

- Emmanuel Adam [emmanuel.adam@univ-valenciennes.fr](mailto:emmanuel.adam@univ-valenciennes.fr)



**Afia**

Association française  
pour l'Intelligence Artificielle

## Prix de Thèse IA 2016

### Le Jury

- Président : Luis Farinas Del Cerro (IRIT, CNRS)
- Membres : Leila Amgoud, Christian Bessière, Marie Odile Cordier, Hélène Fargier, Jean-Gabriel Ganascia, Malik Ghallab, Andreas Herzig, Jérôme Lang, Pierre Marquis, François Pachet, Henri Prade, Marie Christine Rousset

### Les Lauréats

- Prix de Thèse IA (ex-aequo) : Antoine Cully "Adaptation créative par évolution artificielle". 21/12/2015. Dir. : Stéphane Doncieux et Jean-Baptiste Mouret (ISIR, Univ. Paris 6).
- Prix de Thèse IA (ex-aequo) : Marc Law "Apprentissage de distance pour la comparaison d'images et de pages Web". 20/01/15. Dir. : Matthieu Cord et Stéphane Gançarski (LIP6, Univ. Paris 6).

## Prix de Thèse IA 2017

### Le Jury

- Président : Jean-Gabriel Ganascia (LIP6, UMPC)
- Membres : Leila Amgoud, Nicolas Bredeche, Laurence Devillers, Amal Elfallah-Seghrouchni, Jérôme Euzenat, Jean-Gabriel Ganascia, Brigitte Grau, Jérôme Lang, Pierre Marquis, Henri Prade, Guillaume Raschia, Marie Christine Rousset, Julie Velcin.

### Les Lauréats

- Prix de Thèse IA (ex-aequo) : Michaël Perrot "Théorie et Algorithmes pour l'Apprentissage de Métriques à Comportement Contrôlé". 13/12/16. Dir. : Amaury Habrard (LHC, Univ. St. Etienne).
- Prix de Thèse IA (ex-aequo) : Eric Piette "Une nouvelle approche au *General Game Playing* dirigée par les contraintes". 09/12/16. Dir. : Sylvain Lagrue et Frédéric Koriche (CRIL, Univ. Artois).
- Accessit : Fabrice Mayran de Chamisso "Navigation exploratoire au long de la vie : une approche intégrant planification, navigation, cartographie et localisation pour des robots mobiles disposant de ressources finies". 18/11/16. Dir. : Michaël Aupetit et Laurent Soulier (LIST, Univ. Paris 11).





## 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 2016, comprenait Florence Bannay – Rédactrice en Chef –, Olivier Ami, Dominique Longin, Nicolas Maudet et Philippe Morignot. Le comité de rédaction 2017 toujours sous la responsabilité de Florence Bannay comprend Claire Lefèvre, Dominique Longin et Philippe Morignot.

#### **Bulletin 90**                      **Dossier « PFIA 2015 »**                      **Octobre 2015**

- Ce numéro propose deux compte-rendu de journées communes, l'une, EIAH & IA, organisée en coopération avec l'ATIEF, l'autre, Jeux Vidéo et IA, organisée en coopération avec le Pôle de Compétitivité Images et Réseaux.
- Vous trouverez dans ce numéro le compte-rendu de PFIA 2015. Signalons que les vidéos de ces conférences invitées sont accessibles depuis le site de la plateforme (<http://pfia2015.inria.fr/>), tout comme les actes des différentes conférences et ateliers.

#### **Bulletin 91**                      **Dossier « IA et Transport »**                      **Janvier 2016**

- Vous pourrez lire les compte-rendu de nouveaux événements comme « Science & Télévision » et IA, la journée MACS& IA, et enfin la journée RI & IA. Florence Bannay nous fait un compte rendu du premier défi proposé par l'AFIA lors de la nuit de l'info 2015. Philippe Muller présente la deuxième édition de la journée PDIA consacrée aux Apprentissages.
- Le dossier IA et Transport a été réalisé par Flavien Balbo et René Mandiau avec la présentation d'une quinzaine d'équipes de recherche qui travaillent dans ce domaine.

#### **Bulletin 92**                      **Dossier « IA et Humanités numériques »**                      **Avril 2016**

- Dans ce numéro, nous retrouvons les compte-rendu de trois journées bilatérales tenues récemment, Extraction et Gestion des Connaissances & IA avec EGC, Réalité Virtuelle et IA avec l'AFRV, et Traitement Automatique des Langues et IA, avec ATALA.
- Audrey Baneyx nous propose un dossier sur les Humanités Numériques, qui illustre à travers des exemples de projets ou d'équipes de recherches l'originalité et la vitalité de cette thématique, qui mobilise plusieurs facettes de l'IA autour de questions issues des sciences humaines et sociales.

#### **Bulletin 93**                      **Dossier « Equipes de recherche »**                      **Juillet 2016**

- L'AFIA a tenu en Avril 2016 FIIIA 2016, son premier Forum Industriel de l'IA dont un compte-rendu nous est proposé par Bruno Patin. Cet événement est l'occasion de mettre en valeur les membres fondateurs du Collège Industriel de l'AFI qui a été créé en 2016.

- Ce numéro contient un dossier spécial de présentation d'équipes industrielles de recherche en IA.

**Numéro 94**

**Dossier « RFIA 2016 »**

**Octobre 2016**

- Dans ce dernier bulletin de l'année 2016, vous trouverez les compte-rendu des journées Francophones sur les Ontologies, de la journée Rob&IA organisée avec le GdR Robotique, de RF&IA organisée avec l'AFRIF. On y trouve aussi un compte-rendu des JFRB 2016. Le groupe de travail "Compétition" de l'AFIA nous fait un compte rendu de la première compétition IA sur Robots. On y trouve aussi les prix de Thèse IA 2016.
- Ce numéro contient un dossier sur RFIA 2016 (avec un éditorial de Yves Demazeau, Président de notre association), la conférence française sur notre domaine, édité par Florence Bannay.

**Bulletin 95**

**Dossier « Apprentissage »**

**Janvier 2017**

- Au sommaire de ce bulletin, une nouvelle rubrique : "Le mot du Président". Il les compte-rendu de la journée Classification et IA organisée avec la SFC, de celle organisée avec le GdR MaDICS, de la journée PDIA 2016 sur le thème de l'impact social de l'IA. On y trouve aussi un compte-rendu d'une journée du Représentation et Algorithmes en Pratique du pré-GdR IA. Finalement, Florence Bannay nous livre un compte-rendu du défi de l'AFIA pour la Nuit de l'Info 2016.
- Le dossier de ce bulletin porte sur l'apprentissage. Il est composé d'une réflexion de Philippe Muller sur l'apprentissage, et, dans la continuité de PDIA 2015, par la présentation des travaux menés dans deux équipes : Equipe Interaction Algorithme Silicium (Télécom Bretagne) sur le thème de l'acquisition et la mémorisation d'informations mentales, Equipe Sequel (INRIA Lille-Nord Europe) sur le thème transfert et apprentissage multitâches dans l'apprentissage basé sur le renforcement.

**Bulletin 96**

**Dossier « Entreprises de France »**

**Avril 2017**

- Ce dossier contient les compte-rendu de la journée EF&IA organisée avec le MEDEF, de la journée « Philosophie des sciences et IA » organisée avec la SPS et le DEC, de l'intervention de l'AFIA à EduSpot, et celui de la 3e journée IHM & IA organisée avec l'AFIHM.
- Le dossier de ce bulletin contient le verbatim des discours tenus lors de la journée EF&IA. Plus de 510 inscrits à cette journée, qui a été honorée par l'intervention du Secrétaire d'Etat à la Recherche et à l'Enseignement Supérieur, M. Thierry Mandon.

**Bulletin 97**

**Dossier « Equipes de recherche »**

**Juillet 2017**

- Ce bulletin présente les compte-rendu de la journée EIAH & IA organisé avec l'ATIEF et celui du deuxième Forum Industriel de l'IA (FIIA 2017). Il contient aussi les compte-rendu d'une journée commune entre les GdR IM et le pré-GdR AFAIA, ainsi que de la journée Bioss-IA. Il contient enfin deux hommages à deux chercheurs disparus le trimestre dernier: Alain Colmerauer et Daniel Kayser.
- De nombreuses équipes ont répondu à notre appel, le dossier spécial contient la description d'équipes de différents laboratoires français comme le LGI2P, l'IRISA, l'IRIT, le LIRMM, le LIMSI, le LIS, et le LORIA.



**AfIA**

Association française  
pour l'Intelligence Artificielle

# PDIA 2015

L'Association Française pour l'Intelligence Artificielle organise sa deuxième journée sur les PERSPECTIVES ET DEFIS DE L'INTELLIGENCE ARTIFICIELLE.

La thématique de la journée est de prendre du recul sur les progrès en apprentissage machine (apprentissage artificiel), et de discuter de son avenir par rapport au domaine de l'IA. L'IA est traditionnellement associé à des approches symboliques de manipulation de la connaissance, alors que l'apprentissage machine est essentiellement fondé sur des approches statistiques. Les deux versants sont-ils vraiment opposés ou va-t-on vers une synthèse, et si oui, de quelle façon.

## Date et Lieu

- Date : 08 octobre 2015
- Lieu : Salle A709, Université Paris-Dauphine, Pl. du M. de Lattre de Tassigny, 75016 Paris

## Programme

- 09h40 Introduction de la journée par Yves Demazeau (*Président de l'AFIA*) et Philippe Muller (*coordonnateur du Groupe de Travail Recherche de l'AFIA*)
- 10h00 « Apprentissage (et) statistique : de l'âge de raison à l'empire des normes » par Antoine Cornuejols (*AgroParisTech, Paris*)  
*L'intervention se propose de montrer comment l'apprentissage automatique a d'abord épousé le cœur de l'Intelligence Artificielle : représentation des connaissances et raisonnement, puis comment et pourquoi l'approche de l'apprentissage statistique a éclo dans les années 1980 et s'est finalement imposé comme un paradigme dominant. La conférence se terminera avec une réflexion sur la question : est-ce que le schéma "critère inductif basé sur une formalisation statistique + optimisation" est un horizon indépassable ?*
- 10h45 Pause
- 11h00 « 25 ans de noyaux... Et après ? » par Liva Ralaivola (*Université Aix-Marseille, Marseille*)  
*Les méthodes à noyaux, développées dans les années 90 avec l'avènement des machines à vecteurs de support, ont constitué une avancée majeure en l'apprentissage statistique. Elles ont permis d'aborder de nombreuses tâches avec des validités algorithmique et statistique prouvées et une efficacité pratique attestée. Au temps de l'avènement des problématiques big data et des méthodes du type deep learning, nous donnons quelques pistes pour déterminer la place des méthodes à noyaux dans le paysage de l'apprentissage aujourd'hui.*
- 11h45 « Apprentissage par renforcement et transfert de connaissance » par Alessandro Lazaric (*INRIA Lille – Nord Europe*)  
*L'apprentissage par renforcement (AR) permet d'apprendre par expérience directe comment se comporter de manière optimale dans des environnements incertains. Les techniques existantes permettent comportements quasi-optimaux, mais le processus d'AR doit être redémarré chaque fois que les tâches à résoudre changent. Nous montrons comment l'extraction de connaissance de la solution d'une tâche et le transfert de cette connaissance pour la résolution de nouvelles tâches permet de surmonter cette limitation.*
- 12h30 Buffet



- 14h00 « Acquisition et mémorisation de l'information mentale : de Shannon à l'IA » par Claude Berrou (*Telecom Bretagne, Brest*)  
*Le modèle de communication proposé par Shannon en 1948 peut servir de point de départ dans la compréhension des principes de l'acquisition et de la mémorisation de l'information mentale afin d'en tirer des enseignements précieux pour l'intelligence artificielle. En particulier, les notions d'apprentissage statistique (à comparer au codage de source) et d'apprentissage symbolique (à comparer au codage de canal) peuvent être réunies et exploitées dans un même schéma d'acquisition parcimonieuse et de mémorisation robuste.*
- 14h45 « Apprentissages en interaction » par Michèle Sebag (*CNRS, Paris*)  
*Pour certains problèmes mal définis ou sous-spécifiés, l'apprentissage d'un agent doit faire intervenir l'humain dans la boucle ; le rôle de l'humain est d'indiquer ses préférences quant au comportement le plus approprié de l'agent. La situation présente plusieurs difficultés. Le nombre de jugements de préférence demandé à l'humain doit rester limité. Ces jugements peuvent être incohérents, L'agent doit proposer des choix présentant un bon compromis exploration / exploitation. Enfin, l'humain s'adapte au comportement de l'agent.*
- 15h30 Clôture

## Organisation

Cet événement est organisé par Philippe Muller et Yves Demazeau pour l'Association Française d'Intelligence Artificielle (AFIA).

## Inscriptions

L'accès à la journée PDIA 2015 est limité aux seules personnes qui se seront inscrites avant le 05 octobre 2015 sur le site des inscriptions : <http://inscriptions.afia.asso.fr/>

Cette journée est suivie de l'Assemblée Générale Ordinaire 2015 de l'AFIA, incluant bilans 2015, prévisions 2016 et est précédée par une conférence invitée de Salem Benferhat, nouvel « ECCAI Fellow ».

