



HAL
open science

Autonomie et représentation de l'environnement

Frédéric Armetta

► **To cite this version:**

Frédéric Armetta. Autonomie et représentation de l'environnement. Intelligence artificielle [cs.AI]. Université Lyon 1 - Claude Bernard, 2024. <tel-04842891>

HAL Id: tel-04842891

<https://hal.science/tel-04842891v1>

Submitted on 11 Aug 2025

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

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



HAL Authorization

HABILITATION À DIRIGER DES RECHERCHES

AUTONOMIE ET REPRÉSENTATION DE L'ENVIRONNEMENT

PRÉSENTÉE PAR

FRÉDÉRIC ARMETTA

préparée au laboratoire LIRIS

Université Claude Bernard Lyon 1

Spécialité Informatique

SOUTENUE PUBLIQUEMENT LE 02 DÉCEMBRE 2024

COMPOSITION DU JURY :

OLIVIER BOISSIER, Professeur, École des Mines Saint-Étienne Rapporteur
ALAIN DUTECH, Chargé de recherche, HDR, INRIA NGE Rapporteur
GAUTHIER PICARD, Directeur de recherche, ONERA Rapporteur
MARIE-PIERRE GLEIZE, Professeure, Université Toulouse III Examinatrice
MAXIME MORGE, Professeur, Université Lyon1 Examinateur
NICOLAS SABOURET, Professeur, Université Paris-Saclay Examinateur

Remerciements

L'aboutissement de ce mémoire d'Habilitation à Diriger des Recherches marque une étape importante de mon parcours académique et professionnel. Il est le fruit de nombreuses années de travail, de réflexions, de collaborations et de rencontres, et je tiens à exprimer ici ma profonde gratitude à celles et ceux qui m'ont accompagnés tout au long de ce chemin.

Je remercie tout d'abord les membres du jury de mon HDR, pour avoir accepté de lire et d'évaluer ce mémoire, pour leurs remarques constructives, leur bienveillance et l'intérêt qu'ils ont manifesté à l'égard de mes travaux. Merci à Olivier Boissier, Alain Dutech et Gauthier Picard d'avoir rapporté sur ce travail. Merci à Marie-Pierre Gleize, Maxime Morge et Nicolas Sabouret pour leur présence au sein de ce jury.

Je tiens à exprimer ma reconnaissance à mes collègues de recherche, passés et présents. J'adresse une pensée toute particulière à Salima Hassas, fondatrice de l'équipe SyCoSMA, dont les encouragements à rédiger ce mémoire, les nombreuses collaborations, la présence attentive et le soutien constant, dans les moments de doute comme dans ceux d'enthousiasme, ainsi que le positionnement scientifique inspirant, ont largement contribué à la réalisation de cette aventure.

Merci aux membres dits permanents de l'équipe SyCoSMA. Je ne compte pas les heures passées à partager nos positionnements, merci Samir Aknine, Rémy Chaput, Olivier Georgeon, Mathieu Lefort, Laetitia Matignon, Bruno Yun.

Merci également aux doctorantes et doctorants dont j'ai pris part à l'encadrement de thèse. Gaël Clair, Ruiqi Dai, André Fabbri, Maxime Guériau, Olivier Lefevre, Maya Medjad, Quang-Anh Nguyen Vu, Sébastien Mazac, Julien Perier-Camby.

Merci à tous les collaborateurs du LIRIS ou d'autres laboratoires ou entreprises, Marc Bertin, Heike Baldauf-Quilliatre, Romain Billot, Angela Bonifati, Richard Canal, Alice Delorme, Alexia Deporte Richard, Éric Duchêne, Stefan Duffner, Nour-Eddin El Faouzi, Haytham Elghazel, Benoit Gaudou, Mohammed Haddad, Alexandre Meyer, Bruno Malfoy, Lucie Martinet, Emmanuel Olivier, Jean-Charles Quinton, Brooke Stephenson, Raphaël Szymocha, Lucien Tisserand, Nicolas Wiest. Merci à tous ceux que j'ai pu oublier. Merci aux membres de la communauté Multi-Agents ainsi qu'aux membres de la communauté de l'intelligence développementale que j'ai pu côtoyer.

Je n'oublie pas tous les étudiants et étudiantes que j'ai pu encadrer pour leur stages de recherche qui ont été des expériences très enrichissantes, Feras Al Kassar, Anthony Baccuet, Philomène Bobichon, Farah Ben Slama, Pierre Couy, Valentin Chafraix, Etienne Desbois, Tom Dusséaux, Elliot Faugier, Florent Gattoni, Lucas Fournier, Louis Manhès, Maxime Le Duc, Guilhem Marion, Guilhem Marion, Pierre Marec, Nawel Medjkoune, Ciril Moron, Bastien Nollet, Bastien Nollet, Mathieu Regnard, Mohamed Massamba Sene, Benjamin Smati, Kevin Sanlaville.

Merci à tous mes collaborateurs liés à l'enseignement, aux étudiants qui m'ont permis de diversifier mes activités, et plus généralement au personnel permettant le fonctionnement de l'Université.

Un grand merci enfin à ma famille pour leur soutien.

À toutes et à tous, merci !

Préambule

Ce mémoire a pour objectif de synthétiser certains des travaux les plus significatifs au coeur de mon projet de recherche pour le développement d'intelligences artificielles autonomes, développés au sein des laboratoires LIESP, GAMA et depuis juillet 2012 au LIRIS .

Au cours de mes premières années de recherche, je me suis intéressé à ce qu'apportent les Systèmes Multi-Agents pour le domaine, à leurs capacités d'auto-organisation et à l'intelligence collective qu'ils peuvent manifester. Par la suite, je me suis intéressé à des domaines de recherches d'inspirations diverses, telles que les différentes méthodes du domaine de l'apprentissage automatique en informatique, mais également aux inspirations découlant des neurosciences, des sciences cognitives ou même de la philosophie de la connaissance qui peuvent compléter avantageusement notre compréhension du rôle de la connaissance. Je poursuis cette démarche avec beaucoup d'enthousiasme. Les différentes approches du sujet de l'IA sont décrites de façons très différentes suivant les communautés concernées, qui souvent s'ignorent bien que traitant parfois de sujets connexes. De façon générale je constate que les sciences cognitives apportent le recul nécessaire aux problématiques, alors que les méthodes plus appliquées de l'apprentissage en Informatique par exemple, proposent des concrétisations pointues en plaçant cependant quelquefois au second plan le positionnement épistémologique de l'apprentissage réalisé, ce qui à mon sens souligne l'intérêt d'une approche transversale plus globale d'inspirations diverses.

Ce document n'a pas pour objectif de traiter exhaustivement ce très vaste sujet abordé à travers différentes disciplines, mais témoigne des possibilités offertes par le domaine, à partir de différentes applications et axes d'étude basés sur l'expérience de projets auxquels j'ai pu participer. J'espère ainsi être parvenu au sein de ce mémoire à identifier la corrélation entre problèmes applicatifs et axes d'étude permettant de les résoudre ou de tracer une voie vers une résolution à venir. Étant donné la diversité des approches, le vocabulaire descriptif est quelques fois spécifique aux différents domaines, mais afin de faciliter la lecture une terminologie générale sera employée au cours du texte.

Les projets au sein desquels j'ai pu participer ont généralement une déclinaison théorique ainsi qu'une déclinaison applicative. Tout objet de recherche a pour aboutissement un ou plusieurs cadres applicatifs. Il est particulièrement appréciable de voir les différentes propositions se concrétiser au sein d'applications pratiques. L'IA n'est cependant pas suffisamment mature pour être exploitable, c'est la raison pour laquelle on appréciera également les démarches itératives *in vitro* lorsque cela est nécessaire, qui parfois concèdent aux exigences de l'autonomie qui sera définie. J'ai eu l'opportunité de pouvoir participer, au sein de l'équipe Multi-Agents, à des projets des deux types. Certains projets sont plus fondamentaux, et n'ont pas de déclinaison applicative immédiate pour la société. D'autres sont appliqués et très porteurs également. Il me semble

que cet équilibre est un équilibre sain à préserver lorsque la discipline s’y prête.

Finalement, la réification de représentations est un sujet de curiosité que j’apprécie particulièrement à titre personnel et professionnel. À titre personnel, je m’intéresse aux mécanismes psychologiques qui renforcent et nous enlisent dans notre subjectivité tout en nous maintenant dans le déni du caractère artificiel et construit des représentations que nous utilisons, limitant ainsi nos capacités perceptives. Désamorcer ce processus permettrait-il une issue viable ? Du point de vue professionnel, à l’opposé, la démarche s’intéresse à enfermer un système autonome de façon autocatalytique dans ses représentations vers une forme de subjectivité, afin d’accomplir sa tâche efficacement sans plus amples préoccupations. Heureusement, seuls les Hommes s’intéressent à approcher la vérité ce qui nous facilite la conception de nos IA balbutiantes pour quelque temps encore. C’est ce que nous allons voir dès à présent.

Résumé

L'un des enjeux majeurs de l'IA est la conception d'une représentation adaptée pour organiser des connaissances et les exploiter. Dans le cadre du travail présenté, je m'intéresse à la conception autonome de cette représentation, par un agent incarné qui exploite ses interactions avec son environnement afin de construire une représentation qu'il pourra exploiter pour atteindre ses propres objectifs. Le cadre d'étude idéal de ce développement correspond à un environnement ouvert et inconnu préalablement à la conception du système. Les nombreux défis sont discutés et illustrés au cours de ce mémoire : depuis les qualités recherchées pour ce support de représentation jusqu'à sa dynamique qui nécessite des récompenses à définir et des mécanismes d'amorçage. À travers différents projets et différentes méthodes d'apprentissages centralisés ou décentralisés, ce mémoire aborde des enjeux ciblés pour le développement d'un tel agent, de la perception de motifs à partir de signaux bruts continus bas niveau, à l'apprentissage d'algorithmes dans une représentation endogène, mais aussi à l'apprentissage en ligne de concepts sémantiques.

Abstract

One of the major challenges of AI is the design of a suitable representation to organize knowledge and exploit it. In the context of the work presented, I focus on the autonomous design of this representation, by an embodied agent that exploits its interactions with its environment in order to build a representation that it can use to achieve its own objectives. An ideal study framework for this development relies on an open and unknown environment prior to the system's design. Some challenges are discussed and illustrated throughout this research thesis, from the desired qualities for this representation support, to its dynamics which require defining rewards and priming mechanisms. Through different projects and different centralized or decentralized learning methods, this thesis addresses targeted issues for the development of such an agent, from the perception of patterns from raw continuous low-level signals, to the learning of algorithms in an endogenous representation, but also to the online learning of semantic concepts.

Table des matières

Remerciements	i
Préambule	iii
Résumé	v
Abstract	v
Table des matières	vii
Glossaire	xi
Table des figures	xiii
1 Introduction	1
1.1 L'Autonomie au cœur de l'Intelligence Artificielle	1
1.2 Organisation du mémoire	3
2 Positionnement	7
2.1 Problématique de recherche	7
2.2 Support d'inscription de connaissances	9
2.2.1 Un support de connaissance agnostique et endogène	9
2.2.2 Capacités de généralisation	11
2.2.3 Un support d'apprentissage résilient	12
2.3 Processus d'apprentissage	13
2.3.1 Perception et action	13
2.3.2 Le <i>feedback</i>	15
2.3.3 Apprentissage et convergence	16
2.4 Domaines de contributions	18
3 Synthèse des travaux	21
3.1 Auto-organisation de topologie pour le partage de données d'un réseau pair-à-pair	21
3.1.1 Réseaux pair-à-pair et partage de documents	22
3.1.2 Marquer les chemins les plus efficaces	22
3.1.3 Une topologie adaptative issue de l'activité du système	23
3.2 Découverte d'environnement décentralisé et mécanismes de contrôle de fiabilité	28

3.2.1	Application à la cartographie collaborative	29
3.2.2	Application à la régulation d'un ensemble de véhicules connectés	30
3.3	Apprentissage de régularités en environnement continu et Intelligence Ambiante	32
3.3.1	Intelligence Ambiante	33
3.3.2	Des évènements et des signaux continus	34
3.3.3	Une dynamique de contrôle décentralisée auto-catalytique	35
3.3.4	Simulateur	36
3.4	Extensions de MCTS pour les jeux combinatoires	37
3.4.1	Jeux combinatoires et jeu de Go	38
3.4.2	MCTS (Monte Carlo Tree Search)	40
3.4.3	Un processus d'acquisition de connaissances autonome pour MCTS	41
3.4.4	Réification d'objectifs intermédiaires pour MCTS	43
3.5	Apprentissage et régulation du trafic routier	46
3.5.1	Des véhicules connectés pour la régulation	47
3.5.2	Simulation et lois de poursuite	47
3.5.3	Construction d'une représentation par apprentissage concurrent	48
3.6	Apprentissage continu non supervisé d'objets à partir d'un flux vidéo	52
3.6.1	Exploiter les régularités spatio-temporelles d'un flux vidéo	52
3.6.2	Rôle de l'action pour l'apprentissage de représentations visuelles	55
3.6.3	Détection d'objets en apprentissage continu	58
3.7	IA et apprentissage d'algorithmes	62
3.7.1	Un domaine émergent?	63
3.7.2	Exécution étape par étape	64
3.7.3	Algorithme de bout en bout	65
4	Bilan et perspectives	69
4.1	Bilan	69
4.2	Perspectives	71
4.2.1	Vers un support de connaissances toujours plus qualitatif	71
4.2.2	De nouveaux <i>feedbacks</i> pour nos IA	72
	Bibliographie	82
	Curriculum Vitae des activités de recherche et d'enseignement	83
1	Synthèse du parcours professionnel et contexte d'exercice	83
2	Activité scientifique	83
2.1	Présentation synthétique des thématiques de recherche	84
2.2	Liste des encadrements scientifiques	84
2.3	Publications et productions scientifiques	87
2.4	Diffusion et rayonnement	92
2.5	Responsabilités scientifiques	95

3	Investissement pédagogique	97
3.1	Présentation synthétique de l'activité d'enseignement	97
3.2	Présentation des enseignements	97
3.3	Responsabilités pédagogiques	98
4	Responsabilités collectives et d'intérêt général	99
4.1	Responsabilités administratives	99
4.2	Responsabilités et mandats locaux ou régionaux	99
4.3	Autres responsabilités	100
5	Annexes	101
5.1	Tableau des enseignements	101
5.2	Liste des encadrements scientifiques de master	102

Glossaire

Approche inspirée du constructivisme ou développementale. Approche selon laquelle la représentation est construite par le système de façon incrémentale et auto-entretenu.

Cognitivisme. Conception de la cognition comme un processus de traitement de l'information. Du point de vue de l'intelligence artificielle, ceci correspond à une manipulation de représentations symboliques effectuée par un ensemble de règles.

Jeu de données (*Dataset*) représentatif et étiqueté. Ensemble de données structurées et annotées qui est utilisé pour entraîner et évaluer des modèles d'apprentissage automatique (*machine learning*). Le jeu de données doit refléter la diversité des cas et des situations que le modèle rencontrera dans le monde réel. Chaque élément du *dataset* (exemple de données) est associé à une ou plusieurs étiquettes qui fournissent une information de référence. Ces étiquettes servent de réponses correctes pour l'algorithme de machine learning pendant l'apprentissage supervisé.

Environnement ouvert. Qui n'est pas connu ou étiqueté à priori, qui contient ses propres objets et sa propre dynamique.

Représentation endogène. Qui est produit par le système considéré, en dehors de tout apport extérieur.

Représentation subjective du monde. Cette formulation fait référence à un agent incarné dans son environnement qui construit une représentation de son environnement utile à son exploitation.

Table des figures

2.1	Agent - Environnement - <i>Feedback</i>	14
2.2	Du signal (capteurs) au contrôle (effecteurs)	14
2.3	Domaines de contributions	19
3.1	Mécanisme de création de topologie phéromonale	24
3.2	Un réseau pair-à-pair en 3 couches	26
3.3	Topologie à invariance d'échelle du réseau engendré par les dépôts de phéromones	27
3.4	Résilience du processus de recherche de documents	28
3.5	Dynamique des réseaux de confiance (<i>Trust Net</i>)	30
3.6	Vue du simulateur multi-agents <i>Danger Mapping</i>	31
3.7	Architecture en 3 couches	32
3.8	Intelligence Ambiante : des capteurs et des effecteurs	33
3.9	Processus de discrétisation de variables continues	35
3.10	Un contrôle décentralisé à base de multi-agents pour l'apprentissage de régularités	36
3.11	Simulateur multi-agents du processus de construction de motifs	37
3.12	Évaluation des positions pour le jeu de Go	39
3.13	Algorithme MCTS	41
3.14	Background History Reply Forest process	42
3.15	Taux de victoire : MCTS + BHRF(ϵ) - MCTS + simulations avec heuristiques	44
3.16	Réification d'objectifs intermédiaires pour MCTS	45
3.17	Gain obtenu avec l'ajout de stratégies locales pour le Clobber Solitaire suivant le nombre de simulations	46
3.18	Régulation de trafic par véhicules autonomes	48
3.19	Aperçu du modèle de régulation	49
3.20	3 simulations de flux de véhicules	51
3.21	Vue générale de la reconnaissance spatio-temporelle d'objets	54
3.22	Déplacement par une action de la fenêtre de perception	56
3.23	Modèle d'apprentissage de deux tâches : reconstruction et prédiction	57
3.24	Précision des classifications résultantes suivant le nombre de perceptions accumulées	58
3.25	Aperçu de l'architecture neuronale de CURL [Rao19]	60

3.26	Matrices de confusion entre la valeur réelle et les composants inférés pour CURL sur l'ensemble de test MNIST et l'ensemble de test Fashion-MNIST	61
3.27	Matrices de confusion entre la valeur réelle et les composants inférés pour CURL sur l'ensemble de test MNIST et l'ensemble de test Fashion-MNIST	61
3.28	Exemple de représentation d'une multiplication de deux opérandes à 2 chiffres	64
3.29	Architecture <i>Seq2Seq</i>	66
3.30	Évolution du taux d'erreur pour chacune des tâches	67

1

Introduction

Dans ce chapitre, nous discutons des caractéristiques pouvant qualifier une Intelligence Artificielle et introduisons ensuite l'organisation de ce mémoire.

1.1 L'Autonomie au cœur de l'Intelligence Artificielle

La question centrale de l' « Intelligence » (sa définition, sa compréhension et son émulation) a été abordée par de nombreuses disciplines (Intelligence Artificielle en Informatique, Neurosciences, Psychologie, Linguistique, Philosophie et Anthropologie) dans ses différentes facettes regroupées sous l'appellation de sciences cognitives. L'IA informatique qui est une discipline supportant l'implémentation de processus "intelligents", connaît elle-même de nombreuses écoles et des outils variés. La thématique de la cognition a donc un périmètre très vaste, et présente en son sein de belles opportunités de travaux interdisciplinaires.

L'Intelligence Artificielle s'intéresse à la génération d'un comportement dit "Intelligent". Le concept d'intelligence est cependant particulièrement difficile à définir. Plusieurs essais de définition ont été avancés sans jamais être totalement satisfaisants. La première proposition de définition pour Turing présentée dans [Turing50], a été de confronter un homme à une machine par l'intermédiaire d'un canal de communication restreint. Si l'homme ne parvient pas à identifier son interlocuteur comme étant une machine, le logiciel est alors considéré comme intelligent. Cette définition peut aujourd'hui être remise en question, depuis par exemple le développement de Deep Blue qui a vaincu le champion du monde d'échec Garry Kasparov en mai 1997. En effet, Deep

Blue combine des capacités de calcul et des algorithmes spécifiques plus performants que l'humain sur la tâche sélectionnée, mais répond-il à des critères d'intelligence, et si oui lesquels ?

Une des ambiguïtés dans la quête d'une définition satisfaisante provient aussi du rapprochement naturel qui s'opère avec le plus proche concept connu, bien que plus vaste : l'intelligence humaine, elle aussi difficile à circonscrire. On pourra par exemple définir l'IA comme un "ensemble de théories et de techniques mises en œuvre pour réaliser des machines capables de simuler l'intelligence humaine"¹. L'objet de cette recherche s'autorise donc à utiliser l'ensemble des supports du domaine de l'Informatique aussi bien matériels avec toujours plus de capacités de calculs, qu'algorithmiques avec des modèles de plus en plus élaborés.

Lee Sedol, considéré alors comme le plus fort joueur du jeu de Go, a été battu par un nouveau type d'IA en mars 2016. Le jeu de Go est connu pour sa combinatoire supérieure à celle du jeu d'Échecs, mais également pour ses difficultés spécifiques dans l'estimation des positions. Pour permettre cela, la capacité de calcul a bien progressé, les méthodes à base de réseaux de neurones l'exploitent pour apprendre à partir de grandes bases de données de parties de professionnels, puis ensuite simplement par simulation avec Alpha Zero qui apprend par renforcement sans exploiter de données d'entraînement. Il est désormais possible d'estimer des positions bien plus efficacement que cela a pu être fait avec différentes règles issues de systèmes experts. Un système peut désormais apprendre efficacement par simulation. Qu'en est-il de l'intelligence ?

Ces dernières années, l'agent conversationnel ChatGPT a beaucoup impressionné par ses capacités à s'appropriier de très larges corpus textuels, afin d'être consulté comme un modèle prédictif (par *prompting*), c'est-à-dire comme un modèle qui prédit du texte plausible statistiquement sachant le texte d'amorçage entré souvent sous forme de questions. Les capacités à agencer les mots sont très impressionnantes, malgré par exemple différents biais pour la plupart issus du corpus d'apprentissage. ChatGPT reste limité par le corpus qu'il exploite, et ne parvient pas toujours à enrichir celui-ci et à s'appropriier pleinement les concepts qu'il contient. Nous verrons par exemple que les larges modèles de langue ont des difficultés à réaliser certaines inférences complexes et qui nécessitent plusieurs étapes. Un enfant apprend à poser et exécuter des multiplications, et donc s'appropriier le raisonnement, alors que l'approche probabiliste inhérente au fonctionnement des larges modèles de langue ne fonctionne pas lorsque la combinaison opérandes-résultat n'a jamais été rencontrée. Ceci souligne donc des difficultés de généralisation, mais sur un aspect algorithmiques. Pourtant, les agents conversationnels sont désormais capables de s'appropriier des concepts sémantiques grâce aux propriétés apportées par les plongements lexicaux, mais peut-on parler d'intelligence ?

Chacun pourra estimer ce qui peut correspondre ou non à une intelligence artificielle avec prudence.

1. Définition de l'intelligence proposée par le dictionnaire Larousse

Dans le cadre de mon travail de recherche, j'étudie la capacité de ces systèmes à développer de l'autonomie, c'est-à-dire à pouvoir interagir en autonomie dans un environnement inconnu et non-stationnaire afin d'atteindre différents objectifs. Mes recherches au sein de l'équipe SyCoSMA (Systèmes Cognitifs et Systèmes Multi-Agents) du LIRIS s'inscrivent ainsi dans la perspective de donner plus d'autonomie à nos IA.

Je positionne ainsi le concept d'autonomie comme essentiel à toute forme d'émulation d'intelligence. C'est bien sur ce point que l'on peut déprécier les systèmes actuels qui ne peuvent être considérés comme étant pleinement autonomes. Un système autonome ne devrait pas seulement résoudre des problèmes individuellement, mais devrait en effet être doté de capacités de résolution de problèmes qu'il découvre sans connaissances préalables dans un environnement ouvert. Idéalement, un tel système permettrait de répondre à des besoins applicatifs de plus en plus nombreux, spécifiques et dynamiques.

Afin de désormais considérer le système dans son environnement, on pourra s'intéresser au paradigme d'*embodiment* ou de cognition incarnée qui considère la cognition comme résultat de l'interaction d'un agent et de son environnement. Telle que le propose Piaget dans ses travaux sur le développement cognitif de l'enfant [Piaget55], la cognition doit littéralement être construite et doit être considérée en tant que processus dynamique. Ce positionnement définit et motive pour le domaine de l'Informatique le développement d'Intelligences Artificielles dites Développementales (IAD). Dans ce contexte, j'étudie les procédés permettant de faire évoluer l'interprétation courante du monde par un système apprenant, en intégrant le fait que le système a une perception fondée sur ses propres structures cognitives (le système étant autonome).

Cette perspective théorique de la cognition soulève deux thèmes et défis centraux que j'adresse ou qui participent à mes réflexions dans mes recherches :

- **Le problème de la représentation interne** que le système exploite (du monde, de l'individu lui-même ainsi que des processus visant à faire évoluer la cognition (*praxis* de second ordre). C'est un problème que j'aborde dans les différents projets et applications que je porte ou auxquels je suis associé.
- **Le problème de la dynamique d'apprentissage**. Le problème de l'amorce (*bootstrap*), central de par l'aspect incrémental de la cognition. Il s'agit d'un problème de type oeuf/poule complexe. En effet, on souhaite élaborer des structures utiles à la perception de régularités observées dans le monde, sachant que les capteurs des éléments constitutifs de ces régularités sont motivés par la reconnaissance presque préalable de ces mêmes régularités.

1.2 Organisation du mémoire

Dans ce mémoire, j'identifie différents axes de recherche et des verrous scientifiques, ces axes sont utiles à l'élaboration d'un système que l'on souhaite capable d'interagir

de façon autonome avec son environnement.

Le positionnement sous-jacent à ce travail est introduit dans le chapitre 2. Ce chapitre définit la problématique de recherche en positionnant le support d'inscription de connaissances et la dynamiques d'apprentissage comme essentielles au développement d'un système autonome.

Mon parcours de recherche m'a amené à travailler sur des applications variées (affectation de ressources, jeux combinatoires, régulation du trafic, intelligence ambiante, traitement naturel du langage, reconnaissance d'images, etc.), ce qui m'a donné la possibilité d'expérimenter différentes méthodes informatiques (Systèmes multi-agents, arbres de recherche combinatoires, réseaux de neurones, apprentissage supervisé ou non, etc.).

Chacun des projets met l'accent sur diverses propriétés utiles à l'élaboration d'une IA autonome. Parfois, l'effort a porté sur la capacité du système à mémoriser des connaissances complexes de différentes natures. Ceci m'a parfois amené à sortir du cadre strict d'autonomie forte, en fournissant artificiellement au système l'information à acquérir, à travers un signal de supervision. Pour d'autres projets, l'accent a porté sur la dynamique d'apprentissage en interaction avec l'environnement et sans supervision, avec ainsi la notion d'autonomie du système mise en exergue par rapport à sa capacité à trouver un *feedback* pour son apprentissage, par un apprentissage non supervisé ou une motivation définie par le système lui-même ou intrinsèque.

L'ensemble des projets sont présentés dans le chapitre 3. Les projets sont résumés de façon à suivre la démarche des études et faciliter la lecture (motivation, modélisation, résultats), sans nécessairement reporter l'ensemble des expérimentations menées. Les articles détaillés associés sont précisés dans chacune des parties.

Nous voyons en section 3.1 un modèle décentralisé pour optimiser la recherche de documents dans un réseau pair-à-pair. Dans le cadre de la découverte d'environnement par un système multi-agents, différents mécanismes de contrôle de fiabilité sont décrits en section 3.2 pour être appliqués à la découverte d'environnement en situation de crise ou à la régulation du trafic dans le cadre de véhicules autonomes.

Nous nous intéressons ensuite en section 3.3 à la construction par un système multi-agents de motifs sensori-moteurs de bas niveau, pour le domaine de l'intelligence ambiante. Dans ce projet, on s'intéresse ainsi à capter les régularités comportementales de l'environnement. Nous nous intéressons dans la section 3.4 aux jeux combinatoires et proposons des améliorations pour les méthodes de résolution basées sur MCTS (*Monte-Carlo Tree Search*). Dans ce contexte, l'environnement est simulé (le jeu de Go ou jeu du Clobber), nous nous intéressons aux stratégies de mutualisation de la connaissance pour MCTS, mais également à la réification pour MCTS d'objectifs stratégiques inter-médiaires pour structurer la découverte de l'environnement.

La section 3.5 s'intéresse également à l'apprentissage par renforcement, pour la régulation de trafic autoroutier. L'environnement est simulé, différents mécanismes sont

étudiés sur l'aspect perceptif et pour le contrôle exercé par le système de régulation.

La section 3.6 s'intéresse à un apprentissage sémantique de plus haut niveau d'objets à partir d'un flux vidéo sans supervision. L'exploitation des régularités spatio-temporelles, le rôle de l'action et la discrimination statistique sont abordés dans cette section.

Pour terminer, la section 3.7 s'intéresse aux traitements algorithmiques comme le sujet de l'apprentissage, et comme une brique utile au développement d'une IA. La proposition faite permet de limiter les problèmes d'apprentissage d'algorithmes par des réseaux de neurones *end-to-end*.

Le chapitre 4 discute des différents projets parcourus, et trace les grandes lignes de mon projet de recherche dans la continuité du travail engagé.

2

Positionnement

Dans cette partie, après avoir énoncé la problématique de recherche, je présente différents axes qui permettent de préciser mon positionnement. Lorsque cela est possible, j'illustre les différents axes sur les projets sur lesquels j'ai été amené à travailler.

J'ai en effet pu explorer un large éventail de méthodes (algorithme de recherche opérationnelle MCTS (*Monte Carlo Tree Search*), réseaux de neurones profonds, réseaux de neurones à base de prototypes, apprentissage de schémas, systèmes multi-agents, etc.) sur lesquels je m'appuie pour identifier des axes de conception pour l'élaboration d'une IA développementale.

2.1 Problématique de recherche

Le sujet principal de la recherche que j'entreprends s'intéresse à l'élaboration par un système informatique d'une représentation interne du monde (qui correspond à l'environnement du système) utile à l'accomplissement d'un ensemble de tâches. Ce travail s'intéresse ainsi à la constitution d'une structure, ainsi qu'à la constitution de traitements associés que l'on peut englober indifféremment dans la notion de représentation.

On s'intéresse à des environnements ouverts qui ne peuvent être connus exhaustivement préalablement à la mise en service du système. Ainsi, le système doit nécessairement développer sa représentation de son environnement ou de son monde de façon autonome.

Le cadre générique ainsi défini constitue un défi ambitieux sur le long terme. Nous verrons qu'il soulève des problèmes énigmatiques pour lesquels plus on les connaît, plus ils semblent hors d'atteinte, tel que par exemple le problème de l'amorçage d'un tel système. Afin d'aborder ce problème, j'essaie d'identifier des propriétés utiles à la constitution d'un tel système et de proposer itérativement différentes contributions.

Selon cette démarche, l'accent est souvent mis sur certaines propriétés au détriment des autres. C'est d'ailleurs ce qui est généralement fait sachant que le problème abordé reste actuellement non résolu dans sa globalité.

J'adopte ainsi volontiers la possibilité de simuler un environnement en prêtant attention aux éventuelles, et concrètement fréquentes, distorsions occasionnées interdisant la mise en application réelle (*reality gap*), ce qui est par exemple parfois le cas pour les approches d'apprentissage par renforcement pour lesquelles l'environnement peut être simulé.

Ce travail s'inscrit dans une perspective constructiviste, ou développementale, selon laquelle le système perçoit à travers la représentation du monde dont il dispose, le monde lui-même étant inaccessible directement. Dans le cadre de ce positionnement, l'environnement est d'une part inaccessible car il n'est pas connu à priori, mais également parce qu'il n'existe pas de représentation absolue de celui-ci tel que le propose le constructivisme radical [Von Glasersfeld84] dans la lignée du relativisme philosophique. La représentation recherchée est efficace et dédiée pour un usage donné, différents systèmes pouvant de par leurs caractéristiques propres (capteurs, effecteurs, objectifs, capacités représentationnelles) élaborer des représentations distinctes leur permettant d'exploiter toutes les potentialités des objets du monde ou affordances.

La perspective constructiviste adoptée pour ce mémoire fait l'hypothèse que cette représentation doit évoluer et accroître sa complexité itérativement, de la même façon qu'un enfant apprend des concepts simples utiles à l'acquisition de concepts plus élaborés. Ceci soulève différents défis et objets d'étude :

- le cheminement de construction de cette connaissance à définir qui peut être simplement guidé par une motivation intrinsèque et des mécanismes de curiosité [Oudeyer07] déclenchant les différentes itérations avec l'environnement (perceptions et actions). Ce cheminement s'appuie sur les deux objets d'étude suivants.
- **le support informatique** permettant de mémoriser la connaissance ou représentation (données et processus), discuté en section 2.2.
- **la dynamique d'un tel système**, son amorce et le maintien d'un mécanisme rétroactif entre données et processus à la fois libre de pouvoir s'adapter, mais qui manifeste également des propriétés de stabilité et de convergence, discutées en section 2.3.

2.2 Support d'inscription de connaissances

De très nombreuses méthodes permettent d'inscrire des connaissances sur un support informatique, des plus simples et intelligibles, par exemple une simple base de données ou même un graphe étiqueté, aux plus élaborées et parfois qualifiées de boîtes noires tels que les biens connus réseaux de neurones. Il est bien entendu difficile d'envisager le support sans le processus qui l'exploite et lui permet d'évoluer, je décris en section 2.3 une façon d'envisager ce couplage.

Dans cette section je m'intéresse tout d'abord à la nature du support d'inscription de connaissances pour un système apprenant autonome, à ses capacités de généralisation comme essentielles à son apprentissage, mais également à la malléabilité du support d'apprentissage pour un apprentissage incrémental en environnement ouvert.

2.2.1 Un support de connaissance agnostique et endogène

Certainement influencé par mon approche initiale de l'IA par systèmes multi-agents réactifs¹, l'idée selon laquelle les approches les plus efficaces proviendraient de leur simplicité m'intrigue, leur complexité et richesse résultant de leur dynamique. Quels défis alors pour que le système autonome puisse dénouer ses liens avec les experts qui le contrôlent et le programment, à la manière d'un Geppetto apprenenti!

La structure serait ainsi agnostique, car elle ne présumerait pas de connaissances préalables sur lesquelles s'aligner, et endogène, car issue de la propre activité du système.

Je me démarque ainsi des approches expertes selon lesquelles la capacité de résolution des problèmes découle d'un ensemble de règles définies a priori. L'idée sous-jacente étant d'aller au-delà la conception de la cognition comme simple processus de traitement de l'information dans la lignée du paradigme cognitiviste. À travers la notion de règles se retrouve tout un ensemble de paramètres que nous, informaticiens, souvent sans alternative connue, sommes amenés à envisager à regret par la complexité qu'ils occasionnent pour chacune de nos conceptions.

Parmi les différents supports envisageables, les règles expertes ont la qualité d'être intelligibles et donc de favoriser l'interopérabilité avec l'expert qui le programme. Afin de matérialiser une représentation du monde pour le système, une possibilité consiste ainsi à décrire de façon experte le problème, le logiciel exécutant alors un ensemble de règles ou inférences pour atteindre ses objectifs.

1. Au cours de mon travail de thèse sur la résolution décentralisée de problèmes de partage de ressources [Armetta06], selon une approche par système multi-agents réactifs, des agents relativement simples interagissent localement et ont pour objectif de résoudre collectivement les problèmes complexes auxquels ils sont confrontés.

Ce type d'approche n'est pas dépourvu de qualités, et se prêtera très bien à des cas d'utilisation de complexité d'expression restreinte selon lesquels le comportement de l'environnement ou du problème peut s'exprimer à travers un ensemble de règles contenues en nombre et suffisamment isolées dans leurs conditions d'application. Le problème doit par ailleurs être préalablement connu.

Par exemple, dans le cadre du développement d'un agent conversationnel orienté but, après avoir identifié l'intention de l'utilisateur (ce qui correspond au développement d'un composant NLU (*Natural Language Understanding*)), on pourra dérouler de façon séquentielle le traitement associé à ses besoins, tel que préciser un produit et une adresse de livraison. Ces cheminements préétablis peuvent s'apparenter à une modélisation experte. Mais comment appréhender alors les interactions spontanées des utilisateurs non cartographiées ?

Afin de traiter de tels problèmes moins linéaires, les connaissances explicites peuvent être étendues par inférences, tel que cela a été fait en programmation logique avec par exemple le langage Prolog. Suivant cette même inspiration, mais selon une perspective multi-agents et distribuée, les agents BDI [Rao95] modélisent des connaissances et des règles de comportement visant à réaliser des inférences.

Une des limitations de ce type d'approche provient du manque de lisibilité pour l'expert lorsque les règles se chevauchent et deviennent trop nombreuses pour être appréhendées de l'extérieur du système, c'est-à-dire par l'expert qui les met en œuvre.

L'approche cognitive n'a par exemple pas permis de vaincre les meilleurs joueurs lors des premiers moteurs de calcul à base de règles expertes pour le jeu de Go. Malgré le support de très bons joueurs experts, seul un niveau de débutant a pu ainsi être atteint, car l'explosion du nombre de règles visant à décrire la complexité et la diversité du jeu créait des interférences difficiles à corriger et formaliser manuellement. Finalement, c'est bien un support endogène, à travers les réseaux de neurones, qui fournissent aujourd'hui de très bons résultats.

Si un expert enrichit manuellement les connaissances du système, la difficulté revient de nouveau au fait que le système ainsi formé emprunte «ces» connaissances, et ne se les approprie pas par son expérience selon une construction subjective itérative, ce qui correspond à un manque d'ancrage des connaissances manipulées, tel que le décrit l'expérience de la chambre chinoise [Searle80]. La représentation apportée est constituée de symboles et de règles exogènes au système.

Suivant le cadre présenté, un système ainsi suppléé par une intelligence extérieure est incomplet du point de vue de son autonomie, et donc de son "intelligence".

L'approche développementale ou constructiviste dans laquelle je m'inscris **privilegiera l'auto-développement du support de représentation**. Après avoir motivé la qualité de la structure idéale, il faut reconnaître le manque de systèmes artificiels manifestant pleinement les qualités recherchées.

Une des inspirations parmi les plus convaincantes pour le domaine est de type bio-inspirée. C'est par exemple le cas dans le cadre d'une activité décentralisée, pour la phéromone comme support déposée dans l'environnement par les insectes sociaux tels que les termites permettant de trouver les plus courts chemins vers les sources de nourriture ou de construire et réparer la termitière [Grassé59]. C'est suivant cette inspiration que j'ai pu proposer une approche également basée sur la stigmergie, pour le référencement de documents dans un réseau pair-à-pair, en section 3.1.

Sachant que la structure sera manipulée par le système, elle pourra ainsi être non intelligible pour un expert. C'est le cas pour l'intelligence humaine, si l'on considère qu'elle s'exprime à travers leurs systèmes neuronaux, ceux-ci sont encore partiellement compris. Il est encore difficile de les interpréter, de la même façon qu'il est difficile de comprendre les algorithmes exécutés par un ordinateur équipé d'un simple volt-mètre. Les réseaux de neurones informatiques, plus simples que le modèle biologique, semblent cependant de bons candidats comme support d'inscription, avec notamment leurs capacités de généralisation appropriées telles que précisées dans la section suivante.

2.2.2 Capacités de généralisation

Une des propriétés recherchées correspond à la forte expressivité de la structure d'accueil des connaissances. Dans cette expressivité, pour une IA développementale généraliste, la possibilité de représenter et positionner des concepts les uns par rapport aux autres est centrale.

Une caractéristique incontournable des réseaux de neurones correspond à leurs capacités à abstraire de l'information sémantique ou à généraliser, et ainsi définir les positions de ces concepts les uns par rapport aux autres.

Si l'on se réfère aux travaux de Piaget [Piaget54], le processus d'apprentissage chez l'enfant intervient à travers un processus d'assimilation (reconnaître un objet comme appartenant à un concept déjà existant qui doit être mis à jour) ou d'accommodation (création d'un concept nouveau). Le processus d'assimilation repose sur une capacité à généraliser les objets enregistrés sur le support choisi.

J'ai eu l'opportunité au cours de ces dernières années d'éprouver l'utilisation des réseaux de neurones, comme au cours du projet présenté en section 3.6 dans le cadre d'un problème de classification par apprentissage continu. Je m'intéresse par exemple à identifier des objets à partir d'un flux vidéo. L'objectif étant alors de créer des catégories d'objets (ou classes) à partir de leur apparence, afin de les ancrer ontologiquement et fonder par la suite² un processus de perception pour le système.

2. Nous précisons dans la section 2.3 que cette perception ne devrait pas dans l'idéal être découplée et antérieure à l'activité du système

Un support de représentation idéal devrait permettre d'appliquer la généralisation à différents niveaux, du plus abstrait au plus concret. Il devrait également permettre de «projeter» la représentation suivant les contextes d'application tel qu'abordé en perspective de ce mémoire. Le concept «chaise» peut ainsi être associé aux concepts plus généraux ameublement ou décoration. Les larges modèles de langue sont dans une certaine mesure capables de matérialiser et apprendre ces représentations multidimensionnelles, par l'utilisation de modèles attentionnels qui permettent de sélectionner le sens des concepts pertinents pour le contexte d'utilisation [Vaswani17a].

La capacité de généralisation des réseaux de neurones peut également être envisagée d'un point de vue algorithmique [Graves14a]. C'est ce que j'étudie en section 3.7, en faisant des propositions sur les problèmes d'entraînements spécifiques à ce type d'apprentissage.

La section suivante s'intéresse au support d'apprentissage du point de vue de sa dynamique, dans un environnement ouvert.

2.2.3 Un support d'apprentissage résilient

La direction que je donne à ma recherche ne s'intéresse pourtant pas uniquement au problème d'apprentissage du point de vue *machine learning*, mais bien à ces méthodes comme pour le support d'apprentissage vers plus d'autonomie.

L'apprentissage en *machine learning* nécessite généralement une infrastructure complexe incluant des données de qualité représentatives du problème à traiter, et repose souvent sur des techniques de prétraitement et d'optimisation. L'élément déterminant du succès de ces méthodes repose sur les jeux de données disponibles.

Les modèles de *machine learning* présupposent souvent le fait que les données d'apprentissage soient i.i.d. (indépendantes et identiquement distribuées) ce qui n'est généralement pas le cas. Ces données doivent être représentatives, en quantité équilibrée et idéalement présentées au système apprenant de façon aléatoire.

Dans le travail présenté en section 3.6.3 sur l'apprentissage continu d'objets précédemment évoqué, on s'intéresse à s'affranchir de l'aspect structuré des données d'apprentissage présentées au système, tout en maintenant de bonnes capacités d'assimilation et d'accommodation sans supervision, c'est-à-dire sans étiquetage préalable du corpus. L'organisation faite des objets repose sur les similarités d'aspects observées.

On pourra reprocher à un tel apprentissage de construire une représentation à partir de ressemblances qui peuvent être qualifiées d'arbitraires, car on pourra imaginer discriminer les objets par rapport à leur couleur, taille, contraste ... et pourquoi pas leur fonction pour le système. On évalue en effet le système pour sa capacité à retrouver les classes d'objets identifiés dans des *datasets* du domaine du *machine learning*, mais plusieurs organisations concurrentes sont envisageables, chaque objet pourrait appartenir

à plusieurs classes. Quelle représentation résiliente se prêterait le plus à autoriser cette expressivité multiple ?

La section suivante s'intéresse à préciser la qualité de ce qui doit caractériser un objet, d'un point de vue dynamique et interactionnel et comme un complément indispensable au support de représentation d'une IA autonome.

2.3 Processus d'apprentissage

La théorie constructiviste de l'apprentissage [Piaget54] propose une vision de l'apprentissage comme un processus actif de construction et d'adaptation d'une représentation par le sujet, en interaction avec son environnement, plutôt que comme la simple acquisition d'un modèle figé du monde. Ce mode d'apprentissage est un cadre idéal, il repose sur une capacité à accueillir des connaissances et un processus capable de l'exploiter et le faire évoluer. Le travail de thèse présenté en section 3.3 illustre et détaille les fondements de ce paradigme d'apprentissage pour un habitat connecté. Celui-ci est présenté comme un robot autonome percevant et agissant à travers différents capteurs brutes et effecteurs afin d'identifier les régularités de son environnement.

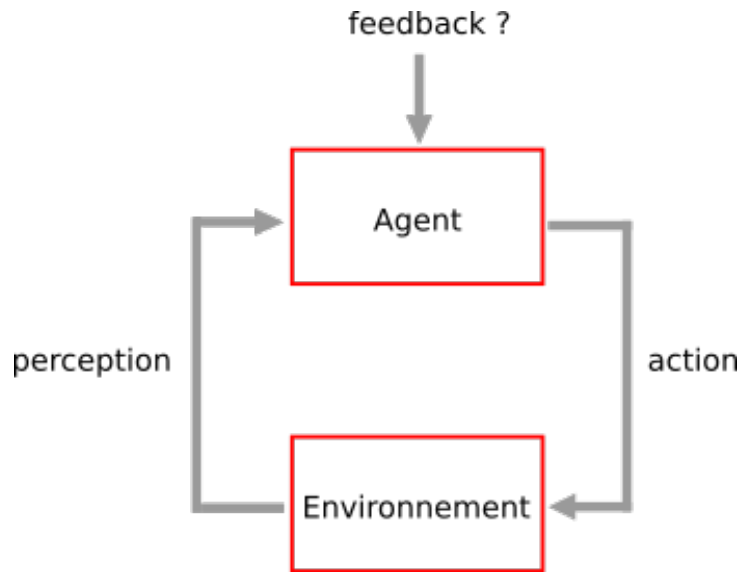
Dans le cadre de mes recherches, je m'intéresse en section 2.3.1 au système apprenant comme interagissant avec son environnement à travers ses perceptions et ses actions. J'ai pu étudier les capacités d'acquisition de connaissances selon différents modes présentés en section 2.3.2. L'apprentissage et sa convergence sont discutés en section 2.3.3.

2.3.1 Perception et action

Un agent apprenant structure sa perception et ses actions à travers une boucle d'interactions sensori-motrices. Il va ainsi faire évoluer la perception qu'il a de son environnement tout en apprenant les actions lui permettant d'atteindre ses objectifs : survie, divers gains ou simple satisfaction d'apprendre diverses régularités par curiosité. Cette boucle d'interaction initialement décrite du point de vue de la psychologie (premier stade d'apprentissage chez l'enfant [Piaget54]), a été reprise par les méthodes d'apprentissage par renforcement en informatique en introduisant une fonction de récompense ou *feedback* [Sutton98].

On présente ainsi dans cette section l'apprentissage comme un enchaînement de perceptions et d'actions, alimentées par un *feedback*, c'est-à-dire une valeur guide de récompense de différents types suivant les approches tel que présenté sur la figure 2.1.

D'une part, le système apprend à percevoir, mais il apprend également une politique d'actions appliquée selon ce qui est perçu. Tel qu'illustré sur la figure 2.2, on comprend qu'apprendre ce couplage n'est pas trivial, dans le cas le plus général ceci implique une perception brute de signaux continus et des actions multiples, en respec-

FIGURE 2.1 – Agent - Environnement - *Feedback*

tant une contextualisation liée à l'historique des perceptions et actions (ce cadre d'étude correspond au projet présenté en section 3.3).

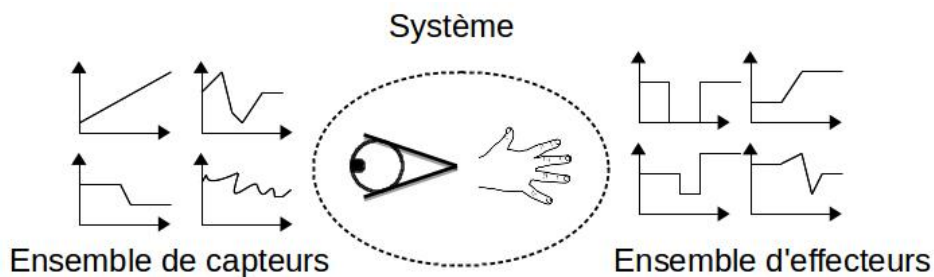


FIGURE 2.2 – Du signal (capteurs) au contrôle (effecteurs)

Dans le cas d'un réseau de neurones qui prend en entrée les différents capteurs de l'environnement et qui apprend une politique d'actions déclenchées à sa sortie il est pourtant difficile ou même inutile de différencier le processus de perception du processus d'action. Ce couplage fort peut même lui-même intervenir comme élément perceptif de second ordre en tant qu'invariant sensori-moteur, tel que le proposent certaines théories de contingences sensori-motrices [O'Regan01b]. Pour illustration, le concept de «ligne» pour un sujet est alors défini selon cette théorie par l'interaction avec la ligne, caractérisée par une régularité perception/action : perception d'un point sur la ligne, déplacement du point de perception sur la ligne, perception du point suivant,

etc. Ces théories soulignent la fonction centrale portée par la dynamique de perception et d'action pour un agent autonome.

2.3.2 Le *feedback*

Au cours de son apprentissage, le système apprend à partir de ses interactions avec son environnement. Si le système parvient à estimer de façon correcte la "direction" à suivre pour parfaire son apprentissage, alors il va s'améliorer, parfois sur le long terme, dans la résolution de la tâche qu'il adresse.

Feedback fixe et supervisé

Pour un réseau de neurones, on applique une correction aux différents poids du réseau afin qu'il retrouve à sa sortie des étiquettes connues pour les objets placés à son entrée. L'apprentissage est alors supervisé.

Pour le projet présenté dans [Faugier24], nous réalisons par exemple un apprentissage de finition sur un large modèle de langue, afin de lui permettre d'inférer le type de relation (*support* ou *attack*) entre arguments formulés en langage naturel. L'approche utilise ainsi un jeu d'arguments dont les relations sont connues et étiquetées, et qui peuvent ainsi être utilisées comme signal de supervision.

Feedback calculé

Pour l'apprentissage par renforcement, le problème est cette fois souvent simulé et le retour est cette fois le fruit de la simulation bien qu'il soit possible d'appliquer la méthode en environnement réel. Dans la section 3.4, l'algorithme MCTS permet d'estimer la qualité des coups à jouer pour le jeu de Go par simulations Monte-Carlo. Pour cette application, l'accent est porté sur la qualité combinatoire du domaine applicatif, et donc sur une stratégie spécifique pour le calcul du *feedback* (méthodes d'exploration et d'exploitation avec minimisation du regret sur le long terme).

Feedback par immersion

Dans le cadre d'un apprentissage décentralisé, le *feedback* issu de l'activité globale du système est perçu localement.

C'est le cas par exemple sur un réseau pair-à-pair présenté en 3.1, la requête parcourt le réseau suivant les marquages phéromonaux réalisés, et perçoit ainsi localement l'activité du collectif. L'immersion pourrait être considérée comme un processus inverse de l'émergence. Alors que l'émergence concerne l'apparition de propriétés nouvelles à partir d'interactions simples, l'immersion décrit comment des propriétés de

niveau supérieur influencent ou se manifestent dans des composants de niveau inférieur.

Motivation intrinsèque

Dans le cas général pour lequel il n'existe pas d'oracle ou de fonction objectif préalablement définie, ce qui correspond au contexte d'autonomie le plus fort. Le système a alors pour objectif d'acquérir une représentation fiable du monde constituée des régularités sensori-motrices observées (des schémas de régularité entre le contexte perçu, l'action et le résultat constaté), dont l'élaboration est motivée par sa simple curiosité [Oudeyer16, Georgeon12]. Dans la façon dont l'effort d'apprentissage est orienté vers les configurations mal comprises, cet apprentissage se rapproche de l'apprentissage actif du *machine learning* [Settles09].

Suivant les applications, il sera parfois difficile d'obtenir ou de matérialiser ce *feedback*. C'est particulièrement le cas pour le projet présenté en section 3.3.

Pour ce projet, une maison équipée de capteur (température, humidité, chauffage, etc.) est assimilée à un robot qui découvre et apprend le comportement de son environnement à partir d'un ensemble de signaux continus. Par exemple, il apprend le fait qu'ouvrir la fenêtre entraîne une baisse de température en hiver. Ce problème correspond par exemple aux premiers stades de développement pour un enfant qui comprend petit à petit la relation de cause à effet entre le fait de lancer un objet et le son qui s'en suit presque systématiquement.

Dans une certaine mesure, les larges modèles de langue extraient également au cours de leur apprentissage des corpus mis à leur disposition des régularités d'enchaînements de mots, de façon auto supervisée et sans étiquetage préalable. À partir d'un contexte constitué par un ensemble de mots, ces systèmes apprennent les mots suivants dont l'apparition est la plus probable.