## Assemblée Générale Ordinaire de l'AFIA

- Date : 08 octobre 2015
- Lieu : Salle A709, Université Paris-Dauphine, Pl. du M. de Lattre de Tassigny, 75016 Paris
- 15h45 « Evolution et traitement des informations imparfaites » par Salem Benferhat (*Université d'Artois, Lens*)  
*L'exposé aborde le problème de l'automatisation des différents types de raisonnement de sens communs. Il fait le point sur des modèles informatiques qui permettent d'une part de prendre en compte des connaissances / ontologies / croyances / préférences incertaines, imprécises, causales et incohérentes et d'autre part de répondre aux requêtes complexes des utilisateurs. L'exposé présente des perspectives de recherche sur des questions pratiques qui tireront profit de l'abondance des données et de l'accessibilité multiple des technologies.*
- 16h15 Assemblée Générale Ordinaire 2015 de l'AFIA
- 18h00 Cocktail





**AFIA**

Association française  
pour l'Intelligence Artificielle

# FIIA 2016

L'AFIA organise son premier FORUM INDUSTRIEL DE L'INTELLIGENCE ARTIFICIELLE. L'objectif est de permettre aux deux communautés académique et industrielle de mieux dialoguer.

Dans ce cadre, la première édition de cette journée aura pour objectif d'initier ce dialogue par une présentation du Collège Industriel puis quelques présentations de ses membres fondateurs afin de donner une première image des besoins de recherche en Intelligence Artificielle de l'industrie. En fin de journée, une session est réservée à des présentations rapides, pour les industriels invités le souhaitant, des préoccupations de recherche.

## Date et Lieu

- Date : 14 avril 2016
- Lieu : Amphi B312, Télécom ParisTech, 46 rue Barrault, 75013 Paris

## Programme

- 09h40 Introduction de la journée par Yves Demazeau (*Président de l'AFIA*) et Bruno Patin (*Coordinateur du Groupe de Travail Applications de l'AFIA*)
- 10h00 « Le Collège Industriel de l'AFIA » par Bruno Patin (*Dassault Aviation, St-Cloud*)
- 10h45 « Décision dans les systèmes mixtes » par Stéphane Durand (*Dassault Aviation, St-Cloud*)
- 11h15 « Des systèmes experts à l'Ingénierie des Connaissances » par Alain Berger (*Ardans, Montigny*)
- 11h45 « Quelques aspects liés à l'IA chez Huawei » par Arnaud Lallouet (*Huawei, Boulogne s / Seine*)
- 12h15 Buffet
- 13h30 « Le Pôle de Compétitivité System@tic et l'IA » par Jean-Luc Beylat (*Nokia, Paris*)
- 14h15 « La gestion des connaissances dans l'industrie » par Florence Amardheil (*Mondeca, Paris*)
- 14h45 « Les senseurs et l'IA » par Thierry Lamarque (*Thales Services, Vélizy*)
- 15h15 « Introduction à l'optimisation cognitive » par Eric Mazeran (*IBM, Nice*)
- 16h00 Succession de présentations rapides de sociétés concernées par l'IA en « Trois planches »
- 17h00 Clôture

## Organisation

Cet événement est organisé par Bruno Patin et Yves Demazeau pour l'Association Française d'Intelligence Artificielle (AFIA).

## Inscriptions

Les inscriptions au Forum (gratuites pour les membres de l'AFIA, 30 € sinon) sont obligatoires et à effectuer sur le site des inscriptions : <http://inscriptions.afia.asso.fr/>. Le déjeuner est offert à toutes les personnes qui se seront inscrites à FIIA 2016 avant le 11/04/16, et seulement à celles-ci.







**AFIA**

Association française  
pour l'Intelligence Artificielle



# PDIA 2016

L'Association Française pour l'Intelligence Artificielle (AFIA) organise sa troisième journée PERSPECTIVES ET DEFIS DE L'IA sur le thème : IMPACT SOCIAL DE L'IA. Les médias rendent abondamment compte des applications de l'IA, de ses promesses et des craintes qu'elle suscite. D'importantes mutations sont en gestation, la vie sociale va connaître des bouleversements qu'il est déjà possible d'esquisser, leur ampleur et leur calendrier faisant débat. En organisant cette journée, l'AFIA veut contribuer à éclaircir ces horizons nouveaux.

Cette journée se tient sous le patronage du MESR Direction Générale de la Recherche et de l'Innovation et de l'ISCC Institut des Sciences de la Communication avec le soutien de la Fondation Telecom.

## Date et Lieu

- Date : 07 octobre 2016
- Lieu : Amphi Thévenin, Télécom ParisTech, 46 rue Barrault, 75013 Paris

## Programme

- 09h30 Introduction de la journée par Yves Demazeau (Président de l'AFIA)
- 09h40 Ouverture de Pascal Estrailier (*Chargé de mission STIC à la MESR DGRI, Direction Générale de la Recherche et de l'Innovation*)
- 09h50 Ouverture de Pascal Griset (*Directeur de l'ISCC, Institut des Sciences de la Communication*)
- 10h00 « La société du calcul. Enjeux et promesses de la massification des données » par Dominique Cardon (Orange Labs, Paris)  
*Les mondes numériques collectent nos traces et nos actions sur le web, produisant des calculs qui régissent nos vies. On propose une réflexion sur le rôle que les algorithmes sont appelés à jouer dans cet univers de massification des données et sur la manière dont nous sommes calculés. Comprendre favorise une critique éclairée de la manière dont le calcul s'introduit dans nos sociétés.*
- 10h45 Pause
- 11h00 « Blockchain, Registres distribués & IA », Laurent Leloup (Président de France Blocktech)  
*Depuis la naissance de la blockchain Bitcoin (2008), l'Intelligence Artificielle impose son indispensable présence au sein d'une technologie en devenir qui va révolutionner l'ensemble des industries avec plus ou moins de disruption. Nous découvrirons ensemble ces évolutions technologiques et les promesses à venir.*
- 11h45 « Comment développer l'éthique des algorithmes et la loyauté des plateformes ? Peut-on esquisser ces nouveaux concepts » par Thierry Marcou (FING)  
*Lancé en 2013, le programme Digiwork s'est donné comme objectif de repenser la place des individus au travail dans la société numérique. Huit pistes d'action ont été explorées dans le cadre de groupes de travail où ont été associés des grandes et petites entreprises, des syndicats, des organisations publiques. Il s'en dégage une vision des nouvelles pratiques du travail de demain...*
- 12h30 Buffet
- 13h45 « Apprendre à vivre avec les systèmes d'IA » par Aymeric Poulain Maubant (PDG Nereys)



- 14h00 « L'IA, levier de la profonde mutation de l'industrie financière » par François Leclerc (chroniqueur et auteur).  
*Les start-ups foisonnent, les banques engagent des alliances et les Gafa développent leurs méga projets. Confrontées à la nécessité de changer de modèle, les institutions financières voient dans la Blockchain l'occasion de réduire leurs coûts et d'optimiser leur fonctionnement, recherchant d'importants gains de productivité afin d'enrayer la chute de leur rendement.*
- 14h45 « L'IA et la métamorphose de la médecine, nouveaux business et intervenants » par Florence Pinaud (La Tribune Hebdo)  
*Plébiscités par les patients, les applications et objets connectés incitent les jeunes médecins à s'associer au développement du monde digital afin de ne pas se laisser complètement dépasser. On présente quelles sont les nouvelles activités de cette médecine en mode numérique et quels sont les acteurs de ce nouveau marché.*
- 15h30 Pause
- 15h45 « L'information revisitée » par Eric Scherer (Directeur de la Prospective France Télévisions)  
*Journalisme automatique, algorithmes en place du rédacteur en chef, bots conversationnels, nouvelles réalités altérées, sont autant de nouveaux défis pour les professionnels des médias et les journalistes qui vivent déjà difficilement depuis 15 ans la mutation numérique.*
- 16h30 Clôture

## Organisation

Cet événement est organisé par Patrick Albert et Olivier Boissier pour l'Association Française d'Intelligence Artificielle (AFIA), avec l'aide de François Leclerc.

## Inscriptions

Les inscriptions à la journée (gratuites pour les membres de l'AFIA ou de la Fondation Telecom, 30 € sinon) sont obligatoires et à effectuer sur le site : <http://inscriptions.afia.asso.fr/> . Le déjeuner est offert à toutes les personnes qui se seront inscrites à PDIA 2016 avant le 04/10/16, et seulement à celles-ci.

La journée est précédée d'une conférence invitée de Thomas Schiex, nouvel « EurAI Fellow » français élu en 2016. Et de l'Assemblée Générale Ordinaire 2016 de l'AFIA, incluant bilans 2016 et prévisions 2017.

## Assemblée Générale Ordinaire de l'AFIA

- Date : 06 octobre 2016
- Lieu : Salle A709, Université Paris-Dauphine, Pl. du M. de Lattre de Tassigny, 75016 Paris
- 14h00 « De la résolution de problèmes NP-difficiles en intelligence Artificielle : optimisation, modèles graphiques et applications » par Thomas Schiex (INRA, Toulouse)  
*Depuis vingt ans, nous avons développé un ensemble d'outils algorithmiques ciblés sur la résolution de problèmes d'optimisation énoncés sous la forme de modèle graphique. Ces modèles sont utilisés aussi bien dans le cadre de l'apprentissage automatique que de la programmation par contraintes. Je détaillerai les principes fondamentaux de ces outils algorithmiques et comment nous avons pu les mobiliser pour répondre à différentes questions en biologie, en particulier grâce à Toulbar2.*
- 15h00 Assemblée Générale Ordinaire 2016 de l'AFIA
- 18h00 Cocktail



**AFIA**

Association française  
pour l'Intelligence Artificielle

# FIIA 2017

Lors de FIIA 2016, il est apparu que la gestion / l'ingénierie des CONNAISSANCES était essentielle pour soutenir tout système d'IA et pouvait de toute façon participer au déploiement d'applications. Ce deuxième FORUM INDUSTRIEL DE L'INTELLIGENCE ARTIFICIELLE est donc consacré à ce sujet, à travers différents thèmes. Chaque thème est composé de courtes présentations suivies d'une mini-table ronde dont l'objectif est d'aboutir à des éléments de feuille de route et de permettre des échanges accrus entre académiques et industriels. Des appels plus généraux concernant l'IA seront également annoncés. Une dernière session est réservée à des présentations rapides par des industriels invités l'ayant souhaité.

## Date et Lieu

- Date 27 avril 2017
- Lieu Université Paris-Descartes, Amphithéâtre Fourier, 45 rue des Saints Pères, 75006 Paris

## Programme

- 08h30 Accueil
- 08h45 Ouverture de Frédéric Dardel (*Président de l'Université Paris Descartes*) et Pavlos Moraitis (*Directeur du LIPADE*), Introduction par Catherine Faron-Zucker (*Coordinatrice du Collège Science de l'Ingénierie des Connaissances de l'AFIA*) et Bruno Patin (*Coordinateur du Collège Industriel de l'AFIA*).
- 09h00 Présentation des travaux de l'OPECST sur l'IA par Dominique Gillot (*Sénatrice*).
- 09h15 Thème « Gestion vs. Ingénierie » de(s) la connaissance(s), avec Jean Charlet (*APH Paris & INSERM*), Alain Berger (*Ardans*), et Nicolas Dubuc (*Michelin*).
- 10h00 Thème « Agents conversationnels », avec Frédérique Segond (*Viseo*), Henri Sanson (*Orange*), Cyril Texier (*DoYouDreamUp*) et Nicolas Sabouret (*Univ. Paris 11*).
- 11h00 Pause-café
- 11h30 Thème « Prise de décision », avec Bruno Patin (*Dassault Aviation*), Ph. Bonnard (*IBM*), J.-Marc David (*Renault*), Laurent Gouzènes (*Pacte Novation*) et J.-Fabrice Lebraty (*Univ. Lyon 3*).
- 12h30 Buffet
- 14h00 Thème « Transparence et confiance », avec Patrick Constant (*Owant-Pertimm*), J.-Pierre Cotton (*Ardans*), Victor Dermiaux (*CNIL*) et J.-Gabriel Ganascia (*Univ. Paris 6*).
- 15h00 « Les projets d'IA s'annoncent » avec Yves Demazeau (*AFIA*), Fabienne Réveillac (*SNCF*) et Pascal Ollivier (*SOGET*).
- 15h30 Pause-café
- 16h00 Succession de présentations rapides de sociétés concernées par l'IA en « Trois planches »
- 17h00 Conclusion par Yves Demazeau (*Président de l'AFIA*)
- 17h30 Clôture

## Organisation

Cet événement est organisé par Alain Berger pour le Collège Industriel de l'AFIA et Catherine-Faron-Zucker pour le Collège Science de l'Ingénierie des Connaissances de l'AFIA.

## Inscriptions

Les inscriptions au Forum (gratuites pour les membres de l'AFIA, 30 € sinon) sont obligatoires et à effectuer sur le site des inscriptions : <http://inscriptions.afia.asso.fr/>. Le déjeuner est offert à toutes les personnes qui se seront inscrites à FIIA 2017 avant le 24/04/17, et seulement à celles-ci.



**AfIA**  
Association française  
pour l'Intelligence Artificielle

**IJCAI-17**  
MELBOURNE

## France@IJCAI 2017

Cette journée France@IJCAI 2017 propose un avant-goût des recherches françaises qui seront présentées à IJCAI 2017 à Melbourne. Les 42 chercheurs de France qui ont un article accepté à IJCAI 2017 ont été invités à venir le présenter lors de cette journée. Tous n'ont malheureusement pas pu se libérer, mais 7 ont pu répondre positivement à cette invitation.

Cette journée est inscrite au sein de la Plate-Forme Intelligence Artificielle 2017.

### Date et Lieu

- Date 07 juillet 2017
- Lieu Université de Caen Normandie, Campus 2, Côte de Nacre, bd. Maréchal Juin, 14032 Caen

### Programme

- 10h15 Présentation de l'Association Française pour l'Intelligence Artificielle, par Yves Demazeau (*Président de l'AFIA*)
- 10h30 « A Recursive Shortcut for CEGAR : Application to The Modal Logic K Satisfiability Problem » par Jean-Marie Lagniez, Daniel Le Berre, Tiago de Lima, et Valentin Montmirail (*Université d'Artois*)
- 10h45 « Constraint Games Revisited » par Anthony Palmieri et Arnaud Lallouet (*Huawei Boulogne-Billancourt*)
- 11h00 « An Improved Decision-DNNF Compiler » par Pierre Marquis (*Université d'Artois*) et Jean-Marie Lagniez.
- 11h15 « Temporal Sequences of Qualitative Information : Reasoning about the Topology of Constant-Size Moving Regions » par Quentin Cohen-Solal (*Université de Caen Normandie*), Maroua Bouzid et Alexandre Niveau
- 11h30 « Strategically Knowing How » par Andreas Herzig (*CNRS Toulouse*)
- 11h45 « End-to-end optimization of goal-driven and multi-modal dialogue systems » par Florian Strub, Jeremie Mary (*Université Lille 3*), Bilal Piot et Olivier Pietquin
- 12h00 « Belief Manipulation Through Propositional Announcements » par Aaron Hunter, Francois Schwarzentruher (*ENS Rennes*) et Eric Tsang
- 12h15 Clôture

### Organisation

Cet événement est organisé par Christian Bessière et Serena Villata pour l'Association Française pour l'Intelligence Artificielle (AFIA).

### Inscriptions

Les inscriptions se font sur le site de PFIA 2017 : <https://pfia2017.greyc.fr/inscriptions> L'inscription à la Journée inclut la participation au repas et aux pause-café.





**AfIA**

Association française  
pour l'Intelligence Artificielle

# Nuit de l'Info 2015

Un défi pour la Nuit de l'info est une question ou un problème à relever lors de la réalisation d'une application dont on ne connaît le thème que le jour même (ou plutôt la nuit) de la "Nuit de l'info".

## La Compétition

Mettez de l'intelligence artificielle 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'IA dans son projet remportera ce défi.

La récompense est un chèque de 1000 € assorti d'une invitation à présenter une démonstration lors de la compétition "IA sur Robots" organisée par l'AFIA au sein du congrès RFIA 2016 à Clermont-Ferrand.

## Les Equipes

Le défi a été relevé durant la nuit de l'info du 3 au 4 décembre 2015, de 16h40 à 08h02. Parmi les 25 équipes qui s'étaient inscrites pour relever ce défi, 9 équipes ont envoyé des réponses : "#Les Tarkiflettes", "And his name is [JOHNCENA]", "Douglas Gang", "GRO-KLC", "Les 73", "Pony God Master", "Segmintation Fault", "Syndicat Lunatique des Informaticiens Paresseux", et "Team du JEEJ".

## Le Jury

Le jury, composé de Carole Adam, Patrick Albert, Cindy Even, Thomas Guyet, Philippe Morignot, Bruno Patin, Sylvie Saget et Florence Bannay (coordinatrice), s'est réuni par visio-conférence le 4 décembre de 9h à 11h.

## Les Résultats

La première place et la récompense furent attribuées à l'équipe "Team du JEEJ" dont l'application propose d'attribuer un centre de refuge à une personne dans le besoin ou à des bénévoles tels que les médecins disponibles pour agir sur place. Les techniques utilisées sont la classification supervisée de données, la mise en œuvre d'un arbre de décision, l'analyse de la tonalité d'une phrase, et le calcul d'itinéraires. Le jury a tenu à souligner la qualité de ce projet tant sur le plan de la largeur du spectre couvert en IA que sur son applicabilité puisque le projet tourne sur différentes plateformes grâce aux techniques de responsive design. Le jury a également apprécié la qualité des explications fournies et la bonne utilisation de bibliothèques standard existantes.

Le jury a classé en deuxième position l'équipe "Douglas Gang" qui a proposé une application capable de donner des conseils en réponse à des demandes orales. Cette application se base sur des techniques de reconnaissance de la parole et d'apprentissage automatique puisque les conseils appropriés auront été préalablement recueillis par apprentissage.

La troisième place a été attribuée à "Syndicat Lunatique des Informaticiens Paresseux" pour leur proposition d'une intelligence artificielle incarnée dans un robot humanoïde NAO. Leur application se base sur des techniques de reconnaissance de visage et reconnaissance de la parole. Elle permet d'envoyer des messages à des personnes appropriées et d'interroger oralement NAO afin d'obtenir les numéros de téléphone des services d'urgences.



**Afia**

Association française  
pour l'Intelligence Artificielle

## Appel à participation

### Compétition *Intelligence Artificielle sur Robot*

**Lundi 27 & Mardi 28 juin 2016, Clermont-Ferrand**

L'Association Française pour l'Intelligence Artificielle ([AFIA](#)) organise une compétition Intelligence Artificielle sur Robot lors de la conférence [RFIA'16](#). Si une compétition traditionnelle de robotique (e.g., RoboCup, [CAROTTE](#)) favorise l'intégration d'algorithmes liés aux capteurs et actuateurs au sein d'une plateforme robotique (e.g., perception, SLAM, contrôle), la présente compétition se focalise sur l'intelligence des robots : il s'agit d'intégrer au sein d'une plateforme des algorithmes de haut niveau mettant en évidence l'intelligence embarquée des robots.

La plateforme imposée est celle des robots [LEGO MINDSTORM](#) (NXT ou EV3). Un ou plusieurs robots par équipe sont possibles. L'arène est de dimensions 10m x 5m avec un sol plan quadrillé. Elle contient des obstacles et des murs intérieurs dont les positions sont variables et inconnues des compétiteurs avant chaque épreuve. Les [règles de la compétition](#) sont adaptées du jeu « [Robot Ricochet](#) » : à partir d'un point d'entrée dans l'arène, atteindre une destination en rebondissant sur des obstacles statiques (e.g., parpaings, murs) ou dynamiques (autres robots) présents dans certaines cellules de l'arène. Plusieurs niveaux de difficulté sont possibles : selon le nombre d'équipes dans l'arène (1 ou 2) et les informations sur la destination (donnée ou informations partielles). Dans ce dernier cas, le(s) robot(s) de chaque équipe doi(ven)t d'abord communiquer avec le(s) robot(s) de l'équipe adverse pour connaître leur destination cible, avant de chercher à atteindre cette destination par rebonds sur des obstacles.

Cette compétition est ouverte à toute personne intéressée par allier Intelligence Artificielle et robotique : chercheurs, étudiants, académiques, lycéens, ingénieurs, amateurs, industriels, et grand public.

Information et inscription : [site web de l'AFIA](#)

Comité d'organisation : Carole Adam, Florence Bannay, Cédric Buche, Christophe Debain, Cindy Even, Philippe Morignot, Sylvie Saget.

Contact et questions : [pmorignot@yahoo.fr](mailto:pmorignot@yahoo.fr)







**Afia**

Association française  
pour l'Intelligence Artificielle

# Nuit de l'Info 2016

La Nuit de l'Info 2016 a eu lieu dans la nuit du 1er au 2 décembre sur le thème « Venir en aide aux réfugiés ». Le principe de la Nuit de l'Info est très simple : Le jeudi 2, au coucher du Soleil, 16h40, en séance plénière : les organisateurs remettent un sujet (le même pour toute la France) aux participants. Les étudiants s'organisent en groupes (sur un ou plusieurs sites) : ils développent un projet (informatique, marketing, rédactionnel, ...) tout en ciblant un ou plusieurs défis. Le vendredi 3, au lever du Soleil, 08h00, les développements sont figés. Le vendredi matin, pendant que les étudiants dorment des jurys se réunissent (un jury par défi) et examinent les travaux réalisés par les différentes équipes durant la nuit.

## La Compétition

Le défi qu'a proposé l'AFIA s'intitulait « Mettez de l'intelligence dans votre moteur » : « Vous mettez en œuvre une ou plusieurs méthodes d'Intelligence Artificielle (IA) dans votre projet et vous indiquez en quoi ces méthodes rendent votre réalisation plus performante ou pertinente. Vous pouvez faire appel à des techniques classiques, en cours de développement ou futuristes. L'équipe qui aura mis le plus en avant les avantages de l'utilisation de l'IA dans son projet remportera ce défi ». Les réponses devaient inclure :

- Une description des problématiques IA rencontrées dans le projet,
- Une explication de l'intégration de l'IA (effective ou imaginaire) dans l'application
- Une analyse des avantages et inconvénients de cette intégration

La récompense était un chèque de 1000 € assorti d'une invitation à présenter une démonstration lors de la compétition "IA et Jeux Vidéo" organisée par l'AFIA au sein de la Plate-Forme IA 2017 à Caen.

## Les Equipes

Parmi les 19 équipes inscrites pour relever ce défi, 4 équipes ont réussi à remettre un projet ans les temps impartis : « root@pls:~> sudo nano & », « ASpico », « Les fortran chocolat », et « Le Coin C'Net ».

## Le Jury

Le jury était composé en 2017 de Carole Adam, Cindy Even, Thomas Guyet, et Florence Bannay (coordinatrice), s'est réuni par visio-conférence le 4 décembre de 9h à 11h. Les projets ont été jugés selon les critères: applicabilité / mise en œuvre, innovation, IA (évaluation qualitative et quantitative : richesse de la solution proposée, spectre couvert), et qualité des explications

## Les Résultats

La première place et la récompense ont été attribuées à l'équipe « root@pls:~> sudo nano & » dont l'application « Botcher » est bien adaptée au sujet national, elle fournit un chatbot capable de donner des informations aux réfugiés dans plusieurs langues. Elle fonctionne sur mobile, exploite des techniques d'IA existantes (mais sans apporter d'innovation). Les explications de cette solution sont bien rédigées et claires.

Le jury a classé en deuxième position l'équipe « Le Coin C'est net » qui propose l'utilisation d'un robot Nao pour répondre à des questions, mémoriser le visage, le nom et la langue de l'utilisateur afin de personnaliser les interactions. Nous avons noté également que cette application n'est pas bien adaptée au thème de cette année, un robot Nao paraît trop cher pour des réfugiés. De plus nous ne pouvons pas tester le code proposé sans disposer d'un robot Nao.

La troisième place a été attribuée à l'équipe « Les fortran chocolat » pour leur application qui contient un jeu vidéo destiné à amuser les réfugiés. Notons que ce jeu vidéo utilise des techniques d'IA mais ne répond pas très bien au sujet national de la nuit de l'informatique.



## COMPÉTITION

& Intelligence Artificielle  
Jeux Vidéo

Lors de **PFIA'17**

à Caen, entre le 3 et le 7 Juillet 2017

Pour gagner, **DÉVELOPPEZ LE BOT** le plus **HUMANLIKE\***!

Plus d'info : <http://goo.gl/rVZ0Dq>

Contact : [botcontest2017@enib.fr](mailto:botcontest2017@enib.fr)



AfIA

Ne pas jeter sur la voie publique. Imprimé par l'atelier de reprographie de l'UBO. \*Semblable à un humain.



**AfIA**  
Association française  
pour l'Intelligence Artificielle



# RO&IA 2015

Cette troisième journée RECHERCHE OPERATIONNELLE ET INTELLIGENCE ARTIFICIELLE se déroule le 23 Septembre 2015 à Paris. La journée est organisée conjointement par 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 les thématiques liées à la PROGRAMMATION PAR CONTRAINTES.

## Date et Lieu

- Date 23 septembre 2015
- Lieu CNAM, Amphi Friedmann, Accès 33, 292 rue Saint-Martin, 75003 Paris

## Programme

- 09h40 Présentation de l'Association Française pour l'Intelligence Artificielle, par Yves Demazeau (*Président de l'AFIA*) et de la Société Française de Recherche Opérationnelle et Aide à la Décision, par Frédéric Gardi (*Président de la ROADEF*)
- 10h00 Tutoriel « Optimisation dans les modèles graphiques et applications » par Thomas Schiex (*INRA Toulouse*)
- 12h00 Pause déjeuner
- 14h00 « The NValue global constraint » par Hadrien Cambazard (*Grenoble INP*)
- 14h40 « Heuristiques de recherche en programmation par contraintes » par Jean-Guillaume Fagès (*Cosling, Nantes*)
- 15h20 Pause-café
- 15h40 « Solving industrial scheduling problems with constraint programming » par Philippe Laborie (*IBM, Paris*)
- 16h20 « Optimal routing in deterministic delay-tolerant networks » par Ronan Bocquillon (*UTC Compiègne*)
- 17h00 Clôture

## Organisation

La journée est organisée par Nicolas Maudet (Univ. Paris 6) pour l'Association Française pour l'Intelligence Artificielle et Amélie Lambert (CNAM) pour la Société Française de Recherche Opérationnelle et d'Aide à la Décision.

## Inscriptions

L'inscription à la journée est gratuite et inclut la participation aux pause-café. L'organisation et la prise en charge du déjeuner restent à la discrétion des participants. Pour des raisons d'organisation, merci à toute personne intéressée de bien vouloir s'inscrire sur le site : <http://www.lamsade.dauphine.fr/~jfro/Inscription/>



# ST&IA 2015

Cet événement commun organisé par l'AST et l'AFIA a pour objectif d'échanger entre scientifiques et grand public autour la mise en scène de l'Intelligence Artificielle (IA) dans les productions médiatiques, en particulier cinématographiques. L'événement est composé de débats qui suivront la projection de films grand public dont la séquence forme une rétrospective de l'IA, de sa naissance à ses perspectives actuelles.