2.3.3 Apprentissage et convergence

Afin d'améliorer l'efficacité ou simplement de permettre un apprentissage, on pourra appliquer différentes stratégies qui sont discutées dans cette section. Ces stratégies peuvent être spécifiques au support d'apprentissage choisi. C'est le cas de l'apprentissage incrémental par transfert pour les réseaux de neurones présenté ci-dessous. Le problème central de l'initialisation ou *bootstrap* ainsi que les enjeux liés à la qualité des signaux perçus sont ensuite discutés.

Apprentissage incrémental et transfert

Dans le cadre d'un apprentissage constructiviste ou développemental, les connaissances acquises reposent sur les connaissances plus simples de niveau inférieur ou préalablement acquises.

En utilisant un support d'apprentissage tel qu'un réseau de neurones, on s'intéresse alors à la capacité de transfert de l'apprentissage entre tâches de complexité croissante. Dans la section 3.7, on montre que l'apprentissage des multiplications par un réseau de neurones pose des problèmes d'apprentissage ou d'entraînement. L'apprentissage peut alors être stimulé par l'introduction de tâches algorithmiques intermédiaires utiles à l'opération cible.

Initialisation et *bootstrap*

Dans le contexte d'un apprentissage agnostique incrémental, le système n'a aucune connaissance préalable, et apprend les régularités de son environnement à différents niveaux d'abstraction.

Un motif est pertinent s'il contribue à la création de nouveaux motifs, de manière à permettre la construction d'une représentation qui les justifie. En d'autres termes, l'agent doit apprendre à percevoir efficacement afin d'être capable d'apprendre davantage. Trouver un contexte pour caractériser l'association entre une action et un résultat nécessite de connaître préalablement cette association, mais apprendre cette association est difficile sans connaître le contexte.

On peut illustrer ce phénomène pour le problème d'intelligence ambiante abordé en section 3.3. Le système perçoit à travers des capteurs de signaux continus l'état de l'habitat (angle d'ouverture des fenêtres, température des pièces, saison). Afin de percevoir une régularité telle que la température d'une pièce baisse après quelque temps lorsque les fenêtres sont ouvertes en hiver, le système doit pouvoir contextualiser, discrétiser son expérience : fenêtre ouverte/fermée, hivers/été, délais (la relation n'est pas directe). Ces perceptions étant maintenues et développées grâce à leur implication dans les régularités comprises, elles présupposent l'existence de la régularité. Les motifs à apprendre peuvent être d'une complexité très rapidement bien supérieure. Étiqueter par exemple un objet à partir d'un flux visuel continu comme pouvant servir de récipient et être utilisé pour se désaltérer en suivant une suite d'action spécifique dans un contexte favorable engage des processus particulièrement complexes. Une étude sur le rôle de l'action dans l'apprentissage perceptif est présentée en section 3.6.2.

Qualité du signal perçu

Le signal d'apprentissage n'est pour un grand nombre d'applications en environnement ouvert pas toujours clair, fiable, synchrone, ni même complet.

Dans son modèle inspiré du paradigme constructiviste, [Drescher91b] souhaite apprendre les régularités du monde, de type "situation, action \rightarrow prédiction". L'étude de *schémas* proposée par [Drescher91b] identifie bien les difficultés rencontrées : variabilité des effets d'une action dans des circonstances différentes qui doivent être qualifiées, un même résultat peut avoir des causes diverses, le résultat peut être perçu parmi un ensemble d'évènements proches dans le temps, mais sans rapport avec l'action concernée.

Pour le domaine de l'apprentissage par renforcement [Sutton18], on souhaite apprendre une politique permettant de trouver l'action optimale dans un état donné et permettant de maximiser une fonction de récompense définie. Sachant que l'état n'est pas parfaitement connu, on qualifie ce type de problèmes comme partiellement observable (*Partially Observable Markov Decision Process*), avec un apprentissage qui peut en conséquence s'avérer long et difficile.

Pour beaucoup de projets décrits dans le mémoire, la qualité du signal perçu et la façon de l'exploiter est un élément déterminant.

Dans la thèse présentée sur l'apprentissage des meilleurs coups pour le jeu de Go en section 3.4, on exploite l'algorithme MCTS (Monté Carlo Tree Search) qui repose sur une estimation des positions de jeu par échantillonnages Monte-Carlo. L'effort d'estimation est pondéré par la qualité estimée des différentes parties représentées dans l'arbre de recherche. Dans la thèse qui s'intéresse à la régulation de trafic routier pour les voies d'insertion présentée en section 3.5, la complexité du système formé rend difficile l'évaluation des causalités entre la politique de régulation du flux de véhicules et les résultats perçus avec souvent des latences significatives.

2.4 Domaines de contributions

Ma recherche s'intéresse à l'apprentissage d'une représentation de l'environnement et à la dynamique associée à cet apprentissage. Les projets de recherche présentés dans le chapitre suivant se positionnent parmi les domaines de recherche ci-dessous (voir figure 2.3) :

- **Sciences cognitives** : Différents paradigmes d'apprentissage motivent chacun des modèles proposés. Dans le cadre de ce travail, on définit l'aspect cognitif comme ce qui permet au système d'acquérir de façon autonome une représentation de son environnement. Cette construction pourra s'appuyer sur l'organisation résultant d'interactions entre agents dans un contexte décentralisé. On s'intéresse également, dans la cadre du développement d'un agent apprenant, à l'apprentissage sensori-moteur, qui incite à considérer un agent incarné en interaction avec son environnement selon une perspective d'apprentissage constructiviste ou développementale. *Dans le cadre de ma recherche, ce domaine a toujours fait partie de mes motivations scientifiques.*

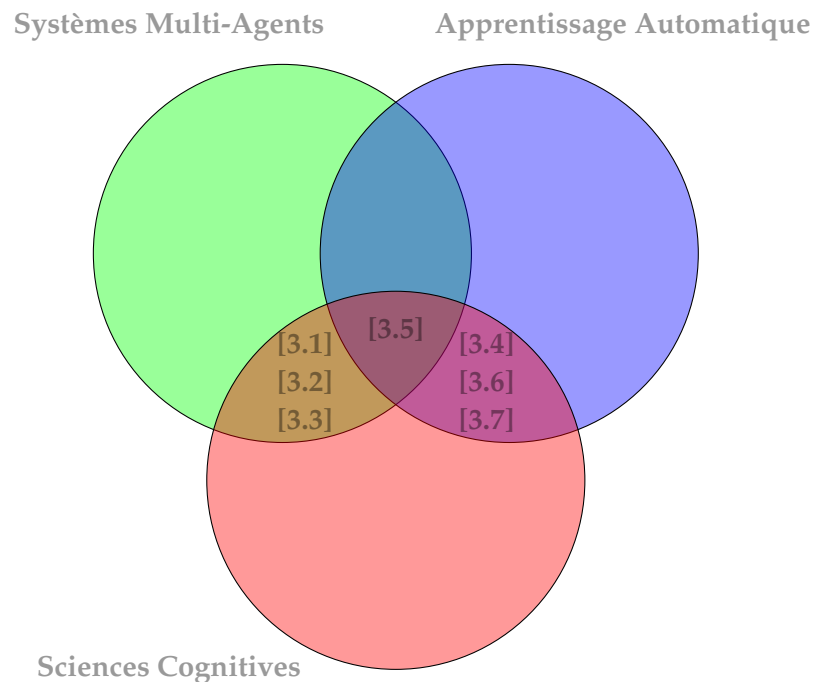


FIGURE 2.3 – Domaines de contributions.

- [3.1] Auto-organisation de topologie pour le partage de données d'un réseau pair-à-pair
- [3.2] Découverte d'environnement décentralisé et mécanismes de contrôle de fiabilité
- [3.3] Apprentissage de régularités en environnement continu et Intelligence Ambiante
- [3.4] Extensions de MCTS pour les jeux combinatoires
- [3.5] Apprentissage et régulation du trafic routier
- [3.6] Apprentissage continu non supervisé d'objets à partir d'un flux vidéo
- [3.7] IA et apprentissage d'algorithmes

- **Systèmes multi-agents** : Le système apprenant est alors constitué ou met en œuvre différents agents interagissant dans un environnement distribué. *J'ai tout d'abord abordé l'apprentissage uniquement de façon décentralisée par ces systèmes multi-agents, puis j'ai ouvert mon expertise de recherche à d'autres méthodes issues de l'apprentissage automatique, notamment par réseaux de neurones . On remarquera que ceux-ci manifestent malgré tout une structuration décentralisée et une dynamique complexe.*
- **Apprentissage automatique** : Ensemble d'outils d'apprentissage statistique très utiles pour identifier des régularités de l'environnement ou proposer une représentation de celui-ci. *Ces méthodes me semblent essentielles pour saisir la dynamique complexe d'un agent apprenant en interaction avec son environnement.*

3

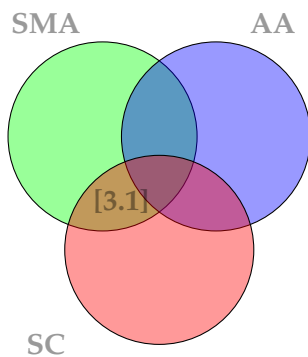
Synthèse des travaux

Ce chapitre introduit les projets principaux sur lesquels j'ai pu travailler, ce qui permet de préciser à partir des descriptions applicatives les enjeux du point de vue de la résolution de problèmes.

3.1 Auto-organisation de topologie pour le partage de données d'un réseau pair-à-pair

Mots-clefs : auto-organisation, contrôle décentralisé, topologie P2P et topologie émergente, SMA, réseaux P2P, partage de documents

Ce projet est issue de la collaboration transversale entre les équipes Multi-Agents et Graphe au sein du laboratoire LIESP. Il a été porté par Mohammed Haddad et moi-même.



Une des motivations de ce projet a été de s'intéresser aux propriétés de contrôle décentralisé que l'on peut déployer par un système multi-agents pour les réseaux pair-à-pair. Le système construit par son activité décentralisée une représentation de son environnement.

Au cours de ce projet, nous proposons une approche auto-organisée qui exploite un algorithme de fourragement inspiré du fourragement observé chez les insectes sociaux [Grassé59], pour un réseau pair-à-pair non structuré dans le cadre du partage de documents. Au cours de cette étude, on s'intéresse à la topologie du réseau pair-à-pair ainsi qu'aux propriétés de la topologie du réseau phéromonal qui se construit au dessus de ce réseau.

L'approche que nous proposons permet d'améliorer la recherche de documents et d'atteindre plus facilement les documents les plus rares tout en maintenant le système robuste face aux perturbations. Les résultats détaillés de ce projet peuvent être consultés par exemple dans [Armetta14].

3.1.1 Réseaux pair-à-pair et partage de documents

Un réseau pair-à-pair est défini comme «un système distribué pour lequel tous les nœuds sont équivalents dans leurs fonctionnalités et tâches» [Androutsellis-Theotokis04]. Ces réseaux sont construits au niveau applicatif, au-dessus de l'infrastructure de communication, telle qu'internet. Il peuvent être décrits par des graphes pour lesquels les nœuds sont les pairs et les arêtes sont les liens entre pairs.

Les pairs sont autonomes et peuvent quitter le réseau de façon asynchrone. Ceci ne doit cependant pas causer d'interruption dans le fonctionnement du système. Dans le cas des réseaux pair-à-pair non structurés étudiés dans ce projet, chaque pair est à la fois un client qui sollicite les ressources du système et également un serveur qui met ses ressources à disposition du système. Chaque pair intègre le système en se connectant à un sous-ensemble de pairs, ce qui forme pour la plupart des applications un graphe de topologie aléatoire. Contrairement au modèle client-serveur, il n'est pas prévu d'indexation distribuée des ressources, qui serait à maintenir continuellement à jour. Ces réseaux non structurés sont ainsi asynchrones, entièrement distribués et très robustes face à la dynamique du système. Cependant, proposer un mécanisme de routage sans centralisation des différentes requêtes émises par les pairs est un défi ambitieux que nous avons souhaité adresser.

Nous proposons ainsi doter le système d'une topologie adaptée au partage de données, sans indexation centralisée de données.

3.1.2 Marquer les chemins les plus efficaces

La méta-heuristique d'optimisation par colonies de fourmis (ACO), décrite dans [Dorigo92], a été proposée comme une généralisation des premiers algorithmes basés sur les fourmis, et a été appliquée avec succès à divers problèmes d'optimisation combinatoire difficiles, y compris les problèmes de routage [Gambardella99].

Plus récemment, [Michlmayr06] a proposé la conception d'une approche basée sur les fourmis pour le routage des requêtes dans les réseaux pair-à-pair.

Dans le cadre des simulations réalisées, on définit le système comme hébergeant différents documents organisés par sujet. Les sujets sont pour certains communs (25 % des pairs), et pour d'autres extrêmement peu représentés (0.1 % des pairs), et donc plus difficiles à retrouver. En effet, dans le contexte de ce projet, il n'existe pas de référencement centralisé des documents, lorsque le sujet est peu représenté, la requête envoyée par le pair doit donc "trouver son chemin" très efficacement vers les pairs hébergeant des documents relatifs au sujet recherché.

Un pair est défini comme intéressé par un sujet d'intérêt principal et quelques sujets secondaires, de sorte que 80% de ses documents partagés sont associés à son sujet principal et 20% sont associés à des sujets secondaires (dans notre cas, fixé à 3). À partir de ces paramètres, chaque requête soumise est dédiée à un sujet va exploiter le mécanisme de fourragement que nous adaptons pour ces réseaux.

Nous voyons sur la figure 3.1 qu'une requête émise par un pair, pour un sujet, sélectionne à chaque étape de son parcours le chemin marqué comme étant le plus prometteur pour les documents du sujet recherché (exploitation) ou envisage un nouveau chemin (exploration). Une fois que la requête atteint sa profondeur maximale, un marquage du chemin parcouru est réalisé suivant la quantité de documents retrouvés pour le sujet en aval.

Ainsi, il se forme au cours de l'activité du système, pour chaque sujet, différentes topologies «phéromonales» dont nous allons étudier les propriétés et pour lesquelles nous allons mesurer les performances.

3.1.3 Une topologie adaptative issue de l'activité du système

Tel que décrit dans [Zhao06], la plupart des réseaux pair-à-pair non structurés actuels présentent des distributions de degrés suivant une loi de puissance. Pour une telle topologie de réseau, la probabilité qu'un nœud ait x liens est proportionnelle à $x^{-\tau}$, où $\tau > 0$ est défini comme l'exposant de la distribution. Plusieurs approches pour effectuer des recherches de requêtes dans les réseaux pair-à-pair exploitent la topologie du réseau pour établir des schémas de routage appropriés. Ainsi, une propriété de petit monde serait intéressante si elle est obtenue dans un système pair-à-pair. Les auteurs de [Guclu08] affirment que la meilleure efficacité de recherche dans les réseaux réalistes pourrait être atteinte lorsque la topologie suit une loi d'échelle (loi de puissance), qui offre des efficacités de recherche de $O(\ln \ln N)$, où N est le nombre de nœuds dans le graphe. Une topologie invariante d'échelle induit une propriété de petit monde. Cependant, la génération et la maintenance de telles topologies invariantes d'échelle sont difficiles à réaliser dans des environnements distribués et potentiellement non coopératifs comme dans les réseaux P2P.

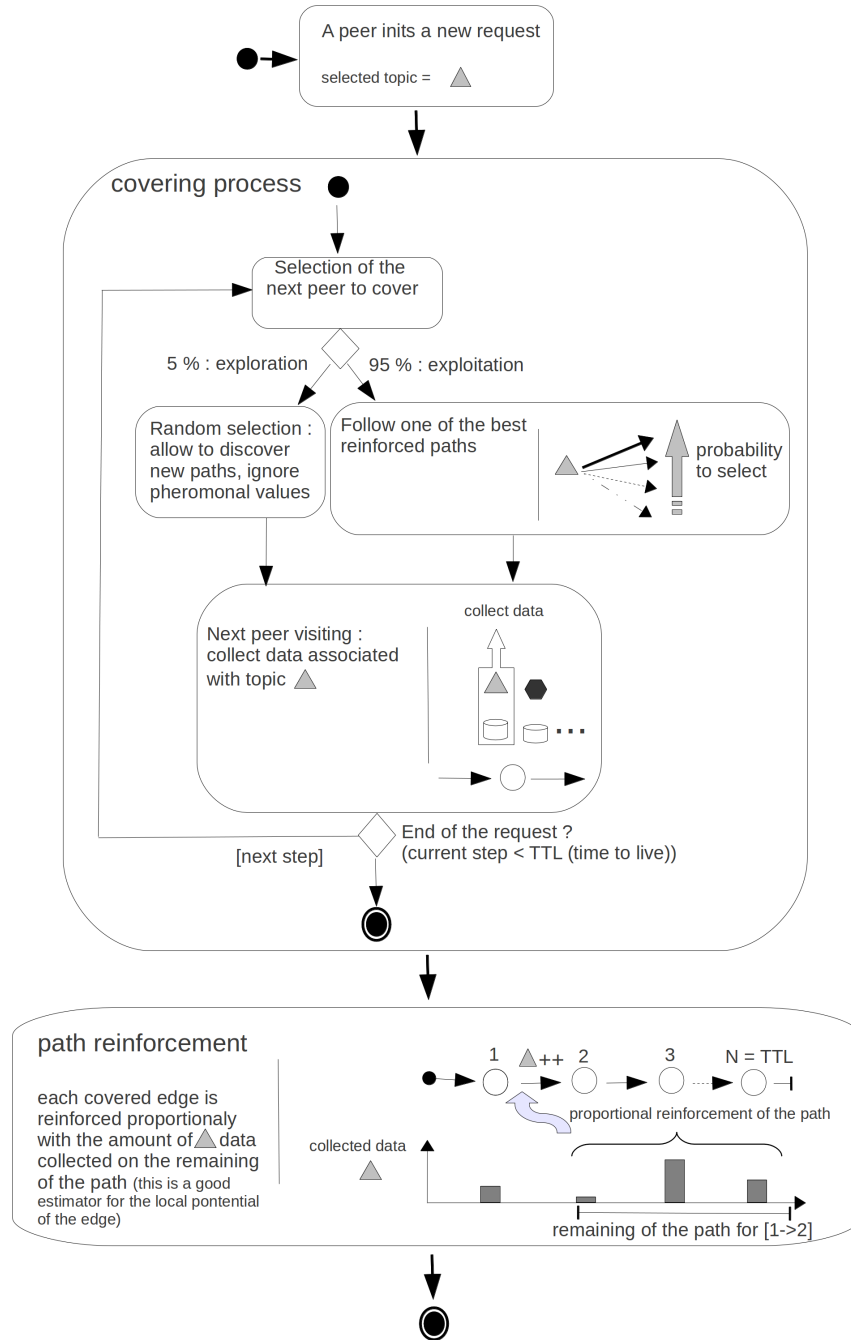


FIGURE 3.1 – Mécanisme de création de topologie phéromonale

Plusieurs études se sont intéressées à l'émergence de réseaux invariants d'échelle à partir de configurations du monde réel comme les marchés [Tseng09], les flux d'information [Wu12] ou les structures motivées par la concurrence [Geng09].

Dans le cadre de ce projet, nous simulons un réseau de 5×10^3 pairs hébergeant des sujets de rareté variable (de 0.1% de présence par pair à plus de 25% de présence).

Tel que présenté sur la figure 3.2, les pairs sont connectés selon plusieurs topologies de réseau : (i) Topologie aléatoire (RND) : la probabilité d'occurrence des arêtes est uniforme, ce qui induit une distribution normale des degrés dans le réseau. Nous faisons varier la probabilité de connexion p afin d'obtenir différentes densités de réseau. (ii) Topologie invariante d'échelle (SF) : une distribution des degrés suivant une loi de puissance, obtenue de manière similaire à [Guclu08]. Plus le degré d'un nœud est élevé, plus sa probabilité d'obtenir de nouvelles connexions est grande. (iii) Topologie sémantique (SEM) : similaire à la précédente, mais la probabilité est calculée selon la similarité sémantique des nœuds plutôt que leur degré structurel (si deux pairs partagent le même premier sujet, la probabilité de connexion p est 100 fois plus élevée). (iv) Topologie sémantique invariante d'échelle (SF + SEM) : obtenue en combinant les deux calculs de probabilité précédents.

Pour ces expériences, nous nous intéressons à la topologie de second niveau résultant du processus d'apprentissage par dépôts de phéromones pour chacun des sujets.

Afin d'extraire le réseau résultant, des filtres sont appliqués localement pour chaque arête comme suit :

- l'arête est activée si sa valeur phéromonale est X fois supérieure à la valeur moyenne des arêtes sortantes (pour le sujet), et est Y fois supérieure à la valeur phéromonale moyenne des arêtes pour l'ensemble du réseau ;
- sinon, l'arête est ignorée.

Pour notre étude, nous extrayons les chemins les plus renforcés, et fixons X et Y respectivement à 10 et 100.

Nous avons pu mesurer qu'indépendamment de la topologie initiale connectant les pairs (aléatoire, invariante d'échelle, sémantique, sémantique invariante d'échelle), la topologie formée par le réseau phéromonal manifeste une distribution de degrés en loi de puissance (de type invariant d'échelle). C'est ce que nous observons sur la figure 3.3, qui repose sur un graphe aléatoire qui suit une distribution normale et manifeste suite au filtrage phéromonal réalisé les caractéristiques d'un réseau à invariance d'échelle (*scale-free*).

Obtenir et maintenir cette topologie de façon décentralisée et asynchrone est un résultat de recherche en soi. Nous mesurons dans la section suivante les bénéfices applicatifs apportés par une telle topologie du point de vue de la résilience et de la performance du système formé.

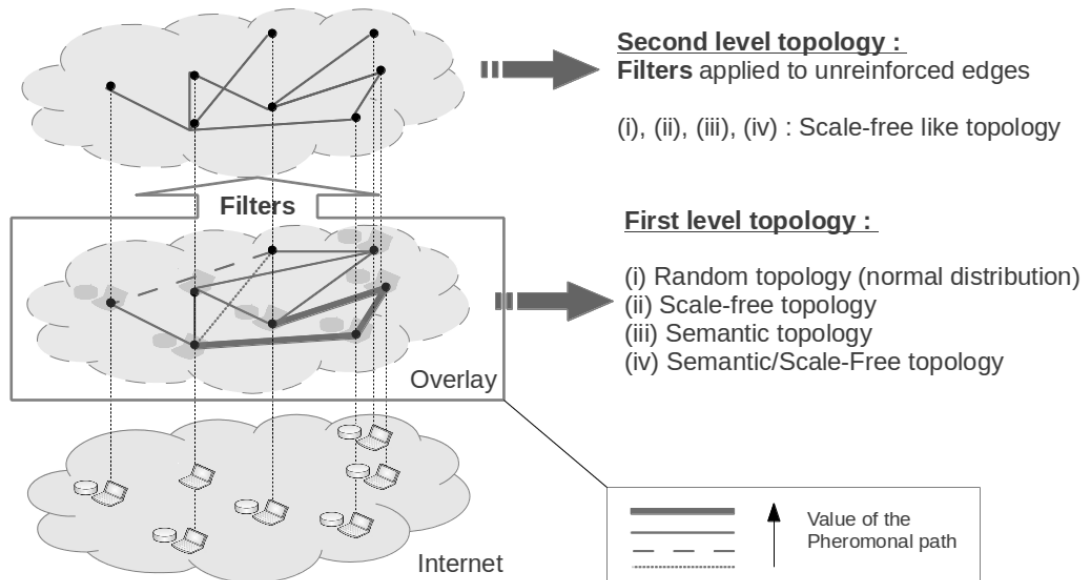


FIGURE 3.2 – Un réseau pair-à-pair en 3 couches

3.1.4 Résilience et performance

Afin d'étudier la robustesse du système, tel que présenté sur la figure 3.4, nous simulons une attaque du réseau et retirons 2/3 des pairs hébergeant principalement des données associées au sujet 6¹ au cours de son fonctionnement.

On constate que les sujets non concernés explicitement par cette perturbation (leurs sujets principaux sont différents) ne sont pas affectés, même si ils partagent exploitent le même réseau pair-à-pair. Les topologies invariantes d'échelle sont bien connues pour être résilientes face à des attaques non ciblées. Cela explique le fait que lorsqu'un sujet est perturbé, les autres sujets sont rarement affectés. Le système retrouve par ailleurs rapidement des performances proches de celle initiale pour le sujet 6, par réparation, sans réinitialisation de l'apprentissage. Les performances sont bien meilleures que celle proposée par une marche aléatoire, ce qui peut s'expliquer par la propriété de petit

1. les sujets 2, 4, 6 concernent respectivement 0,3, 1, 3 % des pairs en sujet principal

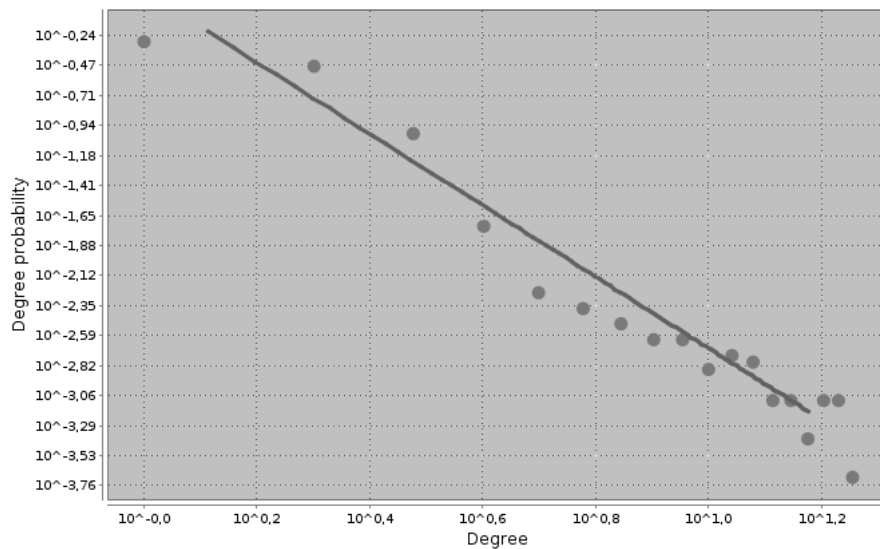


FIGURE 3.3 – Topologie à invariance d'échelle du réseau engendré par les dépôts de phéromones (échelle logarithmique)

monde (*small world*) des réseaux à invariance d'échelle, qui permet de rapprocher les pairs, ce qui est très utile pour les requêtes dont la profondeur est bornée.