## Date et Lieu

- Date 02-05 octobre 2015
- Lieu Cinéma Grand Action, Salle 4, 5 rue des Ecoles, 75005 Paris

## Programme

02 octobre 2015

- 19h00 – 21h30 « [The Imitation Game](#) » (2015) de Morten Tyldum.

Alain Turing, mathématicien et cryptanalyste, a eu un rôle important lors de la Seconde Guerre mondiale. Ce film revient sur le destin de cet homme qui a accéléré la chute du nazisme et qui a posé les bases de l'IA. Le réalisateur prend quelques libertés par rapport à la réalité, mais, tout comme en IA, laisse le soin au spectateur de prêter une intelligence à la machine observée. *Avec, pour le débat, Jacques Pitrat, Chercheur en IA (Retraité du CNRS, Paris) et Giuseppe Longo, Logicien et Epistémologiste (Professeur à l'ENS Paris).*

03 octobre 2015

- 18h00 – 21h00 « [2001 : l'Odysée de l'Espace](#) » (1968) de Stanley Kubrick

Embarquez pour la mission Jupiter aux côtés des docteurs Bowman, Poole et du robot intelligent HAL 9000 ! Les liens entre science-fiction et IA ont toujours été très étroits, même si la machine ne pense pas. Le spectateur mesure l'ampleur du chemin parcouru par l'IA en 50 ans et s'interroge la capacité des machines de demain. *Avec, pour le débat, Jean-Gabriel Ganascia, Chercheur en IA et Philosophe (Professeur à l'Université de Paris 6) et Pierre Mounier-Kuhn, Historien (Chargé de Recherche CNRS, Paris).*

04 octobre 2015

- 18h00 – 21h00 « [A.I. Intelligence Artificielle](#) » (2001) de Steven Spielberg

Un robot peut-il devenir un vrai petit garçon ? Au-delà de la mise en scène de l'IA, c'est la question de la place de l'être humain dans son environnement qui est posée. Si l'homme et la machine restent de natures différentes, les systèmes intelligents qui sont construits doivent s'adapter à leur environnement, souvent peuplé d'êtres humains. *Avec, pour le débat, Philippe Mathieu, Chercheur en IA (Professeur à l'Université de Lille 1) et Serge Tisseron, Psychiatre et Psychanalyste (Professeur Université Paris 7).*

05 octobre 2015

- 19h00 – 21h00 « [Eva](#) » (2011) de Kikke Maillo.

Alex, un ingénieur de renom est rappelé par la Faculté de Robotique pour construire un enfant robot. Ce dernier, choisit de prendre comme modèle sa nièce, Eva. L'IA s'est historiquement développée en étant centrée plus sur l'individu que sur le collectif. Les travaux récents en IA essaient de mieux ancrer les robots et les êtres humains dans le collectif qu'ils constituent ensemble. *Avec, pour le débat, Yves Demazeau, Chercheur en IA (Directeur de Recherche au CNRS, Grenoble) et Raja Chatila, Roboticien (Directeur de Recherche au CNRS, Paris).*

## Organisation

Cet événement est organisé par Yves Demazeau pour l'Association Française pour l'Intelligence Artificielle (AFIA) et Marina Vasseur pour l'Association Science et Télévision (AST).

## Inscriptions

Les inscriptions se font sur le site du Festival Pariscience : <http://www.pariscience.fr/fr/program/>

**GRAND ACTION**



**Afia**  
Association française  
pour l'Intelligence Artificielle



# MACS&IA 2015

L'Association Française pour l'Intelligence Artificielle (AFIA) et le groupe de travail Modélisation et pilotage des systèmes de Connaissances et de Compétences dans les Entreprises Industrielles (C2EI) du GDR Modélisation, Analyse et Conduite des Systèmes dynamiques (MACS) organisent une deuxième journée commune FORMALISATION DES CONNAISSANCES ET IA.

Cette journée est fondée sur des interventions se situant à l'intersection de l'IA, du domaine de la formalisation et de la gestion du retour d'expériences. Les interventions portent sur des problématiques industrielles et sociétales, des états de l'art, des avancées scientifiques et des perspectives de recherche. Un focus particulier est porté sur des approches fondées sur des traces d'activités.

## Date et Lieu

- Date 4 novembre 2015
- Lieu Université Paris-Descartes, Salle Turing, 45 rue des Saints Pères, 75006 Paris

## Programme

- 09h15 Présentation de l'Association Française pour l'Intelligence Artificielle, par Yves Demazeau (*Président de l'AFIA*) et du GT C2EI du GDR Modélisation, Analyse et Conduite des Systèmes dynamiques, par Eric Bonjour (*Animateur du GT C2EI*)
- 09h35 Présentation de la journée par Davy Monticolo (*AFIA*) et Eric Bonjour (*GDR MACS*)
- 09h45 Dynamique des connaissances et expérience tracée, par Béatrice Fuchs (*Université Lyon 1*)
- 10h30 L'observation dans les EIAH : le cas des ENT, des microsondes et des jeux sérieux, par Thibault Carron (*Université de Savoie*)
- 11h15 Comment exploiter le retour d'expérience issu de la résolution de problèmes en entreprises ? par Laurent Geneste (*Ecole Nationale d'Ingénieurs de Tarbes*)
- 12h00 Pause déjeuner
- 13h30 Traçabilité des connaissances de projets, par Nada Matta (*Université de Technologie de Troyes*)
- 14h15 Capitalisation de traces d'activité pour la reconstitution de démarches métier, par Eric Bonjour (*Université de Lorraine*) et Davy Monticolo (*Université de Lorraine*)
- 15h00 Gestion des connaissances patrimoniales : une expansion des humanités numériques pour le patrimoine industriel, par Florent Laroche (*Université de Nantes*)
- 15h45 Extraction de comportements récurrents dans des traces d'interaction, par Alexandre Pauchet (*INSA Rouen*)
- 16h30 Clôture

## Organisation

Cet événement est organisé par Davy Monticolo pour l'Association Française pour l'Intelligence Artificielle (AFIA) et Eric Bonjour pour le GdR Modélisation, Analyse et Conduite des Systèmes Dynamiques (MACS).

## Inscriptions

L'inscription est gratuite mais obligatoire, à prendre auprès de [Davy.Monticolo@univ-lorraine.fr](mailto:Davy.Monticolo@univ-lorraine.fr). Le déjeuner est à la charge des participants. De nombreux restaurants se trouvent à proximité de la salle Turing.



**AFIA**  
Association française  
pour l'Intelligence Artificielle



# RI&IA 2015

La recherche d'information (RI), dans la mesure où elle vise à réaliser des tâches relevant de la cognition humaine, a de fait des liens avec l'intelligence artificielle (IA).

L'objectif de cette première journée RECHERCHE D'INFORMATION et INTELLIGENCE ARTIFICIELLE est de mettre en évidence les liens qui existent entre ces deux domaines, et peuvent aider à mettre au point des méthodes qui font avancer l'analyse et la représentation des documents, de quelque type qu'il soit (texte, image, audio, vidéo), ainsi que la modélisation de leur interrogation ou leur exploration.

## Date et Lieu

- Date 01 décembre 2015
- Lieu Délégation CNRS d'Ivry, 27 rue Paul Bert, 94200 Ivry sur Seine

## Programme

- 09h00 Présentation de l'Association Française pour l'Intelligence Artificielle, par Yves Demazeau (*Président de l'AFIA*) et de l'Association francophone de Recherche d'Information et Applications, par Max Chevalier (*Président de l'ARIA*),
- 09h10 Présentation de la journée par Jean-Pierre Chevallet (ARIA) et Brigitte Grau (AFIA).
- 09h15 Recherche d'Information : Etat des lieux, par Mohand Boughanem (*Université Toulouse 3*)
- 10h10 Pause
- 10h30 L'apprentissage en Recherche d'Information par Eric Gaussier (*Université Grenoble 1*)
- 11h25 Structuration des Flux TV : Etat de l'Art par Mohamed Bouaziz (*Université d'Avignon*)
- 12h00 Repas
- 13h30 Concept Adjustment for Constructing Description Logic based Ontologies par Yue Ma (*Université Paris 11*)
- 14h25 Recherche d'Images par Hervé Le Borgne (*CEA Saclay*)
- 15h20 Contexte dans les modèles de langue neuronaux par Benjamin Piwowarski (*CNRS Paris*)
- 16h15 Pause
- 16h35 Analyse de graphes de terrain pour la recherche d'information par Rushed Kanawati (*Université Paris 13*)
- 17h30 FolkRec: un système personnalisé de recommandations pour les folksonomies basé sur les concepts quadratiques par Mohamed Nader Jelassi (*Université Clermont-Ferrand 2*)
- 18h00 Clôture

## Organisation

Cet événement est organisé par Brigitte Grau pour l'Association Française pour l'Intelligence Artificielle (AFIA) et Jean-Pierre Chevallet pour l'Association francophone de Recherche d'Informations et Applications (ARIA).

## Inscriptions

L'inscription à la journée est gratuite mais obligatoire, à réaliser sur la page : <https://www.weezevent.com/journee-commune-ia-et-ri>. L'inscription ne comprend pas le repas de midi.





# EGC&IA 2016

Cette deuxième journée EXTRACTION ET GESTION DES CONNAISSANCES et INTELLIGENCE ARTIFICIELLE réunit les deux communautés autour du thème des « Données Participatives et Sociales ». Ces données sont au cœur de nouveaux défis tant au niveau de la fouille de données que de l'IA.

Les travaux de la littérature sont généralement associés à l'une des deux communautés, sans montrer le lien entre elles. Cet atelier cherche à focaliser sur ce lien du point de vue représentation qu'analyse.

## Date et Lieu

- Date 19 janvier 2016
- Lieu IUT de Reims-Chalons-Charleville, Chemin des Rouliers, 51100 Reims

## Programme

- 09h00 Présentation de l'Association Française pour l'Intelligence Artificielle, par Yves Demazeau (*Président de l'AFIA*) et de l'Association Extraction et Gestion des Connaissances, par Fabrice Guillet (*Président de EGC*),
- 08h45 Présentation de la journée par Arnaud Martin (*Univ. Rennes 1*) et Engelbert Mephu Nguifo (*Univ. Clermont-Ferrand 2*).
- 09h00 « A la recherche des mini-publics : problème de communautés, de singularités de sémantique » par Eric Leclercq (*Univ. Bourgogne, Dijon*), S. Kirgizov et M. Danisch.
- 09h30 « Analyse d'activité et exposition de la vie privée sur les médias sociaux » par Younes Abid (*Inria, Nancy*), A. Imine, A. Napoli, C. Raïssi, M. Rigolot, et M. Rusinowitch.
- 10h00 « Perception de l'usage de matériaux composites pour les véhicules du futur » par Marouane Hachicha (*Univ. Nantes*), N. Cullot, E. Leclercq, Ph. Castel, M.-F. Lacassagne et S Fontaine.
- 11h00 « Analyse des images qui circulent sur Internet : un aperçu du projet ImagiWeb » par Julien Velcin (*Université Lyon 2*).
- 11h30 « Catégorisation et Désambiguïsation des Intérêts des Individus dans le Web Social » par Coriane Nana Jipmo (*Centrale Supélec, Paris*), G. Quercini et N. Bennacer.
- 12h00 « Un système personnalisé de recommandation basé sur le profil des utilisateurs dans les folksonomies » par Mohamed Nader Jelassi (*Univ. Clermont-Ferrand 2*), S. Ben Yahia, et E. Mephu Nguifo.
- 12h30 Pause-repas
- 14h00 Présentations de l'AFIA par Yves Demazeau (*Président de l'AFIA*) et Fabrice Guillet (*Président de EGC*).
- 14h30 « Fabrique logicielle de réseaux sociaux spécialisés : Aspects fonctionnels » par David Fernandez (*Université Montpellier 2*), B. Billet et D. Parigot.
- 15h00 « Worker-Centricity Could Be Today's Disruptive Innovation in Crowdsourcing » par Sihem Amer-Yahia (*CNRS Grenoble*) (conférence invitée).
- 16h30 « Connaissance et collaboration » par Jérôme Euzenat (*Inria Grenoble*) (conférence invitée).
- 18h00 Clôture

## Organisation

Cet événement est organisé par Engelbert M. Nguifo pour l'Association Française pour l'Intelligence Artificielle (AFIA) et Arnaud Martin pour l'Association Extraction et Gestion des Connaissances (EGC).

## Inscriptions

L'atelier se tient dans le cadre de la conférence [EGC 2016](http://egc2016.univ-reims.fr/index.php/Inscription). Les inscriptions se font sur le site de la conférence (<http://egc2016.univ-reims.fr/index.php/Inscription>).





**AFIA**

Association française  
pour l'Intelligence Artificielle



**AFRV**

ASSOCIATION FRANÇAISE  
DE RÉALITÉ VIRTUELLE AUGMENTÉE,  
MIXTE ET D'INTERACTION 3D

# RV&IA 2016

L'objectif de cette deuxième journée RÉALITÉ VIRTUELLE et INTELLIGENCE ARTIFICIELLE est de mettre en évidence les liens entre ces deux domaines.

La matinée est consacrée à la thématique de la prise de décision pour les comportements humanoïdes virtuelles. L'après-midi se focalise sur les mécanismes intelligents au sein des environnements virtuels d'apprentissage humain.