Ce projet a également mené à l'étude des topologies de support (connexion initiale des pairs) les plus efficaces, lorsque l'algorithme de recherche phéromonale présenté est utilisé.

Tel que présenté sur le tableau 3.1, plus le nombre d'arêtes dans le graphe est élevé, plus les chemins marqués sont efficaces. Cependant, cela entraîne une surcharge de communication (nombre élevé de messages échangés) ainsi qu'un processus d'apprentissage plus lent. Nous pouvons observer, cependant, que lorsque le nombre d'arêtes est élevé, la performance par arête diminue considérablement, ce qui signifie que les arêtes ne sont pas bien exploitées. Ainsi, la nature de la topologie sous-jacente est très importante pour la performance du processus d'apprentissage. En fait, les topologies SF, SEM et SF+SEM sont plus adaptées et donnent de meilleurs résultats. Notre algorithme peut exploiter des topologies plus légères pour son fonctionnement. De plus, la topologie SEM semble surpasser la topologie SF puisqu'elle cherche à rapprocher les pairs similaires dans le réseau.

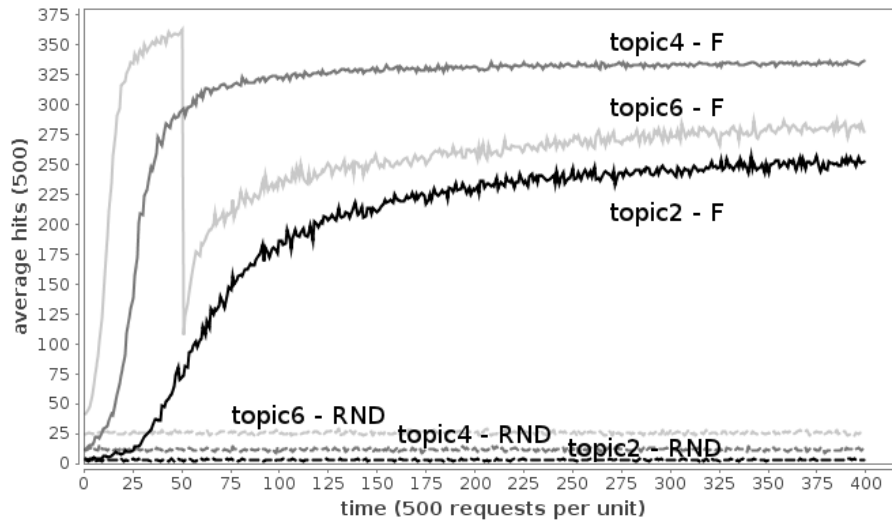


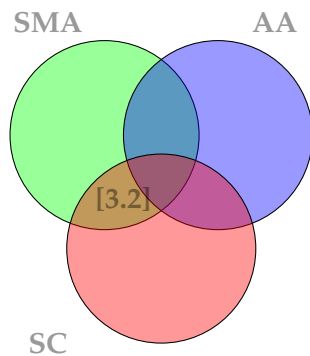
FIGURE 3.4 – Résilience du processus de recherche de documents (comparatif de recherche par parcours aléatoire (RND) en bas de diagramme)

Topology	Edges	Perf./Edg.	P/E gain.	Avg. perf.
RND p=0.01	252801	0.0020	15.62%	508.76
RND p=0.005	128883	0.0031	24.22%	398.90
RND p=0.002	54943	0.0068	53.12%	375.17
RND p=0.0008	25418	0.0128	100%	326.58
SF	24985	0.0138	107.81%	345.362
SEM	24985	0.0185	144.53%	463.16
SF+SEM	24985	0.0187	146.09%	466.51

TABLE 3.1 – Influence de la topologie initiale du réseau

3.2 Découverte d'environnement décentralisé et mécanismes de contrôle de fiabilité

Mots-clefs : construction décentralisée de représentations, fiabilité des informations échangées, cartographie collaborative, véhicules connectés



Ce travail s'intéresse au maintien de la cohérence et de la robustesse dans un système multi-agents décentralisé. Que ce soit dans le cadre de la cartographie d'environnements par une flottille de robots autonomes ou dans le cadre de la régulation de trafic par l'introduction de véhicules autonomes, un des problèmes fondamentaux concerne le partage de données et leur fiabilisation pour le système à réguler.

Dans le cadre de cette étude, on s'intéresse notamment à la résilience du système lorsque des informations inexactes sont propagées. L'approche utilisée consiste à estimer à travers le collectif et différentes agrégations non-linéaires une valeur de confiance pour chaque individu qui forme le système. Le réseau de confiance ainsi constitué permettant de calculer la fiabilité des informations parcourant le système.

Chaque agent construit un réseau de confiance en exploitant l'écart entre ses perceptions directes et ses perceptions indirectes (informations propagées), tel que présenté sur la figure 3.5. À la suite de ces échanges de données, chaque agent peut ainsi calculer une valeur de confiance pour les différents agents du système. Une description de ce procédé est détaillée dans [Schillo99].

Une des contributions de ces travaux a consisté à coupler le réseau de confiance ainsi généré avec le réseau de communication afin de contenir la propagation des informations erronées et d'améliorer l'activité du système, tel que présenté dans [Nguyen Vu12a].

3.2.1 Application à la cartographie collaborative

Ce travail correspond au travail de thèse de Q. A. NGuyen Vu, dans le cadre d'un financement AUF (Agence Universitaire de la Francophonie) en collaboration avec l'IFI (Institut Francophone International).

L'objectif de la découverte d'environnement consiste à coordonner l'ensemble des robots disponibles afin d'établir le plus rapidement et efficacement possible une carte de l'environnement, tel que proposé sur la figure 3.6 pour un problème de détection de zone dangereuse (radioactivité, incendie, etc.). Dans le cadre de cette application, différentes informations sont échangées sur les zones déjà explorées afin d'établir, de façon décentralisée, des stratégies d'exploration individuelles.

Ce travail a pu montrer que la stratégie de partage et de diffusion d'information manifeste comparativement de bonnes propriétés de découverte et de résilience [Nguyen Vu12b]. Le système mis en place permet à chaque agent de s'auto-diagnostiquer. Ce travail s'est également intéressé au couplage entre l'organisation

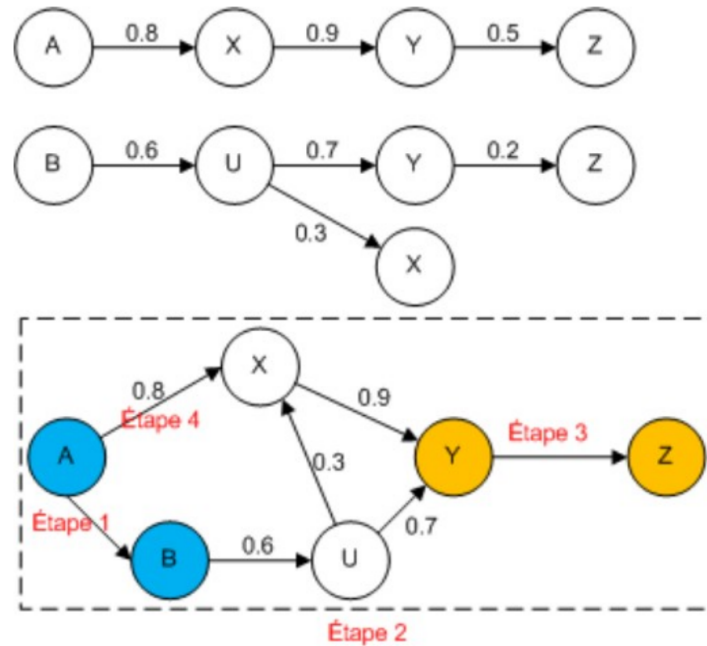


FIGURE 3.5 – Dynamique des réseaux de confiance (*Trust Net*)

Partie supérieure : les agents A et B bénéficient d'indicateurs de confiance directe avec les agents X et U (avec qui ils ont pu partager des données), et indirecte avec d'autres agents (avec qui ils n'ont pas échangé directement)

Partie inférieure : l'agent A reçoit la représentation de l'agent B : (1) il calcule sa confiance envers l'agent B, (2) il se connecte au réseau de confiance de l'agent B (3) il traite les chemins d'informations partagées, (4) il met à jour ses valeurs de confiance directes.

sociale résultant du processus décentralisé d'évaluation de fiabilité et l'organisation spatiale résultant des actions réalisées par les agents.

Cette étude a permis également de soulever quelques difficultés propres à tout système distribué tel que le problème de latences dans le *feedback* perçu à travers le collectif.

3.2.2 Application à la régulation d'un ensemble de véhicules connectés

Les récents progrès dans le domaine des véhicules autonomes et connectés engendrent de nouveaux besoins dans le cadre du partage d'informations entre véhicules pour une régulation collective, ce qui incite à fiabiliser les informations partagées et ainsi à appliquer les mécanismes de confiance présentés.

Tels que détaillés dans la section 3.5.2, les véhicules autonomes qui régulent leur

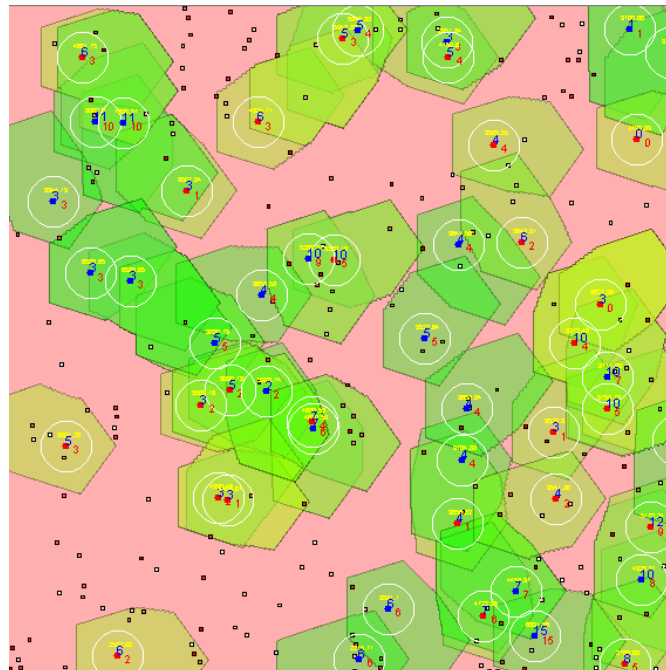


FIGURE 3.6 – Vue du simulateur multi-agents *Danger Mapping*, développé sur la plateforme de simulation GAMA [Taillandier10], dans sa version 1.3 : des agents communiquent, perçoivent et explorent leur environnement.

vitesse appliquent des lois de poursuite exploitant les vitesses et distances inter-véhiculaires.

Une partie du travail s'est ainsi intéressé à enrichir les mécanismes de poursuite en exploitant des informations partagées issues des différents capteurs véhicules connectés. Ces données sont ainsi exploitées dans une loi multi-anticipative. Chaque véhicule utilise des vitesses de véhicules qui sortent de son horizon de perception direct pour réguler sa vitesse.

Le modèle multi-agents proposé se compose de trois couches connectées telles que présentées sur la figure 3.7. D'une part, la couche physique représente les règles qui permettent de modéliser le comportement physique des véhicules. La couche de communication gère les échanges d'informations par des règles de proximité et de fiabilité. La couche confiance qui met en œuvre différents indicateurs de fiabilité des différentes informations partagées.

Les résultats de simulation issus de ce modèle, présentés dans [Guériau15b, Guériau15a], ont montré une amélioration de la fluidité du trafic et une amélioration des conditions de sécurité.

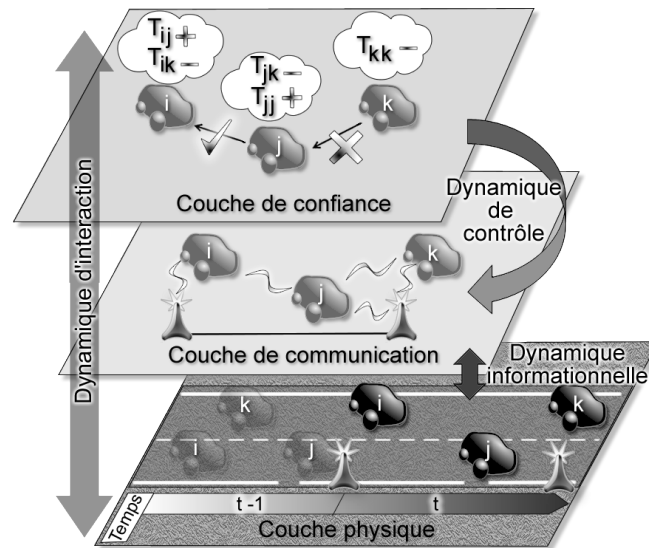
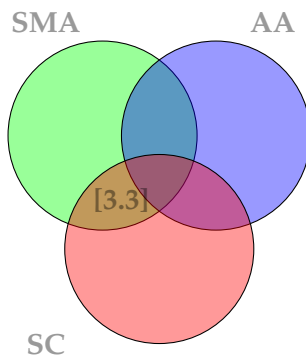


FIGURE 3.7 – Architecture en 3 couches

3.3 Apprentissage de régularités en environnement continu et Intelligence Ambiante

Mots-clefs : apprentissage sensorimoteur, construction décentralisée de représentations, discrétisation de signaux continus, problème d'amorçage, motivation intrinsèque, intelligence ambiante

Ce projet est issu de la collaboration avec l'entreprise Ubiant (<http://www.ubiانت.com/>), qui a été très impliquée dans le développement des technologies pour l'intelligence ambiante en France. Le travail présenté correspond aux résultats de la thèse de Sébastien Mazac [Mazac15a], que j'ai co-encadré avec Salima Hassas.



Ce travail aborde quelques problèmes fondamentaux liés à l'application des méthodes d'apprentissage constructivistes dans des environnements continus. Il propose un mécanisme d'amorce pour la perception de signaux continus, grâce à un système multi-agents. Il s'intéresse à construire sans connaissances préalables des motifs de perception d'évènements significatifs par leurs capacités à saisir les régularités observées dans l'environnement.

Comme pour la robotique autonome, l'intelligence ambiante (AmI) est un sujet complexe qui implique des agents autonomes incarnés qui n'ont pas nécessairement

d'objectifs prédéfinis tout en étant dotés de capacités sensori-motrices évolutives que le système doit s'approprier.

3.3.1 Intelligence Ambiante

Le concept d'intelligence ambiante commence à apparaître dans les années 2000. Il se distingue d'autres notions proches, telles que la *domotique*, l'informatique ubiquitaire, les environnements pervasifs, l'Internet des objets en plaçant la notion d'intelligence au centre de la problématique. En effet, on s'intéresse pour le domaine de l'intelligence ambiante à coordonner efficacement un ensemble de capteurs (température, humidité, luminosité, présence, etc.) et un ensemble d'effecteurs (radiateurs, lampes, machines à laver, aspirateur, etc.) au sein d'un environnement domotique tel que celui présenté sur la figure 3.8.

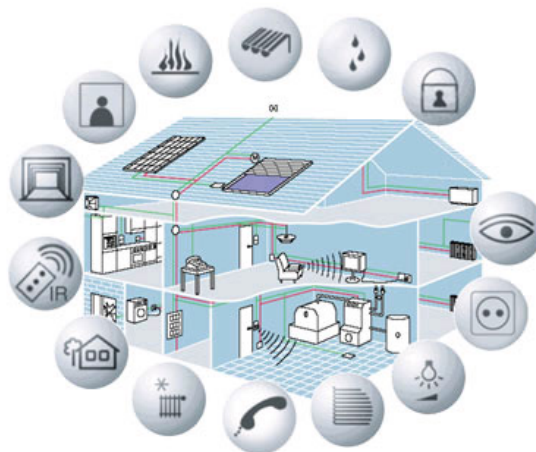


FIGURE 3.8 – Intelligence Ambiante : des capteurs et des effecteurs

Un habitat ainsi équipé pourrait permettre une meilleure gestion énergétique du bâtiment, en assistant l'utilisateur dans sa politique de consommation individuelle, selon les ressources disponibles qui évoluent au cours de la journée. L'objectif étant alors d'accorder les indicateurs de consommation de millions d'individus avec ceux d'un parc de production énergétique très hétéroclite et fluctuant. Ceci répond à des préoccupations écologiques et économiques, les pics de production énergétique nécessitant des «coûts d'absorption» prohibitifs (achat en situation de crise, centrales à gaz ou à charbon, etc.).

L'intelligence ambiante s'intéresse également à la mise en place de nouveaux usages. Les maisons intelligentes (*smart homes*), bureaux, usines, gares, etc. , pourront proposer de nouveaux services. Il est possible que cette évolution soit majeure

(de l'ordre d'une révolution industrielle), même s'il est encore difficile d'imaginer l'ensemble des nouveaux services qui pourront découler de ces avancées.

Qu'il s'agisse de maximiser le confort énergétique de l'utilisateur ou de proposer des services d'assistance inédits, un des verrous scientifiques limitants correspond à la capacité d'autonomie du système dans les propositions faites à l'utilisateur. Il semble en effet naturel qu'un tel système, qui peut s'apparenter à un robot avec des capacités de perception et d'action, soit doté d'une forme d'intelligence.

3.3.2 Des évènements et des signaux continus

Dans une démarche constructiviste ou développementale, des travaux visant la qualification des évènements ont déjà été réalisés à partir de capteurs émettant des valeurs discrètes, tels que proposés dans [Drescher91a] (*schema learning*). Dans le domaine de l'intelligence ambiante, les capteurs ne fournissent cependant pas de signaux pré-discrétisés. Le mécanisme de discrétisation fait en réalité indissociablement partie du processus de cognition.

Les problèmes de discrétisation du signal d'entrée et de recherche de régularité sont en réalité des problèmes interdépendants (de type *oeuf-poule*) qui ne peuvent être abordés indépendamment. On ne peut en effet pas identifier de régularité sans évènements, mais on ne peut pas non plus définir d'évènement stable sans que le motif auquel il participe soit formé. Ceci pose ainsi un problème d'amorce (*bootstrapping problem*), tel que présenté par exemple dans [Kuipers06]. Le système ainsi décrit manifeste des enjeux d'exploration combinatoire parmi l'ensemble des possibilités de structures. Ceci est d'autant plus marqué que, selon une perspective constructiviste ou développementale, les motifs de bas niveau construits sont requalifiés en variables de perception pour le système, afin de construire des motifs cognitifs de plus en plus élaborés à d'autres niveaux.

Dans le cadre de ce travail, nous nous intéressons à doter le système d'une représentation de son environnement. Cette représentation a pour objectif d'améliorer les capacités prédictives du système, en comprenant ses régularités.

L'expérience du système est constituée d'un ensemble de valeurs continues captées, telles que présentées sur la figure 3.9.

L'objectif est alors de construire des motifs de régularité dans le flux de perception. Selon le modèle proposé, trois types d'activité participent au processus de construction des motifs (*patterns*) :

- un processus de découpage (*Cut*) qui permet d'identifier des variations de valeurs ;
- un processus d'association (*Associate*) qui permet d'associer des variations au cours du temps ;

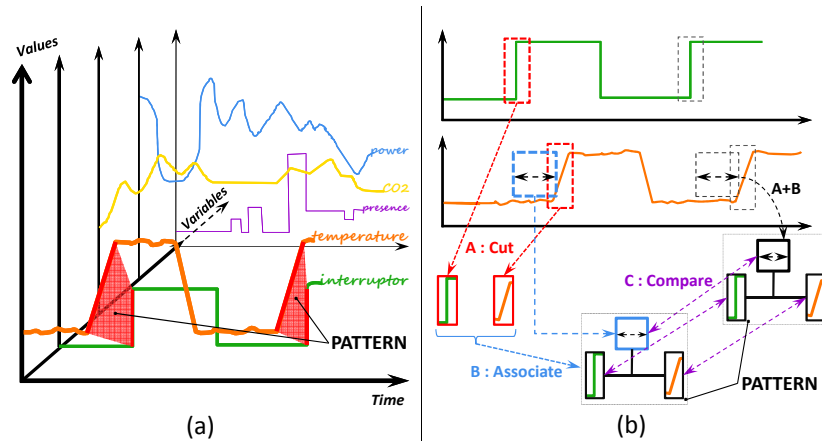


FIGURE 3.9 – (a) Expérience continue du système à travers un ensemble de variables continues (b) 3 types d’opérations couplées pour la création de motifs A : Cibler les variations B : Trouver les associations entre variations C : Trouver les régularités afin de constituer des motifs et maintenir les structures (*feedback*)

- un processus de comparaison (*Compare*) qui identifie les régularités et matérialise des motifs.

3.3.3 Une dynamique de contrôle décentralisée auto-catalytique

Nous voyons sur la figure 3.10 que le système peut être instancié par un système multi-agents, et matérialise ainsi une dynamique de contrôle décentralisé².

Le système ainsi formé est constitué de trois types d’agents (perception par découpage du signal, association entre évènements perçus, similarité entre évènements). Chaque type d’agent peut réaliser sa fonction de génération d’évènements suivant différentes méthodes paramétrées. Les méthodes et paramétrages sont alors validés par la dynamique du système lorsqu’ils sont "utiles", c’est-à-dire lorsqu’ils participent à la construction de motifs récurrents, et sont amenés à varier dans le cas contraire.

Parcourons la figure 3.10 afin de préciser la dynamique ainsi formée. Les agents de perception et les agents d’association produisent des évènements de représentation correspondant respectivement à une description d’une variation et à une description d’une relation entre les évènements (représentées par le centre du cercle rouge, et l’intersection des cercles rouge et bleu). Les agents de similarité comparent les évènements issus d’agents perception ou d’agents associations, ils proposent également des mécanismes de similarité dont la pertinence sera évaluée de façon décentralisée à la suite du

2. Le contrôle décentralisé est à différencier d’un contrôle distribué, car il ne nécessite pas de coordonnateurs centraux

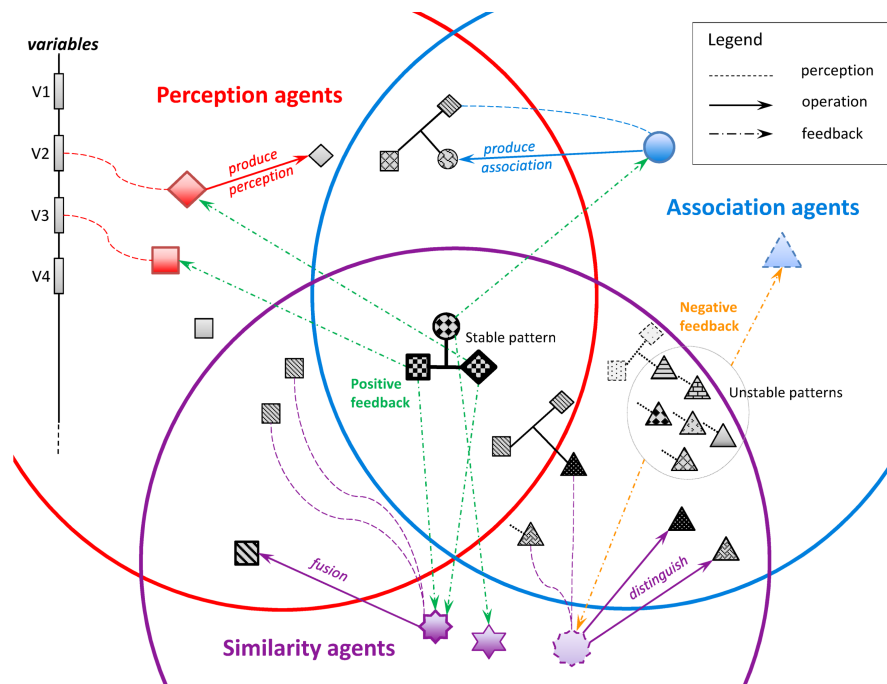


FIGURE 3.10 – Un contrôle décentralisé à base de multi-agents pour l'apprentissage de régularités

processus génératif. Ils peuvent considérer des événements comme similaires (intersection cercles rouge et violet concernant la perception, intersection cercles bleu et violet concernant les associations), et les rassembler ou fusionner. On remarque, tel que représenté (voir *unstable patterns* sur la figure), que si les associations ne sont pas assez régulières et ne participent pas à un motif durable, elles retournent aux agents ayant participé à leur formation un *feedback* négatif. Au contraire, lorsqu'un motif est stable dans le temps (au centre des trois cercles), c'est-à-dire qu'il capture une régularité de l'environnement, alors un *feedback* positif de renforcement est transmis aux agents lui ayant permis d'émerger.

Le motif stable ainsi formé peut être intégré comme variable discrète perçue par le système, qui peut participer à des niveaux de perception de plus en plus élaborés, suivant une dynamique similaire. Ce motif peut alors lui-même recevoir et propager un *feedback* suivant l'utilité révélée à des niveaux supérieurs de perception.

3.3.4 Simulateur

Dans le cadre du simulateur réalisé, nous avons pu observer la formation des motifs de régularité, et visualiser sous la forme d'un graphe dynamique l'ensemble des entités

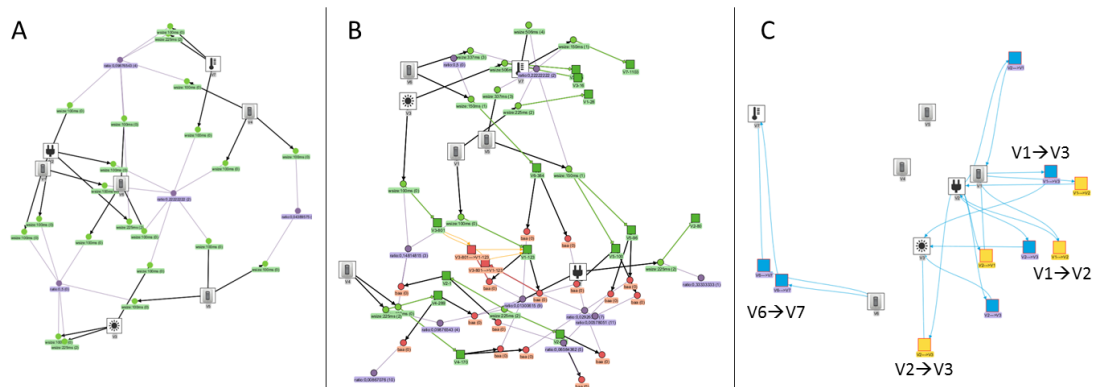


FIGURE 3.11 – A : Etat initial. B : Création d’associations. C : Motifs intéressants.

impliquées.

le système dispose de certains capteurs de signaux continus représentés par des icônes sur la figure 3.11. Sur la partie A de la figure, les capteurs, sont en relation avec les agents de perception (rond vert), eux même en relation avec les agents association (ronds violets). La partie B représente les itérations des agents similarité (ronds rouges) et la génération de motifs (carrés).