## Date et Lieu

- Date 02 février 2016
- Lieu LIP6-UPMC, Tour 25-26, salle 105, 4 place Jussieu, 75005 Paris

## Programme

- 10h00 Présentation de l'Association Française pour l'Intelligence Artificielle par Yves Demazeau (*Président de l'AFIA*) et de l'Association Française de Réalité Virtuelle (AFRV) par Matthieu Lépine (*président de l'AFRV*)
- 10h40 « Modélisation et simulation des comportements humains, utilisations dans les domaines de la sécurité civile ou du militaire » par Caroline Chopinaud (*MAS Group, Paris*)
- 11h15 Pause-café
- 11h30 « Comportement adaptatif de personnages de jeux vidéo. Le modèle CHAMELEON » par Cédric Buche (*ENIB, Brest*)
- 12h05 « Évaluation de la crédibilité des comportements d'entités virtuelles » par Cindy Even (*Virtualys, Brest*)
- 12h40 Pause-repas
- 14h15 « Scénarisation d'environnements virtuels pour la formation en environnements sociotechniques complexes » par Domitile Lourdeaux (*UTC, Compiègne*)
- 14h50 « Pedagogical Collaborative Conversational Agent Architecture for Virtual Reality » par Mukesh Barange (*INSA, Rouen*), Julien Saunier et Ronan Querrec
- 15h25 Pause-café
- 15h40 « Apprentissage de la calligraphie » par Rémy Frenoy (*UTC, Compiègne*), Indira Thouvenin, et Olivier Gapenne
- 16h15 « Versatile Scenario Guidance for Collaborative Virtual Environments » par Guillaume Claude (*Inria Rennes Bretagne Atlantique*)
- 17h00 Clôture

## Organisation

Cet événement est organisé par Cédric Buche pour l'Association Française pour l'Intelligence Artificielle (AFIA) et Ronan Querrec pour l'Association Française de Réalité Virtuelle (AFRV).

## Inscriptions

L'inscription est gratuite mais obligatoire, à réaliser sur: <http://doodle.com/poll/w92za5rn6895u3gh>. Le déjeuner est à la charge des participants.







**AFIA**  
Association française  
pour l'Intelligence Artificielle



Association  
pour le Traitement  
Automatique  
des Langues

# TAL&IA 2016

L'objectif de cette troisième journée TRAITEMENT AUTOMATIQUE DES LANGUES et INTELLIGENCE ARTIFICIELLE est de renforcer les liens entre ces deux domaines, avec un accent sur la thématique des représentations utilisées en TAL et la façon de les acquérir et de les manipuler.

Quels types de représentations sous-tendent les modèles aux niveaux lexical, morphologique, syntaxique, sémantique, quels raisonnements vont de pair avec eux : unification, similarité, implication ? Dans cette perspective comment sont utilisés les modèles et calculs de l'IA et quelles pistes sont prometteuses pour le TAL ?

## Date et Lieu

- Date 24 mars 2016
- Lieu Amphi Hermite, Institut Henri Poincaré, 11 rue Pierre et Marie Curie, 75005 Paris

## Programme

- 09h00 Présentation de l'Association Française pour l'Intelligence Artificielle, par Yves Demazeau (*Président de l'AFIA*), de l'Association pour le Traitement Automatique des Langues, par Patrick Paroubek (*Président de l'ATALA*)
- 09h20 « Représentations lexicales pour l'analyse sémantique : leçons tirées de l'expérience d'un FrameNet pour le français » par Marie Candito (*Université Paris 7*)
- 10h00 « Plongements lexicaux pour la sémantique » par Tim van de Cruys (*CNRS Toulouse*)
- 10h40 Pause
- 11h10 « Modélisation de la langue des signes et représentation sémantique » par Michael Filhol (*CNRS Orsay*)
- 11h50 « Représentation de la dynamique du discours » par Sylvain Pogodalla (*Inria Nancy Grand Est*)
- 12h30 Déjeuner libre
- 14h00 « Apport du modèle BDI Croyances/Désirs/Intentions pour la représentation des opinions, sentiments et émotions » par Patrick Paroubek (*CNRS Orsay*)
- 14h40 « Analyse de Concepts Logiques (LCA) et Traitement Automatique des Langues (TAL) » par Annie Forêt (*Université Rennes 1*)
- 15h20 Pause
- 15h50 « Entités nommées : représentation et structuration » par Sophie Rosset (*CNRS Orsay*)
- 16h30 Table ronde sur les représentations et raisonnements à la frontière du TAL et de l'IA
- 17h30 Clôture

## Organisation

Cet événement est organisé par Philippe Muller pour l'Association Française pour l'Intelligence Artificielle (AFIA) et Yannick Toussaint pour l'Association pour le Traitement Automatique des Langues (ATALA).

## Inscriptions

La participation à la journée est gratuite et ouverte à tous. Pour des raisons de logistique, nous vous demandons néanmoins de vous inscrire auprès de [philippe.muller@irit.fr](mailto:philippe.muller@irit.fr)





**AfIA**  
Association française  
pour l'Intelligence Artificielle



# IM&IA 2016

Cette deuxième journée INFORMATIQUE MEDICALE et INTELLIGENCE ARTIFICIELLE fait se rencontrer les deux disciplines autour du thème de l'ingénierie des connaissances appliquée à l'exploitation des données médicales, dont le nombre n'a cessé d'augmenter au cours de ces dernières années.

L'objectif de la journée est de débattre des verrous liés à la modélisation, la structuration, l'extraction, la recherche et la fouille de données de santé de tous types. Afin d'aborder ces questions, la journée s'appuie sur une conférence invitée et sur un ensemble de travaux qui témoignent de la vitalité et de la diversité de la recherche en France dans ce domaine.

## Date et Lieu

- Date 06 juin 2016
- Lieu Hôtel Mercure, 285 bd de l'aéroport – International-Antigone – 34000 Montpellier

## Programme

- 09h30 Présentations de l'Association Française pour l'Intelligence Artificielle, par Yves Demazeau (*Président de l'AFIA*) et de l'Association Française d'Informatique Médicale par Philippe Cinquin (*Président de l'AIM*).
- 10h00 « Clinical language and scientific language : linguistic contrasts and ontological similarities » par Kevin B. Cohen (Univ. du Colorado)
- 11h00 « Les ontologies pour aider à comprendre les parcours de santé dans le cadre des maladies neurodégénératives » par Sonia Cardoso (IHU-A-ICM Paris) et al.
- 11h25 « OMIConto : une ressource termino-ontologique pour la qualification et l'indexation des outils d'analyse en sciences omiques » Vincent Henry (OMICS Rouen), L. Soualmia et al.
- 11h50 « Approche sémantique pour automatiser le calcul des valeurs nutritionnelles d'une recette de cuisine » par Rabia Azzi (Univ. Paris 13) et al.
- 12h15 Déjeuner
- 14h00 « Tentatives de suicide, prédire la récurrence avec des techniques d'apprentissage statistique » par Philippe Combes (CH Montperrin Paris) et al.
- 14h25 « Reconnaissance des stades de sommeil à l'aide d'un outil de support à la décision basé sur les connaissances et la pratique des experts » par Adrien Ugon (Univ. Paris 6) et al.
- 14h50 « Extraction d'associations d'EIM à partir de dossiers patients : expérimentation avec les structures de patrons et les ontologies » par Gabin Personeni (Univ. Nancy) et al.
- 15h15 Pause
- 15h45 « MuEVo, un vocabulaire multi-expertise dédié au cancer du sein » par Solène Eholié (Univ. Montpellier) et al.
- 16h10 « Interopérabilité sémantique dans le domaine du diagnostic in » par Melissa Mary (Univ. Rouen) et al.
- 16h35 « Evaluation de la SNOMED CT comme support à l'alignement de terminologies diagnostiques en cancérologie » par Jean-Noël Nikiema (INSERM U1219 Bordeaux) et al.
- 17h00 Clôture

## Organisation

Cet événement est organisé par Fleur Mougin pour l'Association Française pour l'Intelligence Artificielle (AFIA) et Pierre Zweigenbaum pour l'Association Française d'Informatique Médicale (AIM).

## Inscriptions

Les inscriptions se font via le site de IC 2016 : <http://ic2016.sciencesconf.org/resource/page/id/11>





**AFIA**

Association française  
pour l'Intelligence Artificielle



# Rob&IA 2016

Le GdR Robotique et l'AFIA organisent leur deuxième journée commune ROBOTIQUE ET INTELLIGENCE ARTIFICIELLE. Cette journée repose sur des interventions invitées et un appel à exposés se situant à l'intersection des deux disciplines. Cette année, la session de l'après-midi se focalise plus particulièrement sur les systèmes multi-robots.

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. La journée se conclut par une table ronde sur l'identification des intérêts communs et sur les supports au développement des projets liant ces deux disciplines.

## Date et Lieu

- Date 27 juin 2016
- Lieu Pôle commun Polytech ISIMA, Campus des Cézeaux, 63177 Aubière

## Programme

- 08h40 Présentations de l'Association Française pour l'Intelligence Artificielle, par Yves Demazeau (*Président de l'AFIA*) et du GdR Robotique par Philippe Bonnifait (*Directeur du GdR*).
- 09h00 Présentation de la journée par Olivier Simonin (INSA Lyon) et Abdel-illah Mouaddib (*Université de Caen*).
- 09h10 « Adaptation Créative par Apprentissage » par Antoine Cully (*Université Paris 6*)
- 10h00 Pause
- 10h30 « Effective path planning based on expanding tree for navigation strategy of vehicle using waypoints » par José Vilca (*Université de Clermont-Ferrand*), L. Adouane et Y. Mezouar.
- 10h50 « Planification et navigation sociale » par Anne Spalanzani (*Université Grenoble Alpes*).
- 11h20 « Humanoid Robotics and human-robot interaction » par Serena Ivaldi (*INRIA Nancy*).
- 11h50 « Robotique et Ethique » par Raja Chatila (*CNRS Paris*).
- 12h20 Déjeuner
- 14h00 « Robotique évolutionnaire : état des lieux » par Nicolas Bredeche (*Université Paris 6*).
- 14h30 « Projet R-DISCOVER : robots d'exploration coopératifs » par Ouidad Labbani-Igbida (*Université de Limoges*).
- 15h00 « Dassault Aviation et la Robotique : un panorama » par Bruno Patin (*Dassault Aviation Saint Cloud*).
- 15h30 « Navigation autonome d'une flotte de véhicules communicants à partir d'une localisation précise et intègre » par Romuald Aufrère (*Université de Clermont-Ferrand*), G. Lelong, L. Delobel et R. Chapuis.
- 16h00 Pause
- 16h30 Table Ronde
- 18h00 Clôture

## Organisation

Cet événement est organisé par Olivier Simonin pour l'Association Française pour l'Intelligence Artificielle (AFIA) et Abdel-illah Mouaddib pour le Groument de Recherche (GdR) Robotique du CNRS.

## Inscriptions

Les inscriptions se font via le site: <http://rfia2016.iut-auvergne.com/index.php/organisation/inscriptions>





# RF&IA 2016

L'AFRIF et l'AFIA organisent leur première journée commune RECONNAISSANCE ET INTELLIGENCE ARTIFICIELLE. La journée est consacrée aux TRANSPORTS INTELLIGENTS.

## Date et Lieu

- Date : 28 juin 2016
- Lieu : Pole commun Polytech ISIMA, Campus des Cézeaux, 63177 Aubière

## Programme

- 08h40 Présentations de l'Association Française d'Intelligence Artificielle, par Yves Demazeau (*Président de l'AFIA*) et de l'AFRIF par Marie-Odile Berger (*Présidente de l'AFRIF*).
- 09h00 Présentation de la journée par Christine Solnon (*INSA Lyon*) et Patrick Rives (*Inria Rennes*).
- 09h10 « PSA sur la route du véhicule autonome » par Vincent Abadie (*PSA Paris*),
- 10h00 Pause
- 10h30 « Retour d'expérience après plus de 1000 Km en navette sans conducteur guidée par vision » par Eric Royer (*Université de Clermont-Ferrand*) et al.
- 11h00 « Cohérence entre la modélisation et les objectifs de contrôle pour les véhicules autonomes » par Philip Polack (*Mines ParisTech*) et al.
- 11h30 « Evidential grids with semantic lane information for intelligent vehicles » par Chunlei Yu (*Université de Technologie de Compiègne*) et al.
- 12h00 « Correction de nuages de points lidar embarqué sur véhicule pour la reconstruction d'environnement 3D vaste » par Pierre Merriaux (*ESIGELEC Rouen*) et al.
- 12h30 Déjeuner
- 14h00 « Évaluation de l'impact du taux de pénétration des systèmes coopératifs sur la consommation énergétique des véhicules » par Louahdi Khoudour (*Université Toulouse 3*) et al.
- 14h30 « Détection de vélos inutilisables grâce aux données ouvertes du système de vélos en libre-service Citibike » par Rémi Delassus (*Université de Bordeaux*) et al.
- 15h00 « Vers l'utilisation de graphes de liens causaux pour l'amélioration de la prévision court-terme du trafic routier » par Julien Salotti (*INSA Lyon*) et al.
- 15h20 « Réflexion sur la notion d'agent markovien dans des réseaux de transports » par Guillaume Lozenguez (*Ecole des Mines de Douai*) et al.
- 15h40 « Un algorithme multicritère pour l'optimisation des routes maritimes en temps réel » par Estelle Chauveau (*Université d'Aix-Marseille*) et al.
- 16h00 Pause
- 16h30 « Descripteurs d'images en contexte routier » et « Visual saliency on the road » par Pierre Duthon (*CEREMA Clermont-Ferrand*) et al.
- 17h00 « Localisation à base d'amers visuels » par Bahman Soheilian (*IGN Saint Mandé*) et al.
- 17h30 « Calibration d'un dispositif stereo-fisheye large baseline pour le diagnostic d'intersections routières » par René Datondji (*Université de Rouen*) et al.
- 18h00 Clôture

## Organisation

Cet événement est organisé par Christine Solnon pour l'Association Française d'Intelligence Artificielle (AFIA) et Patrick Rives pour l'Association Française pour la Reconnaissance et l'Interprétation des Formes (AFRIF).