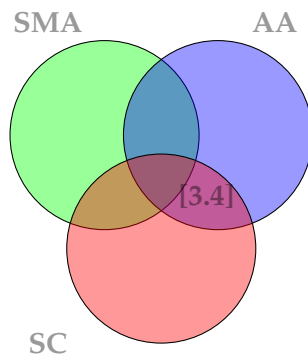
Les motifs sont alors exploités par le système (carrés bleu ou jaune suivant leur niveau de maturité). Sur la figure, le système a appris que certains interrupteurs font varier la consommation électrique et modifient le signal lumineux perçu.

Ces travaux sont détaillés dans cette thèse [Mazac15a] et ces articles [Mazac14], [Mazac15b].

3.4 Extensions de MCTS pour les jeux combinatoires

Mots-clefs : Jeux combinatoires, Jeu de Go, Jeu de Clobber, MCTS, accumulation de connaissances, réification de stratégies, *feedback* par simulation

À l’initiative de ce travail intervient une discussion que nous avons eu avec Éric Duchêne (de l’équipe GOAL du LIRIS) concernant l’intérêt commun que nous avons pour les jeux combinatoires. Le jeu de Go semblait très approprié pour l’élaboration d’une IA constructiviste. D’une part, un bon joueur doit manifester des capacités cognitives de haut niveau, d’autre part, nous pouvons exploiter un *feedback* calculé, tel que proposé par l’algorithme MCTS (Monte Carlo Tree Search), qui était alors avant l’avènement des méthodes à base de réseaux de neurones le plus convainquant. Ce travail a été réalisé par André Fabbri au cours de sa thèse.



Ce travail propose d'enrichir le mécanisme de recherche opérationnelle MCTS pour les jeux combinatoires en mutualisant les configurations perçues pour l'estimation des positions. Ce travail s'intéresse également à la réification de stratégies d'exploration pour des objectifs intermédiaires au sein de l'arbre d'exploration MCTS.

3.4.1 Jeux combinatoires et jeu de Go

Les jeux combinatoires ont été naturellement des applications privilégiées pour l'intelligence artificielle, car ils engagent pour l'humain des processus de cognition de haut niveau. Ce fut notamment le cas lorsque DeepBlue a battu le meilleur joueur mondial d'échecs en 1997. Le système présenté alors n'engageait pourtant pas les mêmes mécanismes que ceux de la cognition humaine. DeepBlue mesurait 1.80 mètre et pesait 1.4 tonne, il basait sa stratégie essentiellement sur ses capacités alors inédites de calcul, essentiellement en projetant sur un horizon élevé les différentes combinaisons de jeu accessibles par anticipation afin de sélectionner l'évolution de jeu la plus favorable. Cette intelligence combinatoire n'a pas été suffisante pour adresser des problèmes qui s'expriment dans des espaces encore plus grands, pour lesquels l'exploration presque systématique était hors de portée. C'est le cas du jeu de Go pour lequel on peut estimer le nombre de configurations de jeu plausibles de l'ordre de 10^{171} , alors qu'elles ne seraient «seulement» que de l'ordre de 10^{47} pour les échecs d'après [Allis94].

Une difficulté que l'on retrouve pour le jeu de Go est que celui-ci sollicite plus la stratégie que la tactique, et s'intéresse plus aux positionnements de dominance qui peuvent s'établir, qu'à la réalisation tactique qui est une simple conséquence de la configuration stratégique. Du point de vue d'un apprentissage, il est plus difficile d'établir des connaissances générales au niveau stratégique, alors que la réalisation tactique est plus accessible d'un point de vue calculatoire. Les aspects stratégiques et tactiques ne peuvent cependant pas être envisagés de façon séparée. Le succès du joueur résulte également dans sa capacité à alterner réalisations tactiques et stratégiques, tout en étant attentif pour le jeu de Go à d'autres préoccupations telles que le maintien de l'initiative ou les menaces Ko, etc.

Dans la communauté IAJ (Intelligence Artificielle Appliquée aux Jeux), les différents algorithmes ont longtemps exploité principalement des méthodes de parcours des positions par anticipation (par exemple par arbre de recherche), afin de sélectionner l'orientation de jeu la plus favorable. Étant donné la quantité très importante de positions de jeux, ces méthodes ne peuvent dérouler l'ensemble des configurations possibles pour atteindre la fin de partie. L'objectif est alors de dérouler l'arbre autant que

possible, et d'utiliser différentes heuristiques permettant d'estimer la branche la plus favorable pour le joueur. Par exemple, aux échecs, une branche prédisant la perte d'une pièce maîtresse adverse sera privilégiée par le système. Ce type d'estimation est cependant plus délicate à réaliser au jeu de Go, pour lequel les pierres "vivantes" ou "mortes" qui constitue les groupes sont uniquement connues en fin de partie. À cause de l'effet d'horizon occasionné par la limitation de la profondeur de la projection, l'estimation intermédiaire d'une position est plus difficile, et ne peut se définir par la perte d'une simple pièce ou l'affaiblissement d'une structure de pièces elle aussi difficile à estimer.

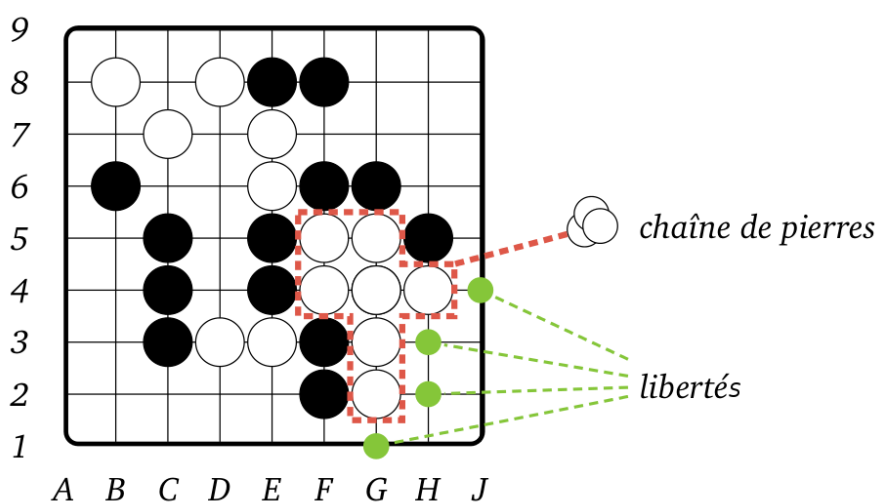


FIGURE 3.12 – Évaluation des positions pour le jeu de Go

Une première approche a consisté à doter les systèmes informatiques de connaissances expertes (ensemble de règles, motifs locaux, etc.), visant à élaguer l'arbre des positions à envisager et estimer la qualité des positions projetées par anticipation. Il est cependant très difficile de caractériser de façon experte, une configuration de jeu formée d'un ensemble de pierres blanches et noires. L'humain peut en effet évoquer les principes de haut niveau qui guident son raisonnement (protéger ses groupes faibles, attaquer les groupes faibles de l'adversaire, conquérir les zones libres, etc.), mais aura beaucoup de peine à les retranscrire pour un support informatique.

Le jeu de Go engage en effet un processus de perception très complexe, qui peut être assimilé à de la reconnaissance de formes, tout en ayant une grande sensibilité aux petites variations : en effet la position d'une pierre peut changer entièrement la qualité d'une position. Bien caractériser le contexte d'application ou percevoir le jeu est nécessaire à tout processus de décision à base de règles. Le problème n'est alors pas la règle elle-même, mais la description fidèle du périmètre d'application de la règle.

Sur la figure 3.12, un joueur débutant percevra un ensemble de pierres. Un joueur

plus expérimenté percevra des groupes connectés, virtuellement connectés, vivants, morts, forts, faibles, des libertés, et même pour les plus avancés un décompte de points. Tous ces concepts abstraits étant le résultat d'un apprentissage cognitif complexe, il est difficile de les expliciter à travers différentes règles. En supposant que nous trouvions une représentation appropriée à la connaissance à décrire, il faudrait juxtaposer un très grand nombre de celles-ci pour imiter la connaissance de joueurs humains expérimentés.

Le travail de thèse présenté s'appuie sur le meilleur algorithme alors connu pour aborder ce problème cognitif et combinatoire qui est présenté dans la section suivante. À la suite de cette thèse, l'introduction de nouveaux algorithmes exploitant des réseaux de neurones pour évaluer les positions de jeu tel que AlphaGo a marqué une évolution majeure du domaine avec des performances inédites.

3.4.2 MCTS (Monte Carlo Tree Search)

Nous avons vu qu'il est particulièrement difficile d'évaluer la qualité d'une position de Go, au sein d'un arbre de parcours. [Brügmann93] propose une évaluation de Monte-Carlo pour le jeu de Go, qui consiste à évaluer la qualité d'une position à partir d'un échantillon de parties jouées aléatoirement jusqu'à leur fin. Utilisée seule, cette méthode ne permettait pas cependant pas de faire face à la complexité combinatoire du jeu de Go. En combinant ce mode d'évaluation à un algorithme de parcours d'arbre, nous permettons une estimation Monte-Carlo plus localisée et donc une meilleure fiabilité, telle que proposée dans [Bouzy04].

MCTS s'intéresse ainsi d'une part à simuler les actions du joueur courant ainsi que les actions de l'adversaire, selon le principe de l'algorithme MinMax. Pour cet algorithme, on réalise un nombre variable de simulations prévoyant 4 phases, telles que présentées sur la figure 3.13.

La phase de descente (*descent*), permet de descendre dans l'arbre en sélectionnant le nœud le plus probant en pondérant le gain estimé et sa fiabilité (nombre d'estimations réalisées). Cette phase permet d'atteindre une feuille de l'arbre, à partir de laquelle un nouveau nœud est créé (*growth*). Ce nœud est ensuite évalué grâce à une méthode d'inspiration Monte-Carlo (*roll-out*) qui permet de terminer une partie. Le résultat de cette simulation permettant de mettre à jour une estimation de la valeur des nœuds de l'arbre (*update*) afin qu'ils puissent être de nouveau considérés pour la phase de descente de la simulation suivante.

Chacune de ces phases est guidée par une politique. Les politiques permettront de définir les stratégies d'exploration et d'exploitation [Auer95], elles permettront également d'inciter certains coups à partir d'heuristiques ou de mettre en place des formes avancées pour le renforcement (*update*).

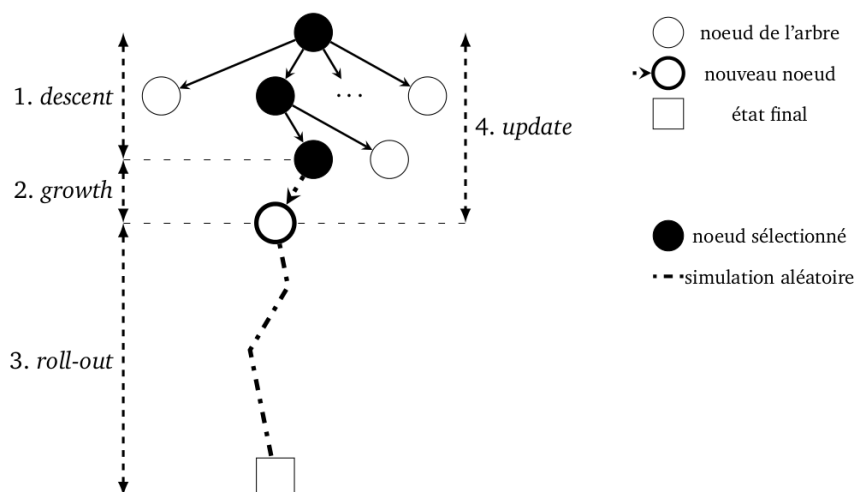


FIGURE 3.13 – Algorithme MCTS

3.4.3 Un processus d'acquisition de connaissances autonome pour MCTS

La phase de *roll-out* a pour objectif d'estimer par une simulation l'état d'une partie suivant la valeur finale obtenue (victoire - défaite). L'accumulation des différentes estimations (plusieurs dizaines de milliers) permet de préciser les chances de gain sur les branches de l'arbre. Il semble naturel de générer des simulations réalistes plutôt qu'aléatoires afin d'améliorer la qualité du *feedback* issue de l'évaluation *roll-out*.

Ainsi, nous avons proposé un processus d'acquisition de connaissance BHRF (*Background History Reply Forest*) [Fabbri12, Fabbri14], permettant d'exploiter les connaissances de l'arbre MCTS pour améliorer la qualité des simulations.

Comme présenté sur la figure 3.14, ces connaissances sont organisées sous la forme d'une forêt construite en même temps que le développement de l'arbre (*growth*). Celle-ci peut être contextualisée et exploitée sur chacune des simulations. Le processus de mise en correspondance retourne ainsi différents arbres dont la racine correspond au coup suivant candidat.

La simulation peut exploiter ces connaissances contextualisées pour améliorer sa pertinence, c'est-à-dire dérouler des coups plausibles. Tel que nous l'avons formulé, la politique de sélection du coup simulé suivant exploite d'une part la valeur des nœuds issue de l'algorithme original MCTS, tout en considérant la correspondance avec les configurations des arbres candidats. Plus la configuration est similaire, plus l'estimation considérée fiable. Ce processus d'association peut être considéré comme une forme simple de généralisation.

On présente sur la table 3.2 des résultats comparatifs entre des versions de MCTS

jeu 9x9, l'exploitation de BHRF permet de meilleurs résultats dès que le nombre de simulations atteint la moitié des simulations réalisées pour l'approche comparative. On compare ainsi 5000 simulations avec BHRF contre 10000 simulations (et croissances de l'arbre) pour l'approche comparative. Ceci souligne bien l'utilité de la connaissance BHRF dans le processus de perception du système formé. Cette efficacité ne compense cependant pas la complexité algorithmique associée à la construction et au maintien de la structure BHRF qui nécessite des optimisations spécifiques (une première optimisation volatile est présentée dans [Fabbri15]).

roll-out simulations	10000	30000	100000
goban 9x9	67.3 ± 1.6	66.9 ± 2	68 ± 2.6
goban 19x19	74.3 ± 3.5	76.6 ± 2.5	77.7 ± 3.4

TABLE 3.2 – Taux de victoire : MCTS + BHRF - MCTS + simulations aléatoires

Nous avons souhaité comparer notre algorithme en dotant cette fois MCTS des meilleures connaissances expertes pour la partie simulation, tel que proposé par le moteur de jeu *Fuego* (version 1.1) [Enzenberger10]. Ces connaissances reposent sur des motifs locaux permettant d'orienter le coup suivant localement.

Nous voyons sur la figure 3.15 le gain apporté par BHRF suivant le taux d'exploitation ϵ de BHRF. Notre proposition permet une très nette amélioration du taux de victoires, pour les problèmes les tailles de plateau les plus combinatoires.

Des résultats complémentaires concernant la nature de la connaissance accumulée par BHRF sont présentés dans [Fabbri16].

3.4.4 Réification d'objectifs intermédiaires pour MCTS

Parmi les propositions faites, nous avons souhaité doter MCTS de capacités à percevoir des objectifs intermédiaires en complément de l'objectif global de gains de parties, les objectifs intermédiaires ayant pour but de supporter la réalisation de l'objectif global. Dans une certaine mesure, la dynamique est la même que celle présentée dans la section précédente : un motif local (sous-objectif) amorce la constitution d'un motif plus global (gain de la partie).

Cette décomposition fait partie des stratégies reconnues pour le jeu de go telles que tuer, encercler, affaiblir des groupes pour gagner. De façon plus générale correspond à la façon d'apprendre incrémentalement et de résoudre les problèmes. Du point de vue perceptif pour le système, nous renforçons ainsi l'exploration visant à un objectif intermédiaire. La récompense ou *feedback* pouvant alors correspondre à l'atteinte d'un objectif intermédiaire.

La figure 3.16 représente une nouvelle structure pour l'arbre MCTS. Le couplage entre les objectifs intermédiaires et l'objectif global nécessite différents ajustements

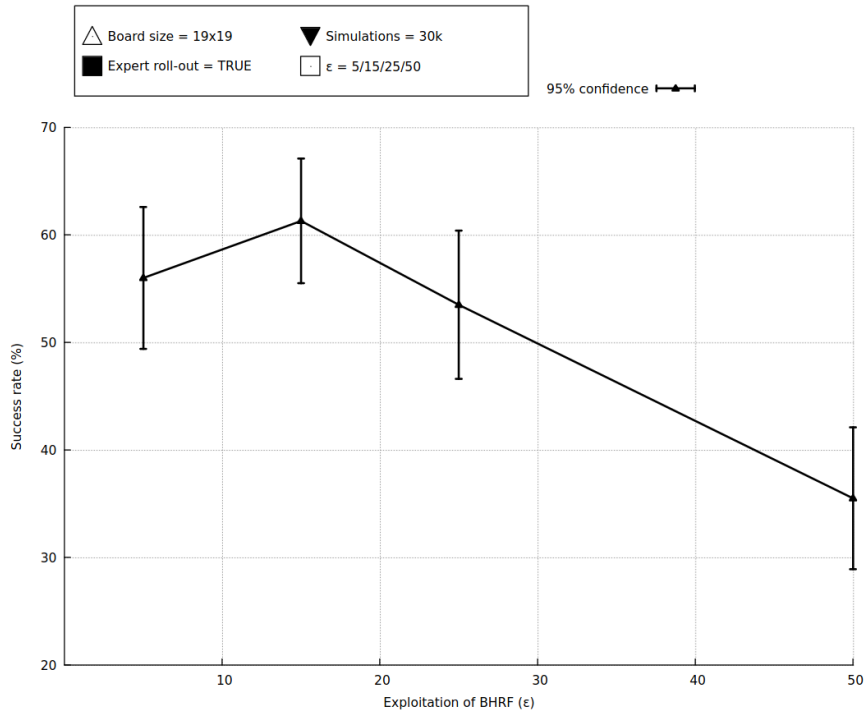


FIGURE 3.15 – Taux de victoire : MCTS + BHRF(ϵ) - MCTS + simulations avec heuristiques

pour les différentes parties de l'algorithme MCTS (phase *descent*, *feedback* local et global, etc.) détaillés dans [Fabbri15].

Une première instanciation a permis de montrer des gains significatifs pour le jeu du Clobber Solitaire présenté sur la figure 3.17. On remarque que l'exploitation d'objectifs intermédiaire est particulièrement pertinente lorsque le nombre de simulations est limité, ce qui souligne une exploitation plus efficace de l'effort de recherche. En contrepartie, lorsque la capacité exploratoire est très importante (beaucoup de simulations), la réification d'objectifs intermédiaire n'a plus d'utilité.

Nous n'avons pas pu obtenir de gains significatifs pour le jeu de Go, mais avons pu mesurer des résultats encourageants montrant des explorations en profondeur consécutives visant les différents sous objectifs, ce qui permet de réduire l'effet d'horizon³ pour les objectifs intermédiaires. Cette première instanciation pourra être améliorée, par une meilleure définition experte des objectifs intermédiaires (vie des groupes de pierres, connexions entre groupes, zones d'influence, etc.) que nous avons dû définir de façon limitée et simplifiée, mais également dans la dynamique par des possibilités à

3. Pour la structure de recherche sous forme d'arbre, l'effet d'horizon correspond à l'aspect limitant de la recherche en profondeur qui est bornée

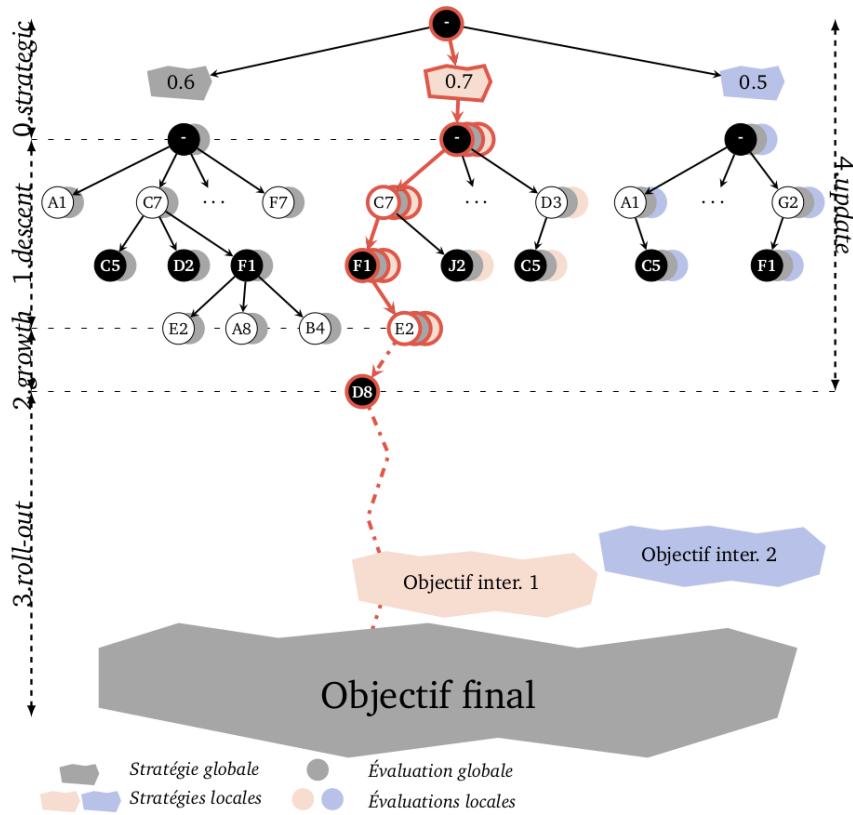


FIGURE 3.16 – Réification d’objectifs intermédiaires pour MCTS

enchaîner en profondeur différents objectifs intermédiaires.

De façon plus ambitieuse, un système cognitif autonome observerait des régularités dans ces différents gains et pertes, il identifierait incrémentalement des motifs récurrents, qu’il pourrait réifier incrémentalement en tant qu’objectifs intermédiaires de lui-même pour la suite de son activité.

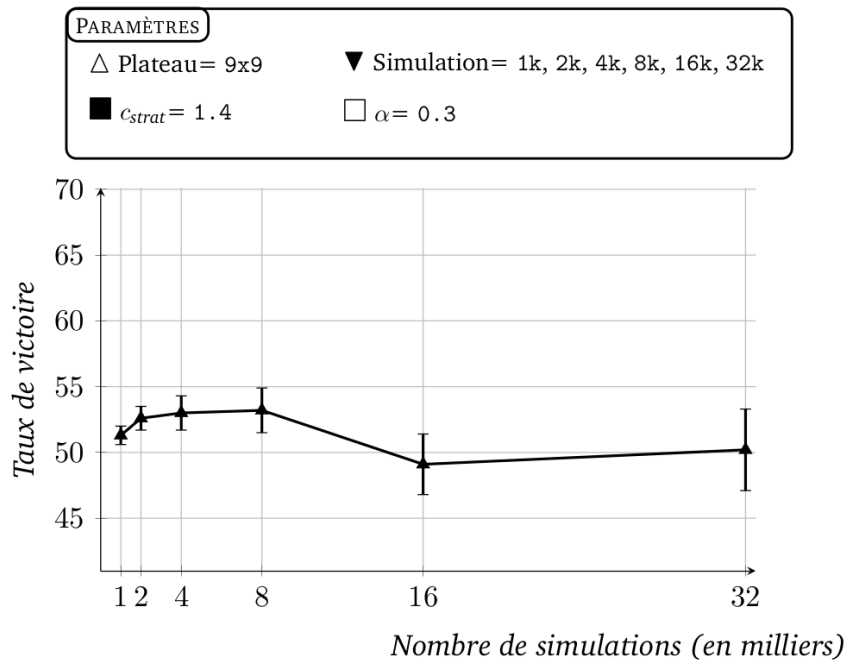


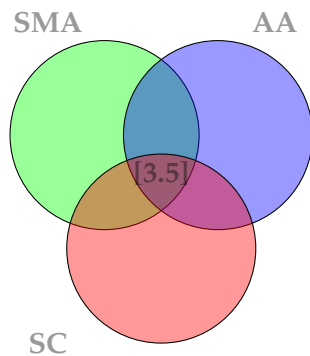
FIGURE 3.17 – Gain obtenu avec l’ajout de stratégies locales pour le Clobber Solitaire suivant le nombre de simulations

3.5 Apprentissage et régulation du trafic routier

Mots-clefs : apprentissage par renforcement, simulation de véhicules connectés, perception/action, régulation du trafic

Les véhicules sont équipés de plus en plus de capteurs et de systèmes de communication. Cette évolution permet d’envisager le trafic routier du point de vue collaboratif, en régulant les comportements des véhicules de manière à optimiser différents critères tels que la réduction des congestions, la limitation de la pollution ou l’amélioration de la sécurité. Au cours de ce projet, nous nous intéressons à développer des mécanismes de contrôle permettant de maximiser ces objectifs.

Tel que nous l’avons vu en section 3.3, la pertinence de la décision prise par un système autonome repose directement sur sa capacité à percevoir son environnement. La politique d’action est également essentielle et doit équilibrer l’exploration et l’exploitation au cours de l’apprentissage tel que présenté en section 3.4.



Au cours de ce projet, nous avons investigué ces deux aspects pour la régulation du trafic routier, avec d'une part la perception par discrétisation de signaux continus du système formé par l'ensemble des véhicules, mais également le développement d'une politique d'action adaptative visant à fluidifier le trafic par la communication de consignes aux véhicules connectés.