## Inscriptions

Les inscriptions se font via le site: <http://rfia2016.iut-auvergne.com/index.php/organisation/inscriptions>





**Afia**

Association française  
pour l'Intelligence Artificielle



# Cla&IA 2016

La classification est au cœur des activités humaines. Lorsque l'on pense le monde, on l'organise et l'on classe les éléments qui le composent. Et, en tant que discipline s'étant initialement fixée pour objet les phénomènes constitutifs de l'intelligence humaine, l'IA s'est naturellement emparée de cette problématique et a produit de très nombreuses approches pour des problèmes variés. Selon les communautés, on distingue la classification de la discrimination, la classification non supervisée de la classification supervisée. La première s'ancre dans la longue tradition de la taxonomie. La seconde plus récente est associée à l'essor de l'apprentissage automatique. Mais aujourd'hui, avec le semi-supervisé, l'interactivité et les passages à l'échelle successifs, les questions posées s'entrecroisent et le dialogue entre les différentes approches semble plus que jamais nécessaire.

Cette journée CLASSIFICATION et INTELLIGENCE ARTIFICIELLE organisée conjointement par l'Association Française pour l'Intelligence Artificielle et la Société Francophone de Classification vise à faire le point sur des avancées récentes aux frontières des différentes communautés.

## Date et Lieu

- Date 16 septembre 2016
- Lieu Salle du Conseil Université Paris Descartes, 12 rue de l'Ecole de médecine, 75006 Paris

## Programme

- 00h00 Présentations de l'Association Française pour l'Intelligence Artificielle, par Yves Demazeau (*Président de l'AFIA*) et de la Société Francophone de Classification par Mohamed Nadif (*Président de la SFC*).
- 09h50 Présentation de la journée par Khalid Benabdeslem (*Université de Lyon*) et Pascale Kuntz (*Université de Nantes*).
- 10h00 « Classification non supervisée sous contraintes et programmation par contraintes » par Christel Vrain (*Université d'Orléans*).
- 11h00 « Classification et découverte de connaissances guidées sur l'analyse formelle de concepts » par Amedeo Napoli (*CNRS Nancy*).
- 12h00 Déjeuner libre
- 13h30 « Apprentissage statistique et optimisation à grande échelle » par Francis Bach (*Inria Paris*).
- 14h30 « Apprentissage de métriques : des algorithmes aux garanties théoriques » par Marc Sebban (*Université Jean Monnet, Saint Etienne*).
- 15h30 Pause
- 16h00 « Classification croisée : les modèles des blocs latents » par Mohamed Nadif (*Université Paris Descartes*).
- 17h00 Discussion
- 17h30 Clôture

## Organisation

Cet événement est organisé par Khalid Benabdeslem pour l'Association Française pour l'Intelligence Artificielle (AFIA) et Pascale Kuntz pour la Société Francophone de Classification (SFC).

## Inscriptions

La participation à la journée est gratuite et ouverte à tous. Pour des raisons de logistique, nous vous demandons néanmoins de vous inscrire auprès de [khalid.benabdeslem@univ-lyon1.fr](mailto:khalid.benabdeslem@univ-lyon1.fr).



## MD&IA 2016

L'AFIA et le GdR MaDICS du CNRS organisent une journée d'étude sur l'articulation entre les problématiques posées par l'Intelligence Artificielle (IA) et celles posées par l'arrivée des données massives et le développement de la Science des Données (SD). L'IA ne constitue-t-elle qu'un terrain d'application aux techniques développées en SD ? Les réflexions menées en IA peuvent-elles impacter la manière dont les recherches en big data sont menées ? Les dernières avancées en SD sont-elles promises à résoudre la question du sens traitée en IA depuis de nombreuses années ? La notion de modèle va-t-elle disparaître au profit de systèmes ascendants fondés sur l'accumulation des données ? Pour le moins, les deux champs de recherche partagent des interrogations communes, comme la place de l'humain et le degré d'autonomie des systèmes mis en place. La journée constituera un lieu naturel de discussions et d'échanges sur ces questions.

La journée se déroulera au travers d'exposés et de posters portant sur des réflexions générales ou des applications particulières, et deux invités apporteront un éclairage complémentaire sur ces questions

### Date et Lieu

- Date 2 décembre 2016
- Lieu Maison International des Langues et des Cultures, 35 rue Raulin, 75007 Lyon

### Programme

- 10h00 Présentations de l'AFIA par Yves Demazeau (*Président de l'AFIA*), du GdR MaDICS par Christine Collet (*Directrice du GdR*) et de la journée par Vincent Claveau et Julien Velcin.
- 10h30 « Big data vs. the right data » par Boi Faltings (*EPFL Lausanne*)
- 11h30 « Modélisation en Intelligence Artificielle au temps du *big data* : où est passée l'intelligence ? » par Fabrice Muhlenbach (*Univ. Saint Etienne*)
- 12h00 « Les machines ont-elles perdu leur esprit ? » par Jean-Sébastien Vayre (*Univ. Toulouse 2*)
- 12h30 Pause Repas
- 14h00 « Learning deep representations in NLP » par Holger Schwenk (*FAIR Paris*)
- 15h00 « Combiner RDF, Datalog et SPARQL pour intégrer et analyser à la demande des données hétérogènes : retour d'expérience » par Marie-Christine Rousset et al (*Univ. Grenoble Alpes*)
- 15h30 Posters
  - « Big data for transport and logistics » par Mourad Abed et al (*Univ. Valenciennes*)
  - « What can I do ? Guiding users in a world of automated decisions » par Matthias Gallé (*XRCE Meylan*)
  - « Clustering sur de grands volumes de séquences d'origine biologique » par Alexandre Bazin et al (*Univ. Clermont-Ferrand*)
  - « From Mobility to Adaptive Services » par Nicolas Gutowski et al (*ESEO Angers*)
  - « Conception d'un modèle généraliste pour l'évaluation d'un test A/B » par Emmanuelle Claeys et al (*MC2I Groupe Paris*)
- 17h30 Discussion et clôture

### Organisation

Cet événement est organisé par Julien Velcin pour l'Association Française pour l'Intelligence Artificielle (AFIA) et Vincent Claveau pour le GdR Masse de Données, Informations et Connaissances en Sciences du CNRS (GdR MaDICS). La journée est organisée avec le soutien du laboratoire ERIC.

### Inscriptions

Les inscriptions se font sur le site <http://bigia2016.irisa.fr/#inscription>. L'inscription est gratuite mais obligatoire. L'organisation et la prise en charge du déjeuner restent à la discrétion des participants.



**AFIA**  
Association française  
pour l'Intelligence Artificielle



# EF&IA 2017

Cet événement commun organisé par le MEDEF et l'AFIA a pour ambition de promouvoir et de favoriser le développement de l'Intelligence Artificielle au sein des Entreprises de France. Des entreprises de toutes tailles et secteurs présenteront les applications et technologies d'IA qu'elles développent et les bénéfices qu'elles en tirent pour leur clientèle et leur propre croissance. L'objectif de la journée est double : informer de l'impact de l'Intelligence Artificielle (IA) pour la compétitivité des entreprises et échanger sur les perspectives de ces technologies pour la création de marchés d'avenir.

## Date et Lieu

- Date 23 janvier 2017
- Lieu 55 avenue Bosquet 75007 Paris

## Programme

- 09h00 Introduction par Pierre Gattaz, Président du MEDEF et Laurent Gouzènes, Président du Comité Recherche-Innovation au MEDEF (*MEDEF 75-Paris*).
- 09h10 Intervention de Thierry Mandon, Secrétaire d'Etat à l'Enseignement Supérieur et à la Recherche (*MENESR 75-Paris*)
- 09h30 « Intelligence Artificielle de France » par Yves Demazeau, Président de l'AFIA.
- 09h50 « Le renouveau de l'Intelligence Artificielle et les Technologies Financières » par Yves Caseau, Directeur du Digital (*AXA Groupe Paris*)
- 10h10 « Ingénierie de la connaissance : mémoire de la société, transfert des savoirs, tremplin de l'innovation » par Alain Berger, Directeur Général (*Ardans 78-Montigny-le-Bretonneux*)
- 10h30 « L'Intelligence Artificielle, levier de progrès pour le développement des produits et services innovants de l'industrie automobile » par Patrick Bastard, Directeur de l'Ingénierie des Systèmes d'Aide à la conduite (*Renault 78-Guyancourt*)
- 10h50 « *We didn't do anything wrong, but somehow we failed*. Que faire pour réussir ? » par Patrick Albert, Président (*SuxccessionWeb 75-Paris*)
- 11h10 Pause-café
- 11h30 « Relation homme-système en aéronautique : l'apport de l'Intelligence Artificielle » par Bruno Patin, Responsable Etudes Amont Autonomie (*Dassault Aviation 92-Saint-Cloud*)
- 11h50 « Réseaux Bayésiens : modélisation des connaissances et apprentissage à partir des données pour l'aide à la décision » par Lionel Jouffe, Président (*Bayesia 53-Laval*)
- 12h10 « L'IA Appliquée aux Infrastructures Télécom et à la Relation Clients » par Henri Sanson, Directeur Recherche Données et Connaissances (*Orange Labs 35-Cesson-Sévigné*)
- 12h30 « Gestion optimisée des ressources : garantir la performance, intégrer l'humain, réduire l'empreinte environnementale des villes et des industries » par Jean-Denis Muller, Directeur Innovation Numérique (*Veolia Recherche & Innovation 78-Limay et Maisons-Laffitte*)
- 12h50 Discussion animée par Laurent Gouzènes, Directeur Scientifique (*Pacte Novation et Taj 92-Neuilly*) et Yves Demazeau, Directeur de Recherche (*CNRS 38-Grenoble*).
- 13h00 Cocktail

## Organisation

Cet événement est organisé par Yves Demazeau et Bruno Patin pour l'Association Française pour l'Intelligence Artificielle (AFIA), et par Laurent Gouzènes et Patrick Schmitt pour le Mouvement des Entreprises DE France (MEDEF).

## Inscriptions

Les inscriptions, gratuites, se font sur le site <http://communication.medef.fr/LP/lhHybQRcPiOO>.



**AfIA**  
Association française  
pour l'Intelligence Artificielle



# PS&IA 2017

Dès son apparition dans les années 1950, l'Intelligence Artificielle (IA) a suscité une intense curiosité chez les philosophes. Soixante ans plus tard, l'IA semble en passe de remplir ses promesses et d'introduire une série de bouleversements économiques, culturels voire anthropologiques. Mais l'IA d'aujourd'hui est-elle celle d'hier ? Qu'annoncent ses succès présents ? Pour aborder ces questions, une réévaluation des fondements de l'IA et de sa place dans le système technoscientifique semble nécessaire.

À cette tâche, l'AFIA et la SPS organisent une journée PHILOSOPHIE DES SCIENCES et INTELLIGENCE ARTIFICIELLE, focalisée sur trois thèmes complémentaires : philosophie de l'IA et IA comme philosophie expérimentale ; impact du tournant pragmatique pour l'IA, en particulier approches 4EA (*Embodied, Embedded, Enacted, Extended, Affective*) ; apport des concepts d'émergence et de complexité.

## Date et Lieu

- Date 02 février 2017
- Lieu Salle Jaurès, Ecole Normale Supérieure, 29 rue d'Ulm, 75005 Paris

## Programme

- 08h30 Accueil des participants
- 09h00 Présentations de l'AFIA, par Yves Demazeau (*Président de l'AFIA*), de la SPS par Stéphanie Ruphy (*Présidente de la SPS*), du DEC par Daniel Andler. Présentation de la journée par Daniel Andler (*ENS Paris*) et Robin Lamarche-Perrin (*CNRS Paris*)
- 09h30 « L'IA comme épistémologie expérimentale : regards sur 40 ans d'histoire » par Joël Quinqueton (*U. Montpellier 3*)
- 10h20 Pause
- 10h40 « GOFAI, NEWFAI et le problème de l'intentionnalisme » par Jean-Michel Roy (*ENS Lyon*)
- 11h30 « Le rôle de la complexité et de la simplicité dans l'émergence des capacités cognitives » par Jean-Louis Dessalles (*Telecom ParisTech*)
- 12h20 Repas (offert)
- 13h40 « D'une intelligence artificielle à une intelligence incarnée dans l'activité humaine » par Catherine Garbay (*CNRS Grenoble*)
- 14h30 « Intelligence artificielle et *cognition 4E* : un changement de paradigme ? » par Pierre Steiner (*UTC Compiègne*)
- 15h20 Pause
- 15h40 « Questionnements éthiques sur la Robotique et l'IA » par Raja Chatila (*U. Paris 6*)
- 16h30 Discussion
- 17h30 Clôture

## Organisation

Cet événement est organisé par Robin Lamarche-Perrin pour l'Association Française pour Intelligence Artificielle (AFIA) et Daniel Andler pour la Société de Philosophie des Sciences (SPS), avec le soutien du Département d'Etudes Cognitives de l'ENS Ulm.

## Inscriptions

L'inscription est gratuite mais obligatoire : <https://www.eventbrite.fr/e/inscription-journee-philosophie-des-sciences-et-intelligence-artificielle-30888523405>. Elle inclut la participation au repas et aux pauses-café.







**AFIA**  
Association française  
pour l'Intelligence Artificielle



# IHM&IA 2017

L'AFIIHM et l'AFIA organisent leur troisième journée commune INTERACTION HOMME-MACHINE & INTELLIGENCE ARTIFICIELLE.

Cette journée s'organisera autour de présentations, posters, démos, échanges sur des sujets à l'intersection des deux champs ou mettant en évidence les liens possibles entre les deux disciplines.

## Date et Lieu

- Date 17 mars 2017
- Lieu UPMC, Amphithéâtre 25, 4 place Jussieu, 75005 Paris

## Programme

- 08h30 Accueil
- 09h00 Présentations de l'Association Française pour l'Intelligence Artificielle, par Yves Demazeau (*Président de l'AFIA*) et de l'AFIIHM par Géry Casiez (*Président de l'AFIIHM*).
- 09h20 « Interface multimodale sur une table graphique multi-touch multi-utilisateurs » par Jean-Paul Barthès (*Université de Technologie de Compiègne*)
- 09h50 « Evolutionary visual exploration using EvoGraphDice » par Nadia Boukhelifa (*Telecom ParisTech*)
- 10h20 Pause
- 10h40 « Co-apprentissage pour les interfaces cerveau-ordinateur : de la science-fiction à l'interaction homme-machine » par Nataliya Kosmyna (*Inria Rennes*)
- 11h10 « Apport des systèmes multi-agents à l'interaction tangible sur table interactive RFID » par Christophe Kolski (*Université de Valenciennes et du Hainaut-Cambrésis*)
- 11h40 « Reconnaître un objet à partir de points de contact sur une surface tactile » par Caroline Appert (*CNRS Saclay*)
- 12h10 Déjeuner
- 14h00 « Vers une meilleure appréciation des algorithmes qui nous entourent » par Fanny Chevalier (*Inria Lille*)
- 14h30 « Conception participative d'un e-learning intelligent » par Alexandra Delmas (*Université de Bordeaux 2*)
- 15h00 « Les systèmes autonomes sont des outils informatiques comme les autres » par Nicolas Roussel (*Inria Lille*)
- 15h30 Pause
- 16h00 Démonstrations, posters, et échanges
- 18h00 Clôture

## Organisation

Cet événement est organisé par Nicolas Sabouret pour l'Association Française pour l'Intelligence Artificielle (AFIA) et Gilles Bailly pour L'Association Francophone d'Interaction Homme-Machine (AFIIHM) avec le soutien du Labex SMART.

## Inscriptions

Les inscriptions se font sur le site <http://ihmia.afihm.org/inscription/index.html>. L'inscription est gratuite et inclut la participation aux pause-café. L'organisation et la prise en charge du déjeuner restent à la discrétion des participants.



# EIAH&IA 2017

L'AFIA et l'ATIEF organisent une troisième journée ENVIRONNEMENTS INFORMATIQUES POUR L'APPRENTISSAGE HUMAIN & INTELLIGENCE ARTIFICIELLE juste avant la conférence EIAH 2017.

L'objectif de cette journée est de rassembler dans un même espace-temps deux communautés, les chercheurs en Intelligence Artificielle et ceux travaillant dans le domaine des Environnements Informatiques pour l'Apprentissage Humain, afin de dialoguer et échanger autour de problématiques, de techniques, de concepts communs et de partager des retours d'expériences.

En 2013 il n'y avait pas de thème pour cette journée alors qu'en 2015 l'ensemble des présentations portaient sur l'utilisation de la fouille de données dans les EIAH. Cette année nous mixons ces deux approches. Tout d'abord, grâce à des présentations de papiers retenus suite à un appel à communications, nos échanges seront multi thématiques. Ensuite nous aborderons la problématique de la conception, la réalisation et l'expérimentation d'EIAH utilisant la réalité virtuelle et la réalité augmentée. Un invité nous fera part de son expertise et de ses retours d'expérience sur ce thème et nous terminerons avec une table ronde.

## Date et Lieu

- Date 06 Juin 2017
- Lieu Hôtel du Département – Place du Quartier Blanc – 67000 Strasbourg

## Programme

- 14h00 Présentations de l'Association Française pour l'Intelligence Artificielle, par Yves Demazeau (*Président de l'AFIA*) et de l'Association des Technologies de l'Information pour l'Education et la Formation par Serge Garlatti (*Président de l'ATIEF*).
- 14h15 « Une architecture multicouche pour l'analyse des compétences non techniques en situation critique » par Yannick Bourrier (*Université Paris 6*), Francis Jambon, Catherine Garbay et Vanda Luengo.
- 14h45 « Le contexte pour personnaliser les systèmes tutoriels » par Monique Grandbastien (*Université de Lorraine*)
- 15h15 « Apprentissage : corriger et visualiser » par Gabriel Illouz (*Université Paris 11*), Anne-Laure Ligozat et Frédéric Vernier.
- 15h45 Pause
- 16h00 « Développer l'apprentissage et les compétences-piétons par la réalité virtuelle : enjeux, perspectives et limites » par Jérôme Dinet (*Université de Lorraine*).
- 17h00 Table Ronde
- 17h45 Clôture

## Organisation

Cet événement est organisé par Davy Monticolo pour l'Association Française pour l'Intelligence Artificielle (AFIA) et Nicolas Delestre pour l'Association des Technologies de l'Information pour l'Education et la Formation (ATIEF).

## Inscriptions

Les inscriptions se font via le site : <http://eiah2017.unistra.fr/>



**AFIA**  
Association française  
pour l'Intelligence Artificielle

PRÉ-GDR IA



# JI&IA 2017

Cette deuxième journée JEUX INFORMATISES et INTELLIGENCE ARTIFICIELLE fait suite au succès rencontré lors de la première journée tenue lors de PFIA 2015 à Rennes. Tous les jeux sont concernés, aussi bien les jeux de plateau que les jeux vidéo ou les jeux à information incomplète.

L'objectif de cette deuxième journée est de rassembler la communauté française en IA des jeux (ou *Game AI*) et de présenter à ceux de nos collègues qui s'y connaissent moins ce qu'est le *Game AI* : ce qu'on y fait, et en quoi les jeux sont des terrains applicatifs idéaux pour l'intelligence artificielle.

Cette journée est inscrite au sein de la Plate-Forme Intelligence Artificielle 2017.

## Date et Lieu

- Date 03 juillet 2017
- Lieu Université de Caen Normandie, Campus 2, Côte de Nacre, bd. Maréchal Juin, 14032 Caen

## Programme

- 10h15 Présentations de l'Association Française pour l'Intelligence Artificielle, par Yves Demazeau (*Président de l'AFIA*) et du pré-GdR IA par Sébastien Konieczny (*Porteur du pré-GdR IA*).
- 10h30 « Stockfish, la meilleure IA d'Échecs au monde » par Stéphane Nicolet (*Université Paris 2*)
- 11h30 Série de présentations courtes
- 12h30 Déjeuner
- 14h00 « SPRITE : un jeu sérieux pour l'enseignement de la gestion des risques naturels » par Carole Adam (*Université Grenoble Alpes*)
- 14h30 Discussion autour des jeux sérieux. Pourquoi les jeux sérieux ? Quels enjeux ? Quelles possibilités de partenariat ?
- 15h30 Pause
- 16h00 « L'industrie du jeu vidéo aujourd'hui, ses besoins et ses attentes » par Eddy Célestine (*Altangames Nantes*)
- 16h30 Réunion du groupe de Travail « IA des Jeux » du pré-GdR IA.
- 18h00 Clôture

## Organisation

Cet événement est organisé par Tristan Cazenave pour l'Association Française pour l'Intelligence Artificielle (AFIA) et Florian Richoux pour le Groupe de Travail « IA des Jeux » du pré-GdR IA.

## Inscriptions

Les inscriptions se font sur le site de PFIA 2017 : <https://pfia2017.greyc.fr/inscriptions> L'inscription à la Journée inclut la participation au repas et aux pause-café.



# E&IA 2017

L'AFIA et le COMETS, en association avec la CNIL, organisent une deuxième Journée ÉTHIQUE et INTELLIGENCE ARTIFICIELLE avec pour vocation de croiser les questionnements et les approches, et de créer une communauté de réflexion sur ces sujets.

La question d'une régulation ou d'un contrôle éthique des agents logiciels ou robotiques ainsi que des systèmes constitués d'humains et d'agents autonomes se pose avec de plus en plus de force. Ces questions sont par nature transverses à de multiples disciplines. Le programme de la journée est constitué de présentations invitées et de présentations courtes relatifs à cette problématique.

Cette journée est inscrite au sein de la Plate-Forme Intelligence Artificielle 2017.

## Date et Lieu

- Date 05 juillet 2017
- Lieu Université de Caen Normandie, Campus 2, Côte de Nacre, bd. Maréchal Juin, 14032 Caen

## Programme

- 10h15 Présentations de l'Association Française pour l'Intelligence Artificielle, par Yves Demazeau (*Président de l'AFIA*) et du Comité d'Éthique du CNRS par Jean-Gabriel Ganascia (*Président du Comité d'Éthique*).
- 10h30 « Contre la transparence. De la valeur du hasard pour un système apprenant » par Alexei Grinbaum (*CEA Saclay*)
- 11h15 « Une théorie formelle des agents moraux » par Emiliano Lorini (*CNRS Toulouse*)
- 11h40 « Approches logiques de la modélisation du raisonnement éthique » par Gauvain Bourgne (*Université Paris 6*)
- 12h30 Déjeuner
- 14h00 « Robots conversationnels empathiques : enjeux technologiques et éthiques » par Laurence Devillers (*Université Paris 4*)
- 14h45 « Pistes de réflexion sur une modélisation de l'éthique pour des agents intelligents autonomes » par Léa Guizol (*Ardans Montigny-le-Bretonneux*)
- 15h05 « ETHICAAA - Éthique et Agents Autonomes » par Grégory Bonnet (*Université de Caen*)
- 15h30 Pause
- 16h00 « Les problèmes juridiques posés par la robotique et l'IA » par Nathalie Nevejans (*Université d'Artois*)
- 16h45 Table ronde « *IEEE Standard Ethics* » avec L. Devillers, A. Grinbaum, N. Nevejans, et C. Tessier.
- 18h00 Clôture

## Organisation

Cet événement est organisé par Olivier Boissier pour l'Association Française pour l'Intelligence Artificielle (AFIA) et Jean-Paul Delahaye pour le Comité d'Éthique (COMETS) du CNRS.

## Inscriptions

Les inscriptions se font sur le site de PFIA 2017 : <https://pfia2017.greyc.fr/inscriptions> L'inscription à la Journée inclut la participation au repas et aux pause-café.