Ce travail s'est réalisé dans le cadre d'une collaboration avec le LICIT (Laboratoire d'Ingénierie Circulation Transport) initiée au cours de la thèse de Julien Monteil, et qui s'est concrétisée par un travail en commun avec la thèse de Maxime Guériau (financement ARC7) sur le développement d'un modèle multi-agents pour la régulation de véhicules coopératifs.

3.5.1 Des véhicules connectés pour la régulation

La figure 3.18 représente l'architecture telle qu'on la projette pour les années à venir sur les voies autoroutières. On distingue deux types de véhicules. D'une part, des véhicules autonomes, qui vont pouvoir accueillir des consignes de régulations (vitesse, interdistance, changements de voies, etc.). D'autre part, des véhicules plus conventionnels non équipés dont le nombre devrait diminuer au cours des années.

Différentes unités de bord de route (UBR) permettent de capter différentes informations, par exemple le nombre de véhicules par minute. Elles permettent également de communiquer des consignes de régulation aux véhicules équipés. Les véhicules autonomes sont eux-mêmes dotés de capteurs locaux essentiels à leur perception et à leur comportement local. Cette perception peut être étendue et fiabilisée lorsque les véhicules connectés échangent leurs propres perceptions, comme présenté dans la section 3.2.2.

3.5.2 Simulation et lois de poursuite

Afin de simuler l'activité routière de façon réaliste, nous avons utilisé pour les véhicules (autonomes ou non) la loi de poursuite *Intelligent Driver Model* (IDM), qui est un modèle sans collision capable de reproduire les instabilités du flux de trafic [Treiber13]. La décision de changement de voie quant à elle est modélisée grâce à la stratégie opportuniste nommée MOBIL (*Minimizing Overall Braking Induced by Lane Changes*) [Kesting07].

Ainsi défini, chaque véhicule peut être simulé (variation de vitesse et changement de voie). Les véhicules connectés peuvent en complément percevoir les signaux émis

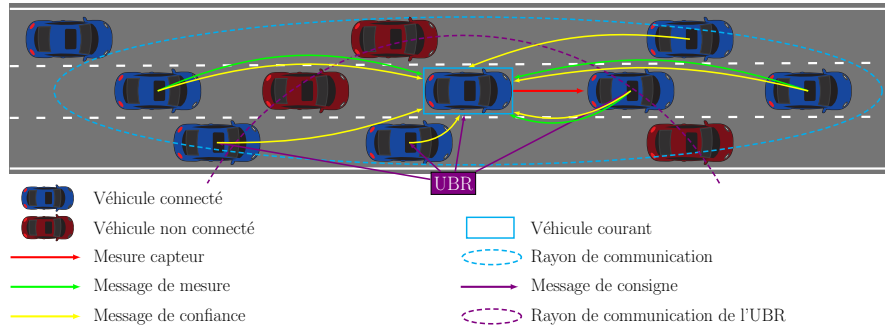


FIGURE 3.18 – Régulation de trafic par véhicules autonomes

par l'UBR afin d'adapter leurs comportements et appliquer les consignes de la stratégie de contrôle apprises pour le système.

Les modèles de suivi de véhicule permettent de calculer l'accélération du véhicule courant i suivant la formule $\ddot{x}_i = f(\dot{x}_i, \Delta x_i, \Delta \dot{x}_i)$ où f est une fonction non linéaire générale, \dot{x}_i la vitesse du véhicule courant, Δx_i l'inter-distance (distance relative au prédécesseur immédiat $i + 1$) et $\Delta \dot{x}_i$ la vitesse relative entre les véhicules courant et précédent.

Le modèle IDM choisi (*Intelligent Driver Model*) exploite une variable de freinage et une variable d'accélération complétées par des paramètres facilement interprétables. La fonction IDM f_{IDM} s'écrit :

$$\ddot{x}_i = a \left[1 - \left(\frac{\dot{x}_i}{v_0} \right)^\gamma - \left(\frac{s^*(\dot{x}_i, \Delta \dot{x}_i)}{\Delta x_i} \right)^2 \right] \quad (3.1)$$

Sachant que $s^*(\dot{x}_i, \Delta \dot{x}_i) = s_0 + \max(0, \dot{x}_i T + \frac{\dot{x}_i \Delta \dot{x}_i}{2\sqrt{ab}})$ est l'inter-distance désirée, a et b sont respectivement l'accélération maximale et la décélération désirée, v_0 est la vitesse désirée, s_0 est la distance minimale à l'arrêt, T est l'intervalle de sécurité temporel désiré, et γ est un facteur d'agressivité de l'accélération.

3.5.3 Construction d'une représentation par apprentissage concurrent

Dans le cadre de cette application, une UBR doit ainsi communiquer des incitations aux véhicules connectés suivant la perception faite de l'état du trafic.

Selon le modèle proposé, cet apprentissage est réalisé par renforcement, tel que présenté sur la figure 3.19. D'une part, différents agents de discrétisation permettent de définir des classes qui constituent des états ou perceptions pour le processus de décision. Ces agents peuvent être connectés suivant différents agencements aux variables captées. Ils exploitent différents algorithmes de classification non supervisés (GNN, K-

means, etc.) afin de définir les états qu'ils proposent. Chaque agent de discrétisation représente ainsi une vue de l'état courant pour le système et pourra participer à la perception sur laquelle s'appuie la politique d'action.

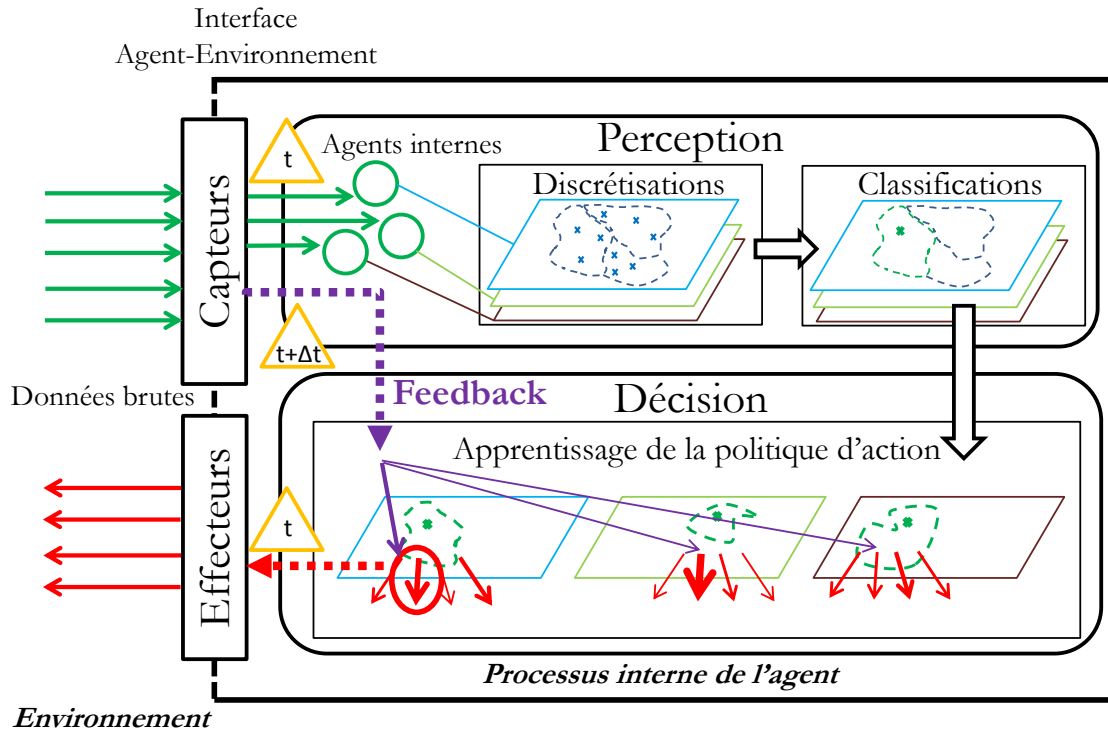


FIGURE 3.19 – Aperçu du modèle de régulation

Le système formé permet ensuite de renforcer les actions évaluées les plus efficaces parmi les couples état-action concurrents disponibles. Pour ce faire, parmi les algorithmes de bandit à n bras de la littérature permettant d'équilibrer les mécanismes d'exploration et de renforcement, nous avons utilisé *Upper Confidence Bound* (UCB) [Auer02].

Ce modèle a été implémenté et simulé sur différents scénarios à partir de données réelles issues de boucles électromagnétiques qui déterminent le flux des véhicules à traiter. Le jeu de données est composé d'un mois (juin 2001) de données boucles, récoltées entre 8h et 9h sur l'autoroute A6 au niveau d'une insertion générative de congestions. Nous évaluons le gain apporté par une stratégie de régulation d'une UBR sur le flux en faisant l'hypothèse d'un pourcentage fixé à 30% de véhicules connectés, répartis de façon homogène parmi les véhicules non connectés.

Concernant la perception du système, nous avons introduit deux agents : 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 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. Un apprentissage préalable non supervisé est réalisé afin de fixer les agents de discrétisation, sur l'ensemble des scénarios représentant les configurations rencontrées.

Toutes les 120 secondes simulées, l'UBR définit la politique d'action à transmettre aux véhicules connectés : pas de consigne, consigne de changement de voie (droite à gauche – gauche à droite), consigne de modification d'interdistance, consigne de limite de vitesse (110 km/h–50 km/h).

L'UBR reçoit régulièrement un *feedback* comme signal qui correspond à la vitesse moyenne glissante mesurée la plus faible parmi les tronçons d'autoroute qu'il doit maximiser. L'apprentissage de la politique d'action est réalisé sur les différents scénarios réalistes récoltés sur l'autoroute A6. Une fois l'apprentissage réalisé, on présente le comportement observé sur 3 scénarios issus de données réelles et présentés sur la figure 3.20.

Un temps total passé faible indique un écoulement fluide du flux. La vitesse moyenne permet d'observer l'homogénéisation. Le pourcentage de congestion fournit quant à lui un aperçu de l'état de trafic et de l'ampleur des perturbations. Plusieurs politiques sont comparées : pas de consigne (pas d'UBR), apprentissage réalisé uniquement avec D_1 , D_2 , D_1 et D_2 simultanément.

Pour la simulation (b), on n'observe pas de congestion, les stratégies de contrôle n'apportent pas d'amélioration. On remarque que les politiques évaluées ne permettent ainsi pas d'améliorer l'approche de référence. Alors que la représentation D_2 seule entraîne une dégradation, la coupler à D_1 permet d'obtenir des résultats similaires à l'approche de référence.

Nous voyons que 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 et dégradent les performances du système. L'utilisation de stratégie de régulation permet de limiter significativement la formation de congestions, et réduit ainsi les temps de parcours. On observe cela par l'indicateur *Temps total passé* sur la section concernée. Nous observons que l'agent D_1 propose une perception pertinente pour la régulation pour la simulation (a). Pour la simulation (c), les meilleures performances sont obtenues en exploitant les agents D_1 et D_2 conjointement.

Nous avons pu constater une grande sensibilité aux appariements réalisés entre les agents perception et les flux d'information disponibles. Une version étendue de ce modèle a été proposée afin de permettre de faire évoluer les agents discrétiseurs eux-mêmes au cours de l'apprentissage, et de les connecter dynamiquement aux informations captées mises à disposition.

Des résultats complémentaires concernant ce projet sont présentés dans [Guéreau17].

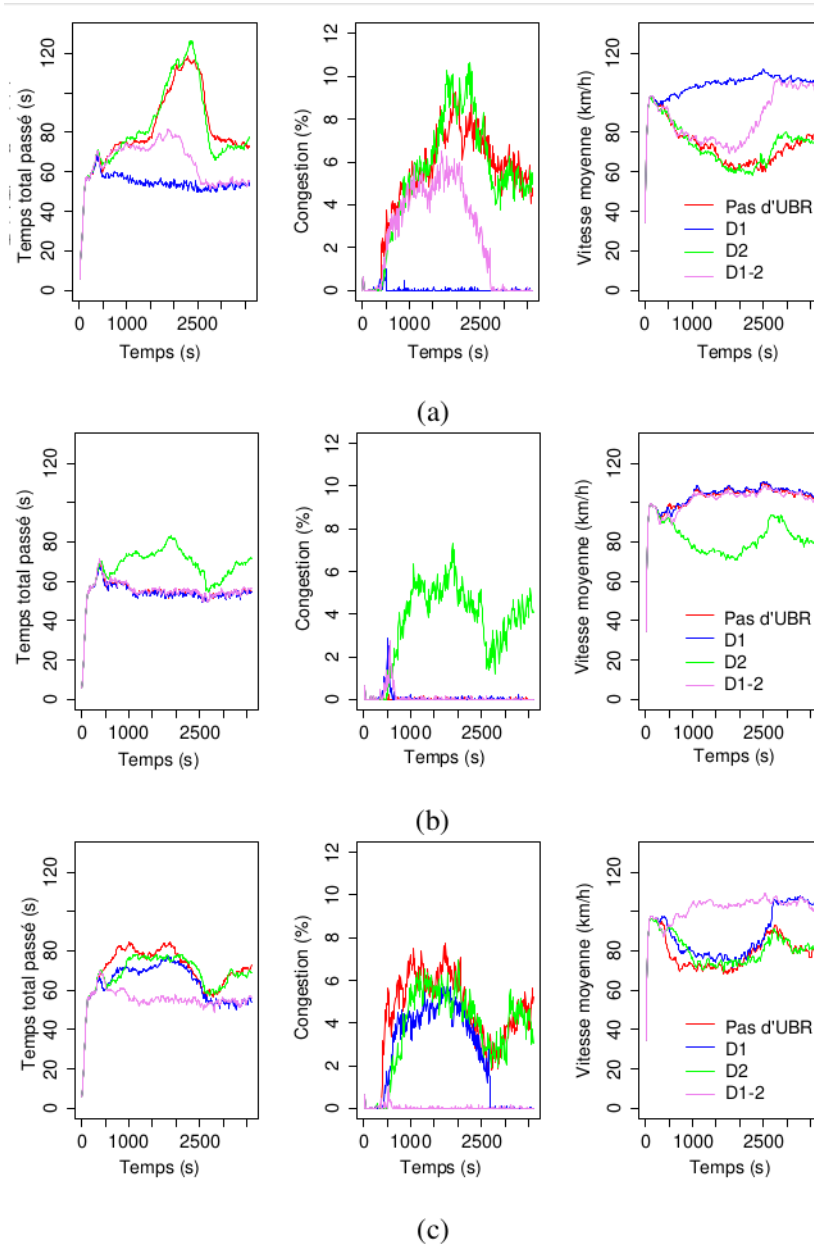
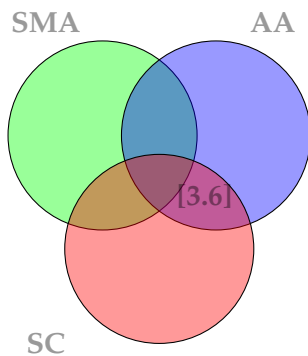


FIGURE 3.20 – Résultats des 3 simulations : (a) Lundi 18, (b) Jeudi 21 and (c) Mardi 28.

3.6 Apprentissage continu non supervisé d'objets à partir d'un flux vidéo

Mots-clefs : apprentissage continu non supervisé, flux vidéo, continuité spatio-temporelle, rôle de l'action, détection de nouveauté



Un système autonome acquiert progressivement des informations du monde extérieur au cours son expérience, à l'image d'un nourrisson qui apprend. Cet apprentissage soulève, entre autres, des problèmes d'apprentissage non supervisé, en ligne et continu dans des environnements non stationnaires. Dans cette section, nous nous intéressons à la capacité d'un robot autonome à découvrir et à reconnaître des objets localisés dans son champ visuel.

Étudier le problème d'apprentissage à partir d'un flux vidéo sans supervision a été un sujet d'intérêt partagé avec Stefan Duffner de l'équipe Imagine du LIRIS et Mathieu Lefort de l'équipe SyCoSMA. Cet élan d'intérêt a donné lieu à deux stages de master en co-encadrement qui nous ont permis de proposer un modèle d'apprentissage par réseaux de neurones siamois présenté en section 3.6.1. Ce modèle exploite les régularités spatio-temporelles du flux vidéo, il a été publié dans [Medjkoune17].

Par la suite, nous avons pu collaborer à travers un stage et une thèse, ce qui a également été l'occasion pour nous d'inviter au sein du projet Mathieu Guillermin, chercheur en humanités (philosophie) à l'UCLy, afin d'échanger nos positionnements respectifs sur les mécanismes de construction de la connaissance.

Comme présenté en section 3.6.2, nous avons ainsi pu étudier à l'occasion de cet encadrement de stage le rôle de l'action dans le cadre d'un apprentissage non supervisé [Devillers22].

Le travail de thèse de Ruiqi Dai présenté en section 3.6.3 a permis de développer des mécanismes de détection de nouveauté et de reconnaissance à partir d'un flux vidéo d'images [Dai21a], [Dai21b].

3.6.1 Exploiter les régularités spatio-temporelles d'un flux vidéo

Au cours de cette étude [Medjkoune17], nous nous sommes intéressés à l'apprentissage d'objets dans un flux vidéo en exploitant la continuité spatio-temporelle des objets présents dans le champ de perception.

Dans ce travail, nous étudions la capacité d'un robot autonome à découvrir et à reconnaître des objets localisés dans son champ visuel. Deux problèmes principaux se

Section 3.6. Apprentissage continu non supervisé d'objets à partir d'un flux vidéo 53

posent dans ce cas : la détection et le suivi des objets dans cet environnement ouvert, ainsi que l'apprentissage de représentations visuelles de manière incrémentale et non supervisée.

De multiples solutions robustes existent pour détecter, suivre et reconnaître des objets ou des personnes dans des images et des flux vidéo [Leibe07, Ess08, Jafari14, Leal-Taixé16, Borji16]. Cependant, ces méthodes sont conçues et pré-entraînées pour des catégories d'objets spécifiques, comme des visages, des personnes ou des voitures, et dans un environnement dédié. Elles ne sont ainsi pas applicables pour toutes catégories d'objets. Pour fournir plus d'autonomie aux algorithmes de perception visuelle, d'autres travaux ont proposé des méthodes d'apprentissage non supervisé, basées sur des réseaux neuronaux, capables d'apprendre des représentations internes discriminantes à partir de données non étiquetées et de s'adapter à un contexte donné en utilisant un apprentissage par transfert.

Par exemple, Wang et al. [Wang15] utilisent un grand ensemble de données non étiquetées de vidéos et des techniques spécifiques d'extraction et de suivi d'objets de bas niveau pour apprendre des représentations visuelles, puis exploitent ce modèle pour apprendre des objets spécifiques avec des données étiquetées dans un second temps. Une autre approche proposée par Liang et al. [Liang15] consiste à mettre à jour continuellement la représentation visuelle apprise en utilisant des objets reconnus à partir de vidéos. Cependant, leur méthode incorpore une quantité considérable de connaissances préalables en utilisant un réseau de neurones convolutionnel (CNN) pré-entraînés sur un grand ensemble de données étiquetées.

Ces méthodes peuvent s'apparenter à notre approche du problème, car elles utilisent des techniques d'apprentissage non supervisées ou faiblement supervisées pour l'apprentissage de représentations visuelles exploitant la cohérence temporelle dans le flux vidéo. Bien qu'elles montrent une haute précision de détection et de reconnaissance d'objets, elles exploitent malgré tout un grand ensemble de données d'images pour être efficaces ce qui limite leur autonomie. Ces approches reposent également sur des pré-entraînements et des architectures complexes de réseaux neuronaux profonds coûteux en calcul.

Dans le modèle que nous proposons, nous abordons ces problèmes en utilisant une approche faiblement supervisée pour la reconnaissance d'objets basée sur un réseau de neurones convolutionnel SNN [Bromley94, Hadsell06] ne nécessitant pas de donnée étiquetée ou pré-entraînement supervisé. Notre méthode n'utilise aucune connaissance préalable sur l'environnement et emploie uniquement des techniques de filtrage visuel relativement simples inspirées par la cognition humaine (comme la saillance) pour extraire des objets candidats.

Sur la figure 3.21, des propositions de proto-objets sont tout d'abord extraites (1). Nous suggérons une méthode simple pour les propositions d'objets inspirée par le système visuel humain. Dans une scène visuelle courante, les objets se distinguent de leur environnement par la couleur, l'intensité, la texture, l'orientation, la profon-

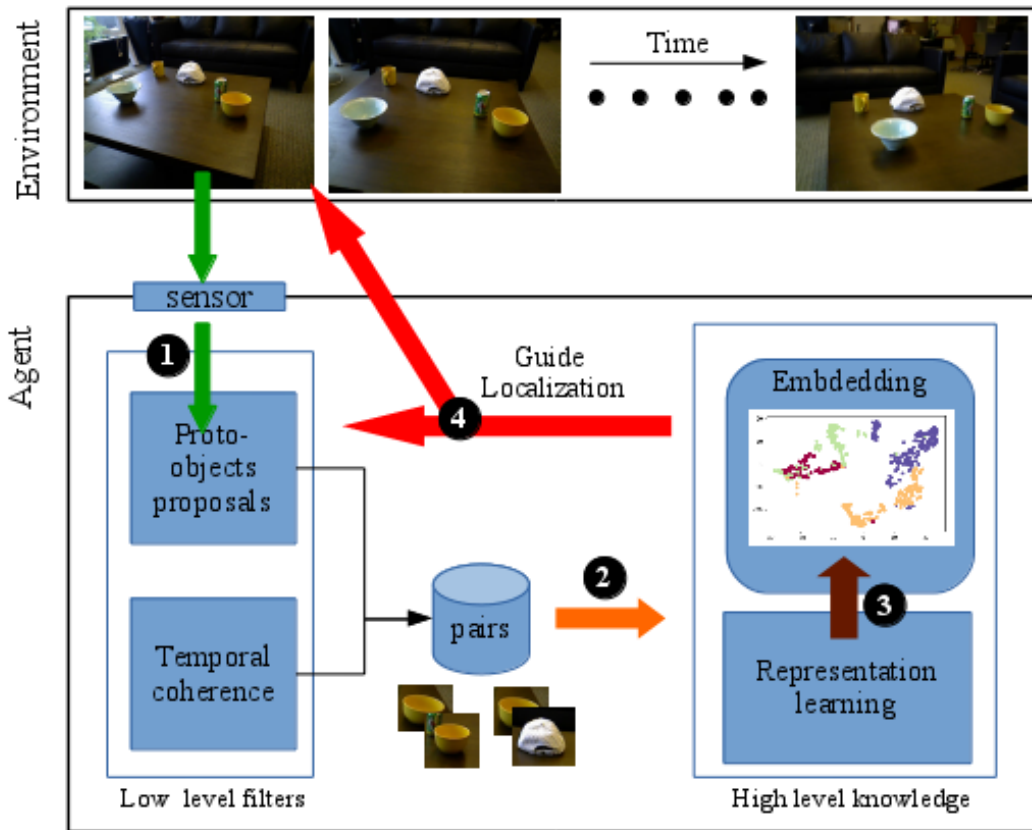


FIGURE 3.21 – Vue générale de l’approche

deur et d’autres caractéristiques. Le système visuel humain est capable de détecter des régions qui diffèrent de leur environnement en utilisant la saillance de bas niveau [Treisman80, Nothdurft05]. En nous basant sur cette observation, nous proposons de calculer la saillance comme un indicateur de présence d’objets à apprendre.

À partir du capteur d’objets ainsi constitué, il est alors possible de calculer des probabilités de similarité d’objet au cours du temps qui exploitent la position des objets dans le champ visuel (2), et de suivre les objets au cours du temps.

Ce suivi permet de constituer des paires similaires (différentes vues du même objet), que l’on exploite comme un signal de supervision pour le réseau, dans le cadre d’un apprentissage contrastif. Différentes images similaires associées à un même objet produisent par un réseau de neurones des représentations latentes que l’on rapproche au cours de l’apprentissage (3). La valeur moyenne de ces représentations correspond alors au concept de l’objet appris. Un nouvel objet peut alors être classé suivant la distance de sa représentation (vecteur latent généré par le réseau de neurones) à l’ensemble des

Section 3.6. Apprentissage continu non supervisé d'objets à partir d'un flux vidéo 55

concepts ou catégories d'objets appris (vecteurs des différents centres de classes).

En perspective de ce projet, la dernière partie du modèle s'intéresse à exploiter les représentations apprises pour aider à la reconnaissance des proto-objets (4), selon une démarche constructiviste ou développementale. Doté de cette fonctionnalité, le système pourrait améliorer de façon auto-catalytique sa capacité à percevoir et sélectionner les objets d'intérêt puis affiner son apprentissage.

Les résultats préliminaires présentés dans [Medjkoune17] montrent qu'il est possible d'exploiter les régularités spatio-temporelles de flux vidéo pour apprendre sans connaissances préalables. Sur le *dataset* étudié, les performances utilisant ce signal de spatio-temporel sont similaires à celles d'un apprentissage supervisé pour lequel les objets sont étiquetés manuellement.

3.6.2 Rôle de l'action pour l'apprentissage de représentations visuelles

Les théories sensori-motrices reposent sur des preuves substantielles en neurosciences, en psychologie du développement et en sciences cognitives. L'affirmation principale étant que les actions, et plus particulièrement les changements sensoriels induits par les actions motrices, jouent un rôle clé dans l'apprentissage d'un modèle prédictif du monde et dans sa perception [Mossio08]. Par exemple, un chaton qui ne peut pas marcher, c'est-à-dire qui reçoit uniquement un flux visuel de manière passive, apprendra des représentations visuelles erronées [Held63]. Le rôle de l'action est également mis en avant dans la notion d'affordance en psychologie [Gibson66], où un objet n'est pas défini par un ensemble de propriétés, mais par les interactions qu'il suscite pour l'agent. Selon la théorie des contingences sensori-motrices, l'action peut même jouer un rôle dans une certaine forme de conscience [O'Regan01a].