# CONGRES RFFIA 2016

## Reconnaissance de Formes § Intelligence Artificielle



27 juin  
1<sup>er</sup> juillet 2016  
CLERMONT-FERRAND

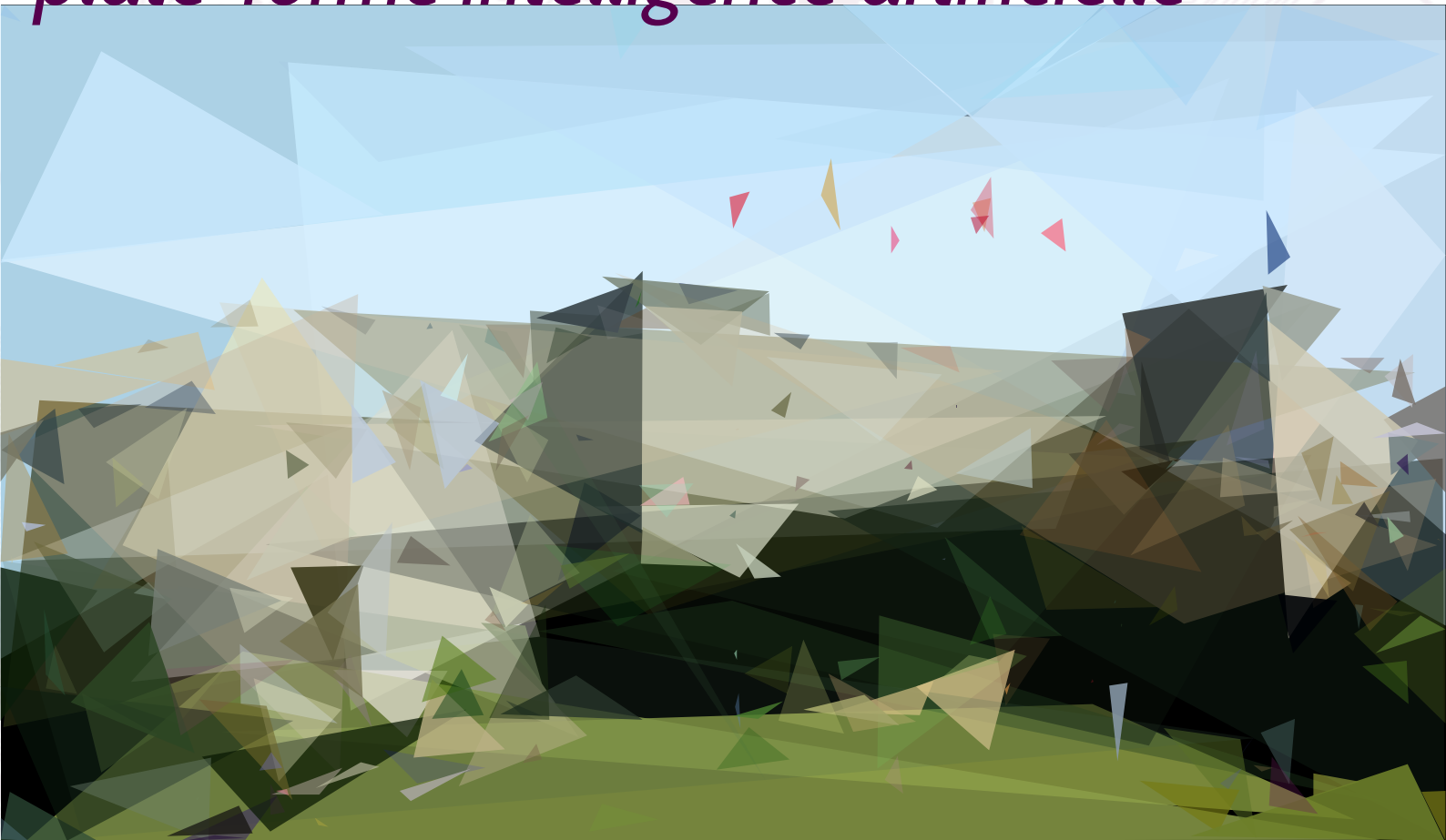


APIA • IAF • IC • JFPDA • JFSMA • RJCIA

jeux informatisés & IA • éthique & IA • compétition d'IA • démos • FR@IJCAI

# PFIA17

*plate-forme intelligence artificielle*



du 3 au 7 juillet 2017

Université de Caen Normandie – Campus 2

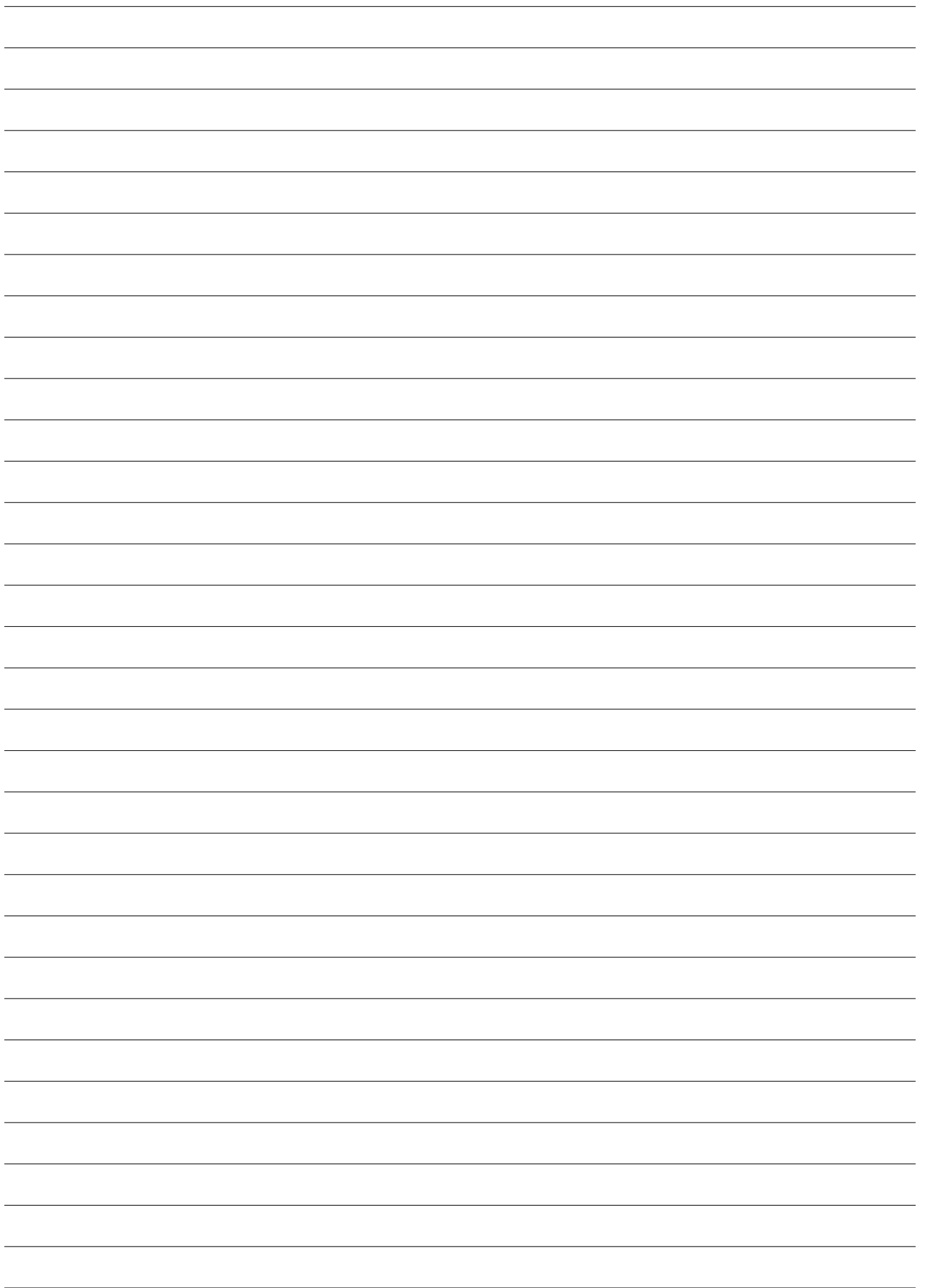


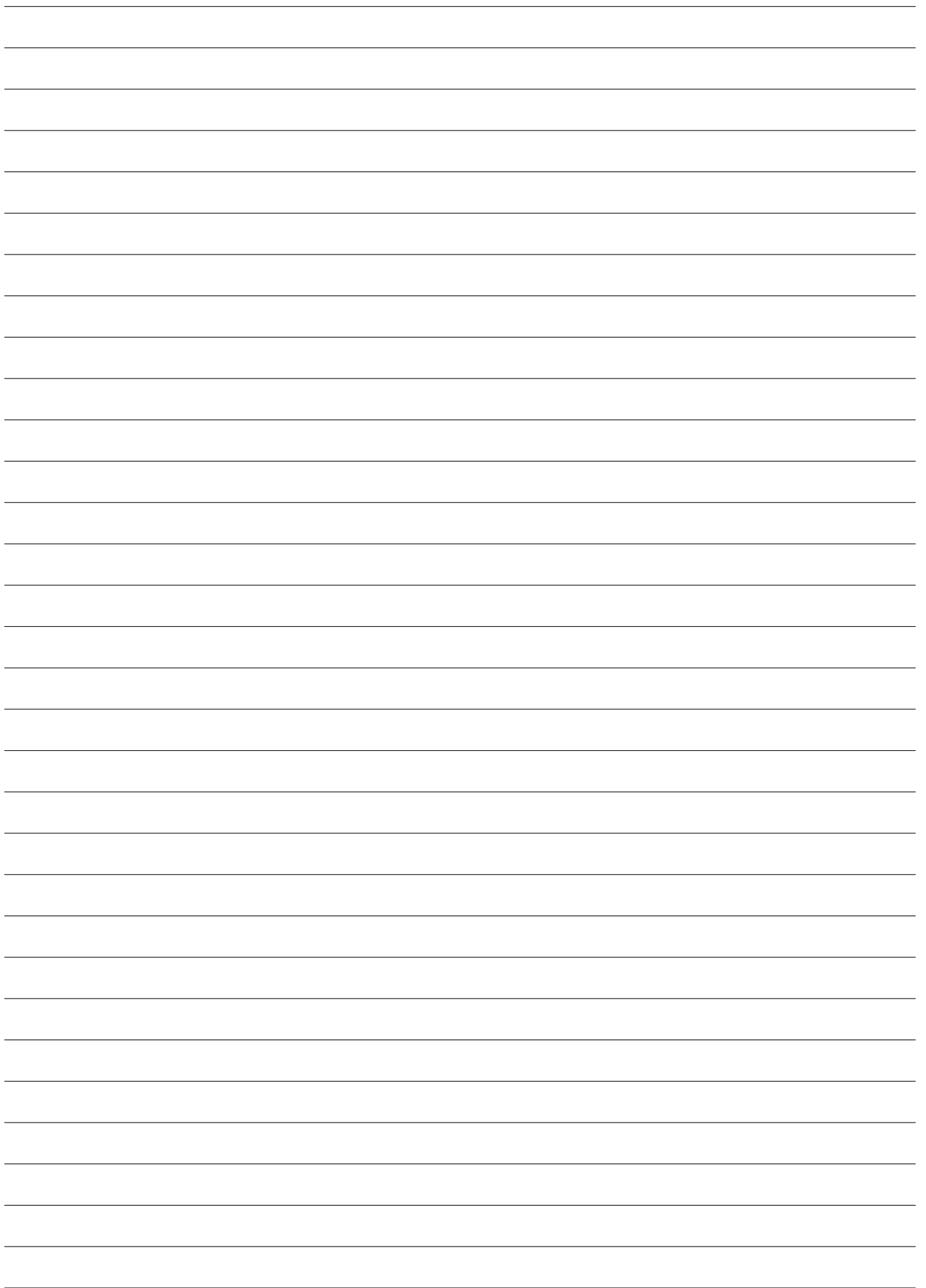
<https://pfia2017.greyc.fr>

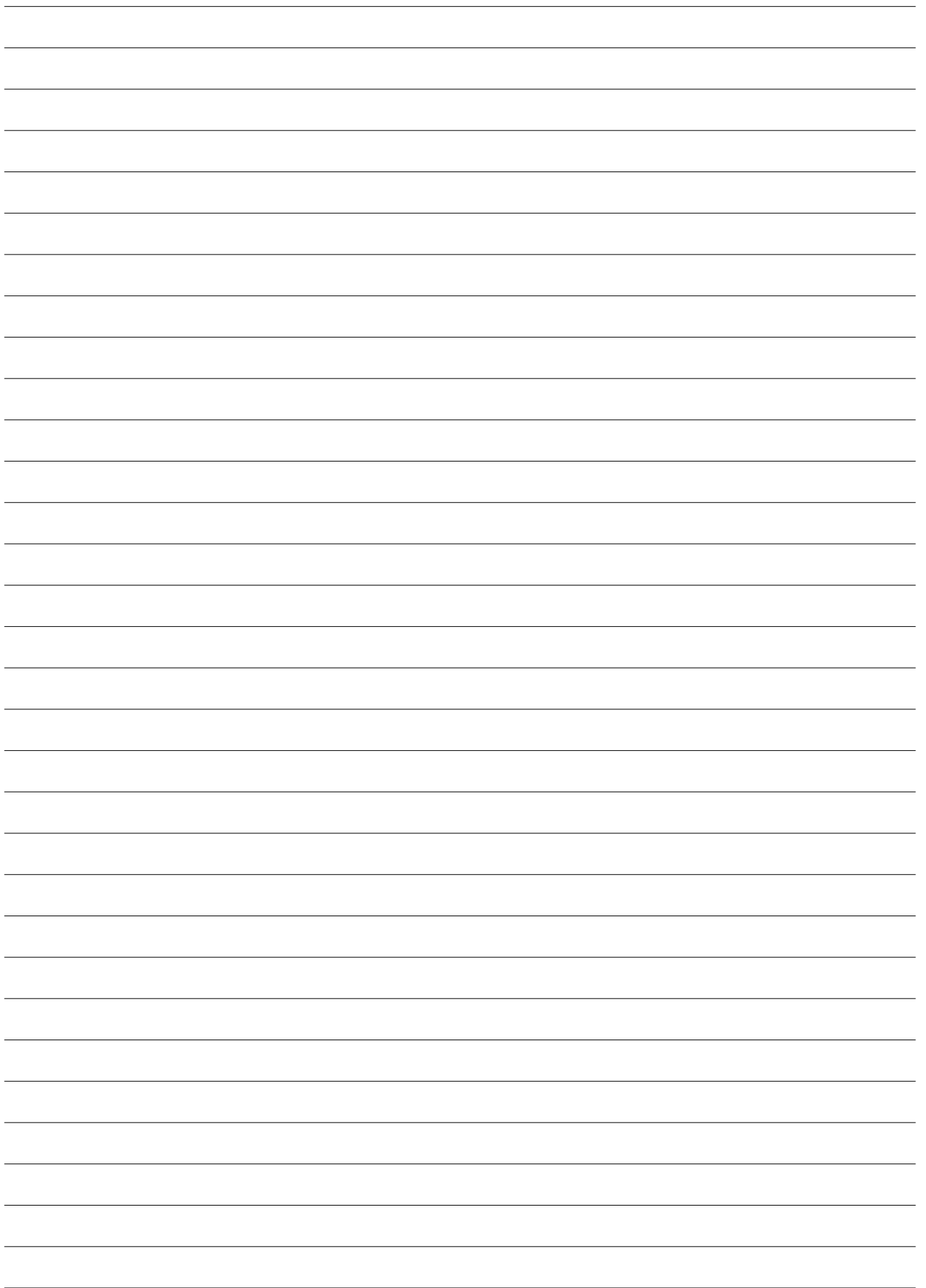














---

En partenariat avec



L'AVENIR EST AUX VALEURS SÛRES

---