Ces concepts sont également liés aux théories de l'enactivisme et de l'incarnation qui stipulent que le corps, en tant que structure permettant d'interagir avec le monde, est nécessaire à l'émergence d'un comportement intelligent [Froese09]. Tout en contribuant à l'apprentissage des représentations, les actions pourraient également viser à percevoir des régions pertinentes de l'environnement, ce qui envisage la perception comme un processus actif. De cette manière, les actions seraient nécessaires pour accumuler des preuves de l'état actuel du monde, comme unifié dans le principe de l'énergie libre, par exemple [Friston11]. En ce qui concerne la vision, c'est par exemple le rôle des saccades qui permettent d'obtenir des aperçus successifs d'une scène visuelle [Friston12].

Pour ce projet, nous étudions ainsi la qualité de la représentation apprise des objets perçus lorsque l'on intègre l'action dans le processus de perception. Afin d'étudier le rôle de l'action dans l'apprentissage, on simule des saccades perceptives pour le système apprenant. La figure 3.22 illustre sur une image du dataset MNIST le procédé perceptif de la perception au temps (t), perception X localisée, vers la perception au temps (t+1), perception Y localisée, suite à l'action réalisée.

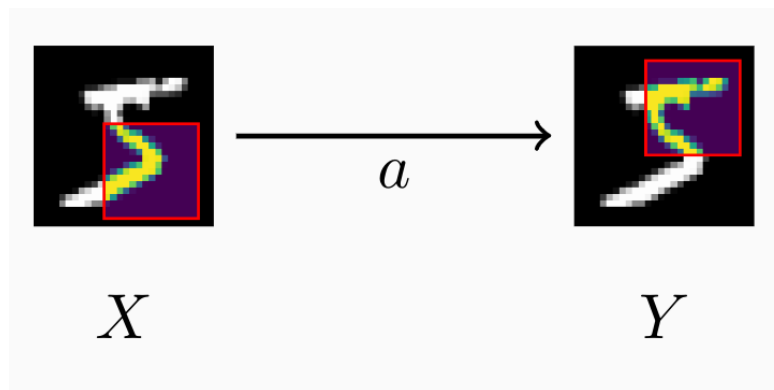


FIGURE 3.22 – Déplacement par une action de la fenêtre de perception

Pour le modèle proposé, nous définissons deux tâches d'apprentissage conjointes par réseaux de neurones. D'une part, une tâche de reconstruction de l'image perçue, mais également une tâche de prédiction de la perception suivante à partir de la perception courante et de l'action réalisée. Pour cette deuxième tâche (prédiction), plusieurs modèles sont étudiés suivant le point d'entrée de l'action dans le réseau.

Selon le modèle utilisé, l'action peut en effet intervenir (signal concaténé) à deux emplacements : PreLSTM ou PostLSTM. Elle peut également être désactivée. Dans ce cas, la deuxième tâche est également une simple tâche de reconstruction, car sans action il n'est pas possible de prédire la perception attendue. Cette désactivation a pour utilité d'établir un point de comparaison avec/sans action.

Sur la figure 3.22, nous utilisons pour la première tâche (partie supérieure) un auto-encodeur variationnel convolutionnel (VAE) [Kingma13], qui réduit la dimensionnalité de l'aperçu actuel en le projetant dans un espace latent, puis en le reconstruisant.

Pour la deuxième tâche (partie inférieure), nous alimentons la branche du modèle avec les perceptions consécutives encodées. Pour cela, nous utilisons un réseau de neurones à mémoire à long terme (LSTM) [Hochreiter97] qui permet "d'accumuler" les perceptions consécutives. Cette accumulation est en effet utile à la construction de la représentation à partir des saccades ou perceptions partielles réalisées. L'action alimente quant à elle le flux à différents points de jonction suivant les scénarios étudiés (PreLSTM, PostLSTM, ou désactivation de l'action).

Le calcul du gradient exploite la fonction d'erreur des deux tâches cumulées, les deux parties du réseau sont intimement liées dans leur dynamique d'apprentissage. Les poids du réseau de neurones du décodeur sont partagés entre les deux tâches.

L'objectif de cette étude est d'évaluer la qualité de la représentation latente pour qualifier les objets. Cette représentation latente est extraite à la sortie du LSTM de la tâche de prédiction.

Section 3.6. Apprentissage continu non supervisé d'objets à partir d'un flux vidéo 57

Nous choisissons comme estimateur de la qualité de la représentation, de réaliser suite à l'apprentissage réalisé que l'on fige, un apprentissage secondaire de classification de la représentation construite. De façon intuitive, plus les performances de classification seront bonnes, plus la représentation construite permettra de discriminer efficacement les objets, et sera donc de qualité.

Ce classificateur a été entraîné sur toutes les représentations produites par le LSTM à partir des aperçus successifs. Ainsi, nous pouvons observer sur la figure 3.23 que la qualité de la représentation s'améliore incrémentalement lorsque de nouveaux aperçus (fenêtre partielle de perception) sont intégrés dans le modèle.

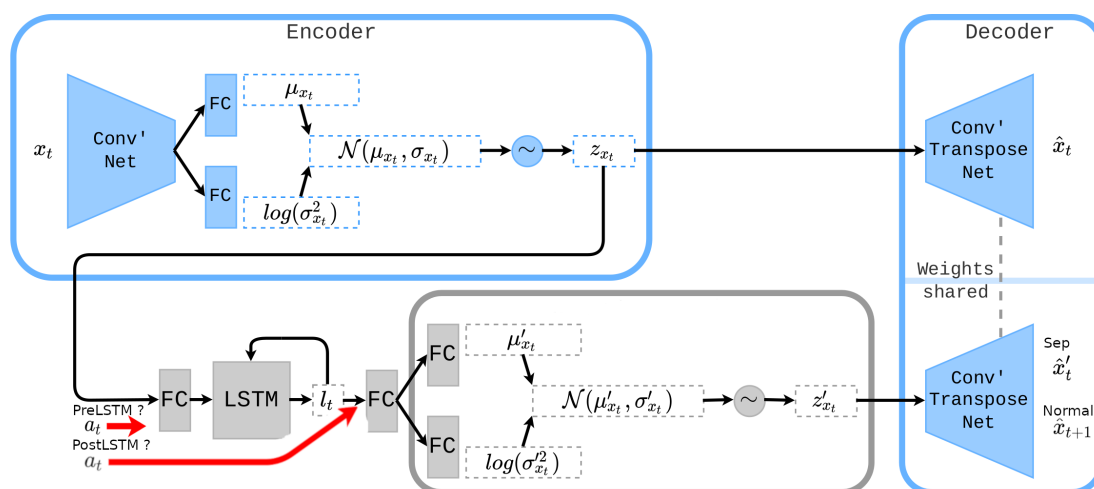


FIGURE 3.23 – Modèle d'apprentissage de deux tâches.

Première tâche (partie supérieure) : reconstruction.

Deuxième tâche (partie inférieure) : avec action \rightarrow prédiction (preLSTM ou PostLSTM), sortie normale prédictive \hat{x}_{t+1} , sans action \rightarrow reconstruction (simple reconstruction de l'entrée \hat{x}_t).

Nous mesurons la précision résultant de différents scénarios sur 3 jeux de données.

Les modèles avec le suffixe "-Sep" réalisent un pré-entraînement de la première tâche de reconstruction, sans utilisation de l'action (et donc avec pour deuxième tâche une reconstruction également). Ils sont ensuite entraînés en ajoutant l'action (action ajoutée en entrée du LSTM (PreLSTM) ou en sortie (PostLSTM), la deuxième tâche est alors bien de prédire la perception suivante. Les modèles sans le suffixe "-Sep" sont entraînés directement sur la tâche prédictive.

Nous observons que les modèles n'utilisant pas l'action lors de l'apprentissage de l'encodeur du VAE (suffixe -Sep) sont bien moins efficaces que leurs homologues utilisant l'action (sans suffixe). La présence de cette tendance pour les deux architectures et pour tous les ensembles de données montre que l'action a son utilité dans l'apprentis-

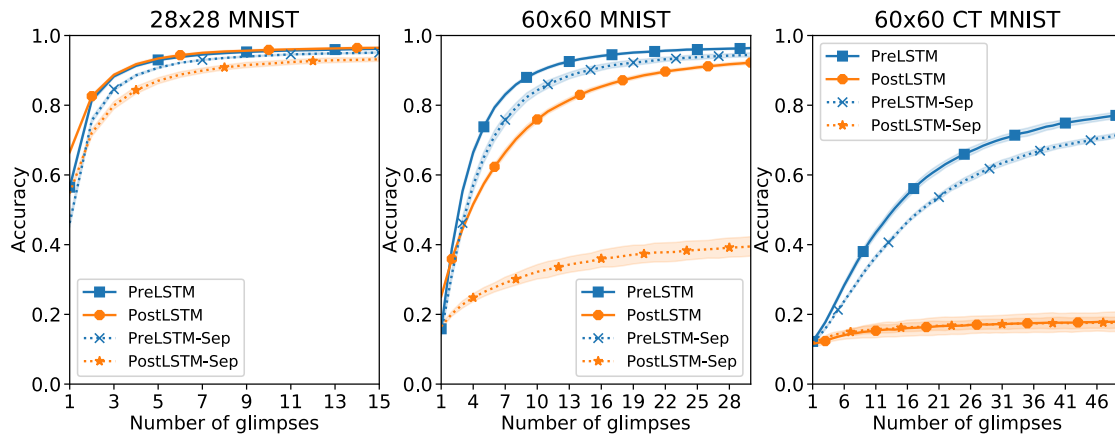


FIGURE 3.24 – Précision des classifications résultantes (comme estimateur de la qualité des représentations apprises) suivant le nombre de perceptions accumulées pour 3 jeux de données

sage des caractéristiques visuelles, et semble être bénéfique pour l'extraction de caractéristiques significatives.

Nous constatons également que tous les modèles intégrant l'action dans le LSTM (architectures PreLSTM) fonctionnent mieux que leurs équivalents intégrant l'action après le LSTM (architectures PostLSTM), sauf pour l'ensemble de données MNIST en 28×28 . Dans ce dernier cas, les performances sont similaires, ce qui peut s'expliquer par la plus grande simplicité de l'ensemble de données. Comme détaillé dans [Deville22], cette différence de performance entre les architectures Pre et Post tend à montrer que prendre en compte l'action dans les représentations, c'est-à-dire dans le LSTM, aide à construire de meilleures représentations de l'environnement. De plus, comme cette tendance s'accroît avec la difficulté croissante de l'ensemble de données, la présence de l'action dans les représentations semble être plus importante pour les tâches complexes.

3.6.3 Détection d'objets en apprentissage continu

De nombreux travaux dans le domaine de l'apprentissage profond traitent souvent le problème de l'apprentissage de la représentation en supposant des données d'entrée i.i.d. (indépendantes et identiquement distribuées), ce qui correspond à un contexte idéal d'apprentissage pour ce type d'approches. Bien que ces travaux obtiennent des résultats très convaincants, cette hypothèse semble peu réaliste lorsque les agents doivent apprendre de manière autonome. Dans le cadre d'un apprentissage selon lequel un agent est en interaction avec son environnement, il est peut probable ou même impos-

Section 3.6. Apprentissage continu non supervisé d'objets à partir d'un flux vidéo 59

sible que l'ensemble des objets à apprendre se présentent de façon «équilibré», selon des apparitions aléatoires parmi tous les objets à apprendre.

Le manque de contrôle sur la manière dont les objets sont perçus par le système conduit à des phénomènes d'oubli catastrophique (*catastrophic forgetting*) pour les objets qui n'ont pas été perçus depuis longtemps, ce qui reste une limitation à surmonter et diminue considérablement la performance de l'apprentissage. La capacité d'apprendre de nouvelles connaissances provient en effet de la plasticité du réseau, ce qui en contrepartie estompe et provoque l'oubli des connaissances acquises sur le long terme (dilemme de stabilité-plasticité).

La littérature sur l'apprentissage continu avec les réseaux neuronaux [Rebuffi17, Rao19] répond partiellement à cette problématique, par des modèles souvent supervisés, qui nécessitent afin de classifier efficacement de nouvelles observations d'objets appris, un étiquetage au cours de l'entraînement, ou après l'entraînement pour attribuer correctement les nombreux clusters identifiés à des catégories d'objets significatives.

Une difficulté rencontrée dans le cas de l'apprentissage continu et non supervisé correspond à la tendance à sursegmenter au cours de l'apprentissage les catégories en de nombreux clusters additionnels [Furao06], ce qui nuit à la qualité de l'apprentissage. La sur-segmentation crée des catégories d'objets artificiellement, et limite ainsi l'autonomie du système en nécessitant un traitement correctif de fusion. Le travail réalisé au cours de cette thèse s'intéresse à introduire avec parcimonie les nouvelles catégories rencontrées au cours d'un apprentissage, à travers des mécanismes de détection de nouvelles catégories (détaillé dans [Dai21a]), mais s'intéresse également la reconnaissance de catégories précédemment observées (détaillé dans [Dai21b]).

Nous proposons ainsi d'enrichir le modèle de réseau neuronal profond CURL [Rao19] présenté sur la figure 3.25. Cette architecture s'appuie sur un auto-encodeur variationnel (tel qu'utilisé dans la section précédente), mais prévoit un composant dédié à chaque nouvelle classe introduite. L'encodage est commun à toutes les classes, les composants sont ensuite activés suivant les composants des classes identifiées (une classification de la valeur de l'espace latent est appliquée).

Ce modèle permet d'atténuer l'effet de l'oubli catastrophique (dans ce cas, sur la partie commune d'encodage des objets) en appliquant un mécanisme de réapprentissage génératif [Rao19] (*generative replay*). L'apprentissage est ainsi complété par un processus générant continuellement des exemples d'entraînement synthétiques des classes précédemment apprises de façon équilibrée, et les utilisant comme données d'apprentissage complémentaires afin de limiter les phénomènes d'oubli du réseau.

Cependant, cette approche a également tendance à sur-segmenter et introduire des composants pour des catégories qui ne devraient pas être créées. CURL utilise un mécanisme de détection de nouveauté (*outlier detection*) assez simple, basé sur la fonction de perte du modèle, plus l'objet perçu est éloigné de ce qui est connu, plus il a tendance à produire une valeur de perte importante.

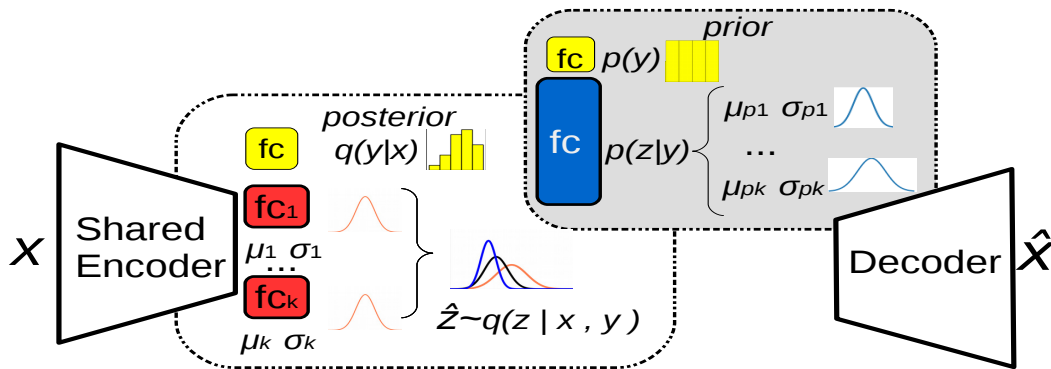


FIGURE 3.25 – Aperçu de l’architecture neuronale de CURL [Rao19] : un auto-encodeur variationnel qui inclut un composant dédié pour chaque classe représentée. X représentant les images d’entrée, y la variable de catégorie. L’encodeur projette les images en entrée vers une représentation partagée pour toutes les catégories. Sa sortie est utilisée comme entrée d’un classifieur (fc, en jaune) qui estime la catégorie de l’objet $q(y|x)$; et met à jour les paramètres μ_k, σ_k du ou des composants correspondants k . La distribution a posteriori de la variable latente $q(z|x, y)$ est approximée pour chaque composant. De plus, la distribution a priori $p(z|y)$ de la variable latente z suit une distribution gaussienne. Ensuite, l’image \hat{X} est reconstruite à partir du \hat{Z} rééchantillonné en utilisant le décodeur.

La première contribution de ce travail s’intéresse ainsi à détecter automatiquement les nouvelles classes introduites plus efficacement. Dans ce contexte, nous introduisons l’utilisation du test statistique de Page-Hinckley [Page54] à la vraisemblance de la fonction de perte. Ce test est reconnu pour son efficacité pour la détection de dérive conceptuelle et de changements abrupts dans des données séquentielles.

Nous avons pu évaluer la pertinence de l’apprentissage, en observant les matrices de confusion sur deux datasets représentant des chiffres et des vêtements. Pour cet apprentissage continu, les objets sont présentés par classes (sans retour sur les classes parcourues dans un premier temps). Les transitions entre classes sont aléatoires, et doivent être détectées efficacement.

Tout d’abord, on constate sur la figure 3.26 que l’approche CURL a tendance à créer beaucoup de composants additionnels. Le test que nous avons appliqué permet de suivre au plus près l’introduction de nouvelles catégories d’objets, tel que présenté sur la figure 3.27. En effet, le nombre de composants créés correspond ou est très proche des composants étiquetés dans les datasets. Des résultats comparatifs complémentaires sont présentés dans [Dai21a].

Ce premier mécanisme n’inclut pas la reconnaissance de catégories précédemment apprises, mais s’intéresse exclusivement à la détection de nouveautés. Afin de com-

Section 3.6. Apprentissage continu non supervisé d'objets à partir d'un flux vidéo 61

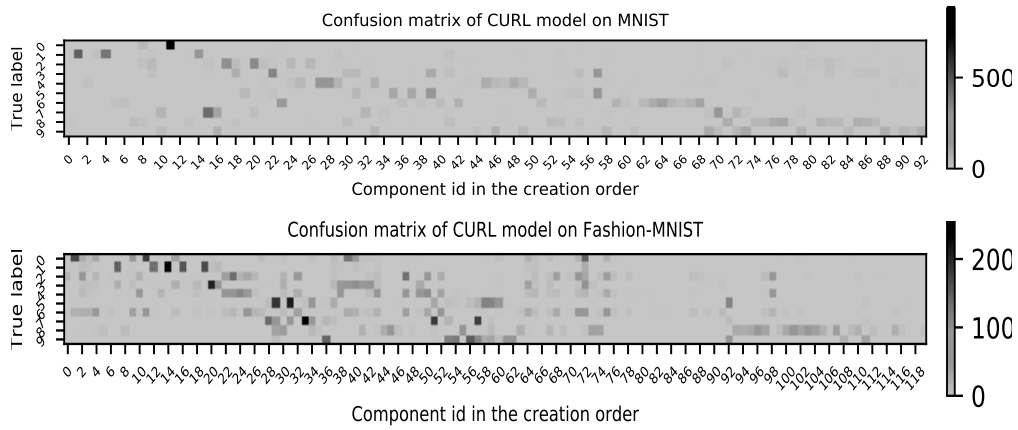


FIGURE 3.26 – Matrices de confusion entre la valeur réelle et les composants inférés pour CURL sur l'ensemble de test MNIST et l'ensemble de test Fashion-MNIST (plus une cellule est sombre, plus elle représente d'instances).

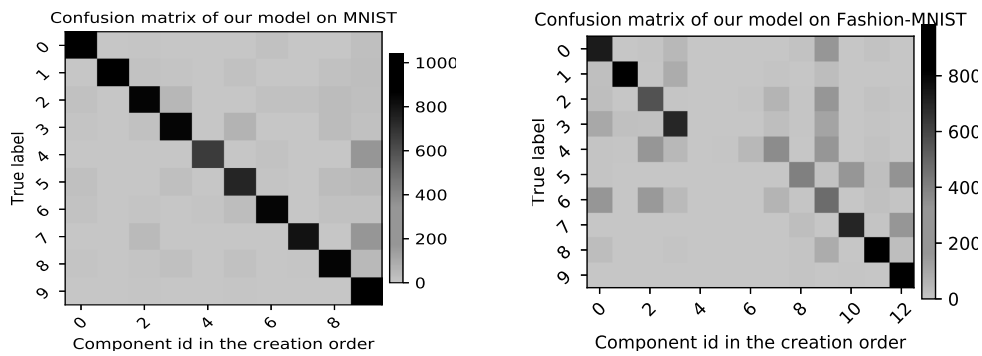


FIGURE 3.27 – Matrices de confusion entre la valeur réelle et les composants inférés pour CURL sur l'ensemble de test MNIST et l'ensemble de test Fashion-MNIST (plus une cellule est sombre, plus elle représente d'instances).

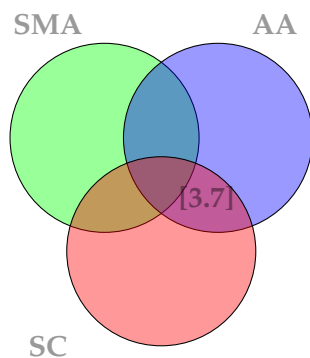
pléter notre proposition et améliorer l'autonomie du modèle, nous avons intégré un deuxième test statistique Hotelling [Hotelling31], permettant également de reconnaître les catégories d'objets déjà identifiées, et qui se prête bien aux distributions normales multivariées manipulées.

Comparativement à d'autres modèles de l'état de l'art, nous montrons dans [Dai21b] que ce deuxième modèle permet de reconnaître des objets appris dans un scénario en ligne tout en évitant de créer et de réaffecter de nouveaux composants pour les catégories déjà créées.

Par rapport à mon positionnement de recherche présenté dans ce mémoire, un agent base son apprentissage sur les interactions qu'il entretient avec son environnement. Bien que cette approche permet d'apprendre de façon continue les objets présentés, la discrimination est cependant réalisée uniquement à partir du flux visuel perçu. Une extension de ce travail consisterait à exploiter un signal sensori-moteur qui inclurait alors l'activité du système en interaction avec son environnement. Dans ce contexte, l'apprentissage réalisé serait plus incarné et moins arbitraire, la capacité à contenir le nombre de composantes créées pourrait être étudiée et être très utile à l'autonomie du système.

3.7 IA et apprentissage d'algorithmes

Mots-clefs : Apprentissage d'algorithmes, apprentissage incrémental, transfert et récurrence, *bootstrap*



La machine de Turing neuronale [Graves14b] publiée il y a quelques années a attiré toute mon attention. Il s'agit d'un modèle original qui exploite un réseau de neurones pour exécuter des algorithmes. La tâche réalisée par le réseau est alors bien différente de l'usage généralement donné aux réseaux de neurones. On souhaite obtenir à partir de paramètres positionnés en entrée du réseau le résultat escompté d'un algorithme en sortie du réseau, et donc manipuler des variables et exécuter des sous-routines. Ceci a attisé ma curiosité et m'a incité à étudier la capacité des réseaux de neurones à exécuter ce type de tâches, avec peut-être à terme des capacités nouvelles de généralisation, mais cette fois algorithmique, pour nos IA. Cette capacité de généraliser des interactions avec un environnement, sur un second ordre, porte en elle de nombreux défis.

J'ai porté ce projet en encadrant 3 stages de master, j'ai associé Mathieu Lefort aux différentes réflexions et à la diffusion scientifique des résultats. Les résultats détaillés de ces travaux sont présentés dans les articles [Nollet20, Armetta23b, Armetta23a].

3.7.1 Un domaine émergent ?

Depuis le premier succès des réseaux de neurones convolutifs sur la classification d'images [Krizhevsky12], les méthodes d'apprentissage profond ont fait progresser l'état de l'art dans de multiples domaines [Goodfellow16]. Cela se traduit par des tâches de plus en plus complexes comme la traduction de langues [Vaswani17b] ou la résolution de positions du jeu Go [Tian19].

La machine de Turing neuronale [Graves14b] a été proposée pour apprendre des procédures algorithmiques, comme le tri de listes. Elle a pour objet de réduire le problème d'entraînement rencontré par les RNN (Récurrent Neural Networks), qui sont turing-complet [Siegelmann95], en ajoutant des mécanismes spécifiques tels qu'une mémoire et des moyens différentiables d'y accéder. Cependant, elle connaît également un problème d'entraînement, avec des performances qui sont difficiles à reproduire et inconstantes [Collier18]. Ces problèmes sont dus à la profondeur de l'architecture proposée [Kaiser15] mais également à la complexité intrinsèque des opérations à apprendre, notamment la dépendance à long terme dans les données.

Les modèles d'apprentissage profond sont malgré tout des approximateurs de fonctions universels [Cybenko89], et donc de bons candidats pour l'apprentissage algorithmique. D'un point de vue théorique, la capacité d'expression d'une architecture neuronale peu profonde peut être suffisante pour toute tâche, mais en pratique, la capacité des réseaux profonds à apprendre des représentations pertinentes à partir des données est bien meilleure [Ba14]. De plus, les RNN (Récurrent Neural Networks) sont complets au sens de Turing [Siegelmann95], mais sont également confrontés à un problème d'entraînement [Graves14b]. Cette limitation tend à être naturellement plus prononcée pour les tâches complexes.

Pour surmonter ces difficultés spécifiques à l'apprentissage d'algorithmes, j'ai engagé des recherches sur l'apprentissage arithmétique pour les multiplications à plusieurs chiffres. Ce problème est d'apparence assez simple, mais ne l'est pas pour une IA autonome.

En particulier, la multiplication à plusieurs chiffres qui peut être résolue par un algorithme. Ce calcul implique de nombreuses opérations et variables dans le flux de calcul (calculer plusieurs multiplications à un chiffre, additionner correctement les résultats intermédiaires afin d'obtenir le résultat final). La complexité d'une telle opération est également liée à la propagation des retenues [Cho19]. L'apprentissage direct de bout en

bout⁴ des multiplications conduit à des performances médiocres [Hoshen16].

Permettre l'apprentissage de cet algorithme spécifique pourra permettre de mieux comprendre les dynamiques d'apprentissages algorithmiques afin de généraliser ensuite à l'apprentissage de tout type d'algorithmes.

3.7.2 Exécution étape par étape

[Hoshen16] montre que la multiplication ne peut pas être apprise de bout en bout (*end-to-end*) en utilisant un perceptron multi-couches.

Un premier modèle proposé dans [Nollet20] s'intéresse à décomposer le problème, et éduquer un réseau de neurones afin qu'il prédise, étape par étape en mimant la façon dont un écolier déroule ses multiplications tel que présenté sur la figure 3.28.

$$\begin{array}{r}
 0023 \quad (1) \\
 \times 0048 \quad (2) \\
 \hline
 0012 \quad (3) \\
 0184 \quad (4) \\
 0010 \quad (5) \\
 + 0920 \quad (6) \\
 \hline
 0110 \quad (7) \\
 1104 \quad (8)
 \end{array}$$

FIGURE 3.28 – Exemple de représentation d'une multiplication de deux opérandes à 2 chiffres. Notez que les signes (+ et ×) et les lignes horizontales ne sont représentées que pour plus de clarté. Les lignes (1) et (2) correspondent aux deux opérandes. Les lignes (3) et (4) (respectivement (5) et (6)) représentent les retenues et le résultat de la multiplication à 1 chiffre de 8 (respectivement 4) par 23. Les lignes (7) et (8) correspondent aux retenues et au résultat de l'addition des lignes (4) et (6), qui est aussi le résultat final de la multiplication globale de 48 par 23, soit 1104.

Le premier modèle décompose l'exécution (inférence des résultats intermédiaires de multiplication à 1 digit, puis à partir de ces résultats intermédiaires inférence du résultat global de la multiplication), et équilibre l'effort d'apprentissage parmi les différentes sous-tâches par un apprentissage actif [Settles09].

Ce premier modèle utilise un perceptron multicouche. Il est alimenté par récurrence manuelle pour les différentes inférences. La représentation des nombres placés en entrée et sortie du réseau est symbolique et utilise un codage de type *one-hot* (codage

4. Une résolution de bout en bout ou *end to end* exploite un réseau par exemple récurrent sans intervention extérieure lors de l'inférence. Une résolution plus manuelle ou *handmade* autorise de replacer et combiner les variables lors de l'inférence récurrente.

selon lequel la position de l'unique bit positionné à 1 dans le vecteur qualifie la valeur représentée).

Les résultats de simulation détaillés dans [Nollet20] montrent que pour des problèmes de taille équivalente à [Hoshen16], avec un réseau de taille et un codage similaire, la multiplication peut être apprise étape par étape avec le modèle que nous proposons.

Ce premier modèle souligne la nécessité de guider l'apprentissage en lui indiquant les étapes intermédiaires, c'est à dire un lui fournissant un "feedback" plus explicite (voir section 2.3.2). Ceci étant, l'algorithme n'est toujours pas exécutable de bout en bout, ce qui fait l'objet de la deuxième contribution ci-dessous.

3.7.3 Algorithme de bout en bout

Une façon courante du *Deep Learning* pour apprendre des tâches complexes est de les décomposer par complexité croissante [Bengio09]. Ceci correspond à un apprentissage par curriculum. Une autre stratégie s'appuie sur l'apprentissage multi-tâches qui consiste à combiner différentes tâches avec des objectifs similaires afin que le réseau puisse mieux généraliser et apprendre la structure sous-jacente des données [Zhang17].

En s'inspirant de cette décomposition, nous montrons dans [Armetta23b, Armetta23a] que l'apprentissage de la multiplication de bout en bout est possible grâce à l'introduction d'un nouveau type de transfert de connaissances basé sur la décomposition chronologique du processus à apprendre (UAT : Unrolling Algorithmic Training), c'est à dire par l'introduction de tâches intermédiaires utiles séquentiellement à la résolution de l'algorithme global.

Le réseau utilisé par notre modèle UAT est désormais récurrent, et de type encodeur/décodeur, tel que présenté sur la figure 3.29.

Les différentes tâches présentées sur la table 3.3) sont exécutées sur le même réseau. L'intuition suivie étant que la tâche globale pourra exploiter l'apprentissage des sous-tâches sur le même réseau pour exécuter la tâche globale de bout en bout.

Tâche	Entrée Encodeur	Sortie Décodeur
st1	(1//2)	(3//4)
st2	(1//2) (3//4)	(5//6)
st3	(1//2) (3//4) (5//6)	(7//8)
tglobale	(1//2) (ligne double vide * 2)	(7//8)

TABLE 3.3 – Sous-tâches et tâche globale associées à une multiplication à 2 chiffres (voir les lignes numérotées sur la figure 3.28, les lignes sont lues et écrites deux par deux ("//"))

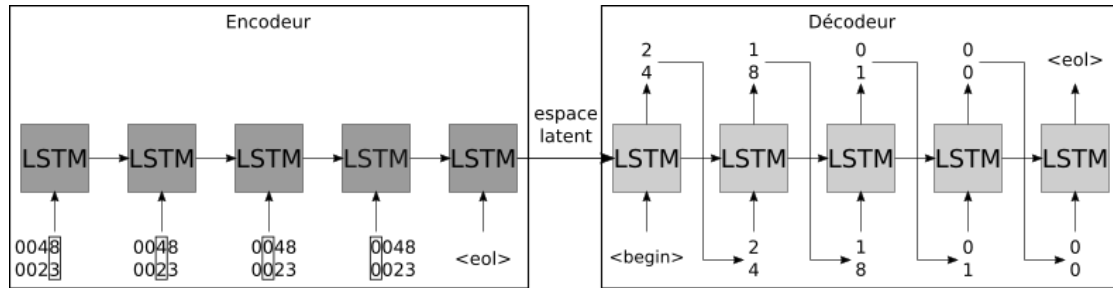


FIGURE 3.29 – Architecture *Seq2Seq*. L’encodeur reçoit successivement les deux chiffres (encadrés) de deux lignes consécutives (les opérands 23 et 48) de droite à gauche. À partir de l’encodage, le décodeur produit de façon récurrente les chiffres, de droite à gauche, des deux lignes suivantes (ici la retenue et le résultat de la première multiplication intermédiaire 8×23). Dans la pratique, chaque chiffre est codé sous la forme d’un vecteur *one-hot*.

Les résultats présentés sur la table 3.4 montrent que la proposition que nous avons faite permet d’améliorer de façon très significative l’apprentissage de la multiplication de bout en bout 4×4 restreinte à un résultats à 7 digits. Nous montrons également qu’un apprentissage de finition (*fine tuning*) peut être appliqué à la tâche globale seule, sachant que l’apprentissage a été préalablement amorcé (*boot-strap*). Un dernier résultat intéressant montre que l’on peut également appliquer un apprentissage de finition à la tâche final tout en modifiant la capacité expressive du réseau (3 lignes de récurrence ajoutées), ce qui souligne la capacité de réorganisation interne du réseau.

	test = tglobale	
	4×4 restreinte (7 chiffres en sortie)	4×4 (8 chiffres en sortie)
Hoshen et al.[Hoshen16]	37.6 %	n.c.
UAT (train = tglobale)	35.87% ($\pm 28.10\%$)	92.42% ($\pm 3.64\%$)
UAT (train = sous-tâches et tglobale)	4.51% ($\pm 1.21\%$)	23.34% ($\pm 14.69\%$)
UAT fine tuning	n.c.	6.68% ($\pm 4.31\%$)
UAT fine tuning (+ 3 récurrences)	n.c.	3.83% ($\pm 1.29\%$)

TABLE 3.4 – Taux d’erreurs pour la multiplication (tglobale)

Je n’ai pas connaissance d’autres travaux exploitant cette forme de transfert algorithmique à travers un réseau récurrent. La dynamique de renforcement entre les sous-tâches peut être observée sur les figures 3.30. La tâche globale est apprise en dernier, suite à l’apprentissage des tâches d’addition et de multiplications. Dans le contexte d’un agent autonome développemental, cette propriété pourrait être intéressante lorsqu’une

sous partie des routines a été préalablement apprise au cours du développement de l'agent, comme un support pour l'apprentissage d'une nouvelle tâche.

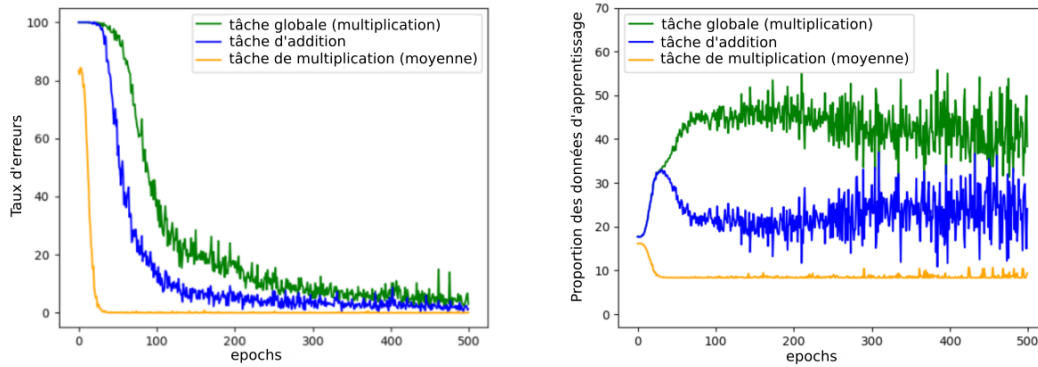


FIGURE 3.30 – (partie gauche) : évolution du taux d'erreur pour chaque tâche (estimé à partir de l'ensemble des données d'apprentissage) utilisé pour l'apprentissage actif. (partie droite) : évolution associée de la proportion des différentes tâches dans l'ensemble des données d'apprentissage

4

Bilan et perspectives

4.1 Bilan

L'intelligence artificielle (IA) a connu des avancées spectaculaires ces dernières années, transformant profondément de nombreux domaines d'application et ouvrant des perspectives inédites pour l'avenir. Ces progrès s'accompagnent cependant de défis considérables.

L'un des enjeux majeurs de l'IA réside dans la conception de structures adaptées pour accueillir et organiser les connaissances. Le positionnement suivi au cours de mes recherches propose au système de construire sa représentation subjective de son environnement, utile à ses propres objectifs. Ce positionnement interpelle, car il ne s'intéresse pas à calquer une représentation réaliste, dont il n'existe pas de version dans l'absolu. Le système s'appuie plutôt sur ses interactions avec son environnement pour valider ce qui peut faire sens.

C'est donc naturellement que les projets de recherche que j'ai sélectionnés et rassemblés dans ce mémoire s'intéressent à cette fameuse extraction et cette fameuse dynamique.

Une des difficultés de ce travail réside dans la diversité des domaines applicatifs étudiés. C'est également une richesse, car ils mettent chacun en avant une spécificité ou une version des défis à aborder du point de vue de la représentation et de la dynamique. Au cours de mes recherches, j'ai également pu expérimenter de nombreuses méthodes informatiques comme des outils d'instanciation pour répondre aux enjeux décrits dans mon positionnement.

J'ai tout d'abord exploré la notion d'apprentissage d'un point de vue décentralisé. Ce fut le cas pour l'émergence d'une structure décentralisée pour le partage de documents dans un réseau pair-à-pair que j'ai beaucoup apprécié par le caractère émergent de la topologie à invariance d'échelle qui se forme dans le réseau, et le contrôle dépourvu de toute centralisation pour ce projet. Les mécanismes de fiabilisation d'informations partagées entre agents m'ont permis d'étudier les phénomènes de robustesse des systèmes multi-agents.

Un de mes centres d'intérêt principaux consiste en l'étude de la construction d'une connaissance (patterns, ontologie, stratégie) que je qualifierais de riche et anthropomorphique, idéalement sans signal de supervision, et que l'on retrouve dans beaucoup des projets auxquels je me suis intéressé. Ainsi, le travail sur l'intelligence ambiante permet d'instancier des mécanismes originaux et agnostiques pour la discrétisation de signaux continus à partir d'une motivation intrinsèque "auto-justifiée". Les *patterns* sont utiles car ils participent à d'autres *patterns*! L'autopoïèse, c'est-à-dire la capacité de se produire soi-même en interaction avec l'environnement, est particulièrement marquée et inspirante pour ce projet.

À un niveau de conceptualisation sémantique plus abstrait, l'étude de l'apprentissage continu de classes d'objets à partir d'un signal vidéo du point de vue des régularités spatio-temporelles, du rôle de l'action et de la similarité statistique a été très instructive sur les capacités d'apprentissage des méthodes à base de réseaux de neurones. Comme présenté dans mon projet de recherche, améliorer la richesse ontologique de ces représentations fait partie des sujets que je souhaite développer.

L'aspect simulation et implémentation intervient en filigrane de beaucoup de projets. C'est notamment essentiel pour les algorithmes d'apprentissage par renforcement étudiés. Dans le cadre de la régulation de trafic routier, le système complexe simulé a pu permettre un apprentissage par renforcement. L'aspect combinatoire est très présent en IA, et abordé sous la forme d'un arbre de recherche complété par une évaluation Monte-Carlo pour les jeux combinatoires a enrichi ma connaissance des méthodes et algorithmes efficaces, pour lesquels nous avons pu proposer une mutualisation des connaissances entre les phases de l'algorithme, mais également introduire un nouveau signal de renforcement issu de réifications de stratégies locales.

L'apprentissage d'algorithmes a également pu faire l'objet de ma recherche, avec pour objectif de peut-être un jour bénéficier de la capacité des réseaux de neurones à généraliser ... des algorithmes! L'algorithme peut à la fois être vu du point de vue de la représentation, mais également du point de vue de la dynamique comme un support d'apprentissage évolutif de niveau supérieur pour une IA tel qu'évoqué en perspectives.

4.2 Perspectives

En perspectives de ce travail, je m'intéresse à l'élaboration de supports de représentation toujours plus qualitatifs, mais également à approfondir l'exploitation de différents *feedbacks* pour guider l'apprentissage.

4.2.1 Vers un support de connaissances toujours plus qualitatif

J'ai été amené à utiliser particulièrement les méthodes à base de réseaux de neurones ou d'autres méthodes comme support de mémorisation. Dans la continuité des travaux que je mène, je souhaite approfondir ces supports du plus bas au plus haut niveau :

- Les possibilités d'**acquisition de connaissances du point de vue de l'apprentissage de motifs sensori-moteurs** : dans la continuité des travaux de discrétisation de signaux continus, avec un accent porté sur l'interaction.
- Les possibilités de **conceptualisation ontologique** : dans la continuité des travaux sur l'apprentissage d'objet par flux perceptif continu. Dans ce domaine, je souhaiterais améliorer l'expressivité de la classification d'objets réalisée suivant plusieurs axes. D'une part, cette classification devrait être contextualisée et non figée. Un objet pourrait être interrogé du point de vue de sa couleur ou de sa forme par exemple, les modèles attentionnels des réseaux de neurones sont pour cela de bon candidats à étudier et étendre. Différents niveaux de représentation (du plus précis au plus abstrait), la notion de composition ou encore d'utilité pour le système pourrait idéalement s'intégrer à la notion de représentation même si les défis sont nombreux.
- Les **possibilités des systèmes à manipuler et traiter leurs propres connaissances**. Les premiers algorithmes que j'ai pu exécuter *end-to-end* manipulaient des variables symboliques (codage *one-hot*). Coupler ces traitements aux données apprises par le système doterait le système de nouvelles capacités utiles à la compréhension de nouveaux motifs algorithmiques. La généralisation algorithmique et les mécanismes de transfert algorithmiques présentés sont également à approfondir, en recherchant de nouveaux domaines d'application pour ancrer et concrétiser la recherche.

Bien que ce défi ne soit pas encore à notre portée tellement la tâche est ambitieuse lorsqu'elle est abordée en autonomie, il serait intéressant de coupler ces différents niveaux de connaissance au sein d'un même système, et de définir ce qui permettrait d'amorcer un tel système. Les questions liées à l'amorçage de l'apprentissage (*cold start problem*) font partie des questionnements essentiels à l'élaboration d'un tel système autonome, car tout apport spécifique à un problème applicatif peut nuire à l'adaptation à de nouveaux problèmes inconnus, en environnement ouvert.

4.2.2 De nouveaux *feedbacks* pour nos IA

La notion d'autonomie implique que le système apprenne en interaction avec son environnement, avec une intervention contenue des connaissances et processus initiaux injectés par son concepteur. Au contraire, le système doit exploiter le plus efficacement possible les signaux perçus de son environnement :

- **Un *feedback* issu de l'interaction avec l'humain.** Lors des différentes phases d'apprentissage, l'enfant est souvent amené à interagir d'abord avec ses parents puis avec ses pairs ce qui participe à façonner la compréhension qu'il a de son environnement. Dans le cadre du développement d'intelligences augmentées, une IA entre en interaction avec un humain pour l'assister dans ses tâches. Les problèmes de co-construction soulèvent leurs propres défis scientifiques. Comment coupler le niveau représentationnel de l'humain avec celui du système? Comment pérenniser la connaissance issue de l'interaction?
- **Un *feedback* issu de l'interaction avec les larges modèles de langue.** Les larges modèles de langues exploitent pour leur apprentissage des corpus de très grande taille, et matérialisent des quantités astronomiques de connaissances. Ils connaissent leurs propres limites, tel qu'un manque d'autonomie face à des situations inconnues. Ils peuvent malgré tout être exploités, par exemple à l'aide d'un agent conversationnel, pour aider un système à construire sa représentation. L'objectif serait encore une fois de coupler différents niveaux de représentation. Afin d'illustrer cela pour le jeu d'échecs, des conseils généraux sur la stratégie seraient alors disponibles par ce biais, de la même façon que les concepts sont manipulés sans ancrage et compréhension dans l'expérience de la chambre chinoise imaginée par Searle. L'enjeu étant alors d'ancrer la conceptualisation du système dans son expérience de son environnement.
- **Un *feedback* issu du calcul.** Les algorithmes à base de renforcement exploitent souvent des modèles simulés de leur environnement. Un apprentissage de l'environnement fidèle permettant cet exercice de simulation virtuel en limitant les biais (*reality gap*) soulève de nombreux défis. Il permettrait cependant de limiter les interactions coûteuses en temps du système avec son environnement. Idéalement, le système pourrait construire sa représentation "mentale" de son environnement, à exploiter comme support de simulation, afin de projeter des tâches dans le monde réel, mais également afin de faire évoluer sa représentation en minimisant les inconsistances avec son environnement "réellement perçu".

Bibliographie

- [Allis94] Louis Victor Allis. *Searching for Solutions in Games and Artificial Intelligence*. Doctorat, Limburg University, Maastricht, Netherland, 1994. 38
- [Androutsellis-Theotokis04] Stephanos Androutsellis-Theotokis & Diomidis Spinellis. *A survey of peer-to-peer content distribution technologies*. ACM computing surveys (CSUR), vol. 36, no. 4, pages 335–371, ACM New York, NY, USA, 2004. 22
- [Armetta06] Frédéric Armetta. *Proposition d'une approche auto-organisationnelle pour le partage de ressources critiques*. Theses, Université Claude Bernard - Lyon I, December 2006. 9
- [Armetta14] Frédéric Armetta, Mohammed Haddad, Salima Hassas & Hamamache Kheddouci. *A self-organized system improving inner topology for data sharing efficiency*. Evolving Systems, vol. 5, pages 109–119, Springer-Verlag, June 2014. 22
- [Armetta23a] Frédéric Armetta, Anthony Baccuet & Mathieu Lefort. *Algorithmic learning, a next step for AI. An application to arithmetic operations*. In Artificial Intelligence and Cognition (AIC), 2023. 63, 65
- [Armetta23b] Frédéric Armetta, Anthony Baccuet & Mathieu Lefort. *L'apprentissage algorithmique, une nouvelle étape pour l'IA. Une application aux opérations arithmétiques*. In CNIA 2023-Conférence Nationale en Intelligence Artificielle, PFIA, pages 31–39, 2023. 63, 65
- [Auer95] Peter Auer, Nicolo Cesa-Bianchi, Yoav Freund & Robert E Schapire. *Gambling in a rigged casino : The adversarial multi-armed bandit problem*. In Proceedings of IEEE 36th annual foundations of computer science, pages 322–331. IEEE, 1995. 40
- [Auer02] Peter Auer, Nicolo Cesa-Bianchi & Paul Fischer. *Finite-time analysis of the multiarmed bandit problem*. Machine learning, vol. 47, no. 2-3, pages 235–256, Springer, 2002. 49

- [Ba14] Jimmy Ba & Rich Caruana. *Do deep nets really need to be deep?* In Advances in neural information processing systems, pages 2654–2662, 2014. 63
- [Bengio09] Yoshua Bengio, Jérôme Louradour, Ronan Collobert & Jason Weston. *Curriculum learning*. In Proceedings of the 26th annual international conference on machine learning, pages 41–48. ACM, 2009. 65
- [Borji16] Ali Borji, Saeed Izadi & Laurent Itti. *iLab-20M : A Large-Scale Controlled Object Dataset to Investigate Deep Learning*. In Proceedings of the IEEE International Conference on Computer Vision, pages 2221–2230, 2016. 53
- [Bouzy04] Bruno Bouzy & Bernard Helmstetter. *Monte Carlo go developments*. Advances in computer games, vol. 10, pages 159–174, 2004. 40
- [Bromley94] Jane Bromley, Isabelle Guyon, Yann LeCun, Eduard Säckinger & Roopak Shah. *Signature verification using a "siamese" time delay neural network*. In Advances in Neural Information Processing Systems, pages 737–744, 1994. 53
- [Brügmann93] B. Brügmann. *Monte carlo go*. White paper, Citeseer, 1993. 40
- [Cho19] Sungjae Cho, Jaeseo Lim, Chris Hickey & Byoung-Tak Zhang. *Problem Difficulty in Arithmetic Cognition : Humans and Connectionist Models*. PsyArXiv, 2019. 63
- [Collier18] Mark Collier & Joeran Beel. *Implementing Neural Turing Machines*. In Artificial Neural Networks and Machine Learning – ICANN 2018, pages 94–104, Cham, 2018. Springer International Publishing. 63
- [Cybenko89] George Cybenko. *Approximation by superpositions of a sigmoidal function*. Mathematics of control, signals and systems, vol. 2, no. 4, pages 303–314, Springer, 1989. 63
- [Dai21a] Ruiqi Dai, Mathieu Lefort, Frédéric Armetta, Mathieu Guillermin & Stefan Duffner. *Novelty detection for unsupervised continual learning in image sequences*. In IEEE International Conference on Tools with Artificial Intelligence (ICTAI), Washington DC (virtual), United States, November 2021. 52, 59, 60
- [Dai21b] Ruiqi Dai, Mathieu Lefort, Frédéric Armetta, Mathieu Guillermin & Stefan Duffner. *Self-supervised continual learning for object recognition in image sequences*. In International Conference on Neural Information Processing (ICONIP), Bali, Indonesia, December 2021. 52, 59, 62

- [Devillers22] Alexandre Devillers, Valentin Chaffraix, Frédéric Armetta, Stefan Duffner & Mathieu Lefort. *The Impact of Action in Visual Representation Learning*. In International Conference on Development and Learning (ICDL 2022), London, United Kingdom, September 2022. 52, 58
- [Dorigo92] Marco Dorigo. *Optimization, learning and natural algorithms*. Ph. D. Thesis, Politecnico di Milano, 1992. 22
- [Drescher91a] G. L. Drescher. *Made-Up Minds : a constructivist approach to artificial intelligence*. MIT Press, 1991. 34
- [Drescher91b] G.L. Drescher. *Made-up minds : a constructivist approach to artificial intelligence*. The MIT Press, 1991. 18
- [Enzenberger10] Markus Enzenberger, Martin Müller, Broderick Arneson & Richard Segal. *Fuego—an open-source framework for board games and Go engine based on Monte Carlo tree search*. IEEE Transactions on Computational Intelligence and AI in Games, vol. 2, no. 4, pages 259–270, IEEE, 2010. 43
- [Ess08] A. Ess, B. Leibe, K. Schindler & L. V. Gool. *A Mobile Vision System for Robust Multi-Person Tracking*. In Proceedings of the IEEE International Conference on Computer Vision, 2008. 53
- [Fabbri12] André Fabbri, Frédéric Armetta, Eric Duchene & Salima Hassas. *A new self-acquired knowledge process for Monte Carlo Tree Search*. In European Conference on Artificial Intelligence, Computer Game Workshop, Montpellier, France, August 2012. 41
- [Fabbri14] André Fabbri, Frédéric Armetta, Eric Duchêne & Salima Hassas. *Knowledge complement for Monte Carlo Tree Search : an application to combinatorial games*. In 2014 IEEE 26th International Conference on Tools with Artificial Intelligence, pages 997–1003, Limassol, Cyprus, November 2014. 41
- [Fabbri15] André Fabbri. *Dynamique d'apprentissage pour Monte Carlo Tree Search : applications aux jeux de Go et du Clobber solitaire impartial*. Theses, Université Claude Bernard - Lyon I, October 2015. 43, 44
- [Fabbri16] André Fabbri, Frédéric Armetta, Eric Duchene & Salima Hassas. *A Self-Acquiring Knowledge Process for MCTS*. International Journal on Artificial Intelligence Tools, vol. 25, no. 01, page 20, World Scientific Publishing, February 2016. Les ressources du centre de calcul de l'IN2P3 ont

- été nécessaires pour obtenir certains des résultats présentés dans ce papier. 43
- [Faugier24] Elliot Faugier, Frédéric Armetta, Angela Bonifati & Bruno Yun. *Assisted Debate Builder with Large Language Models*. In European Conference on Artificial Intelligence (ECAI), demo session, 2024. <https://www.youtube.com/watch?v=KMzqKJlH91E>. 15
- [Friston11] Karl Friston, Jérémie Mattout & James Kilner. *Action understanding and active inference*. *Biological cybernetics*, vol. 104, no. 1, pages 137–160, Springer, 2011. 55
- [Friston12] Karl Friston, Rick Adams, Laurent Perrinet & Michael Breakspear. *Perceptions as hypotheses : saccades as experiments*. *Frontiers in psychology*, vol. 3, page 151, Frontiers, 2012. 55
- [Froese09] Tom Froese & Tom Ziemke. *Enactive artificial intelligence : Investigating the systemic organization of life and mind*. *Artificial Intelligence*, vol. 173, no. 3-4, pages 466–500, Elsevier, 2009. 55
- [Furao06] Shen Furao & Osamu Hasegawa. *An incremental network for on-line unsupervised classification and topology learning*. *Neural networks*, vol. 19, no. 1, pages 90–106, Elsevier, 2006. 59
- [Gambardella99] Luca Maria Gambardella, Éric Taillard & Giovanni Agazzi. *MACS-VRPTW : A multiple ant colony system for vehicle routing problems with time windows*, 1999. 22
- [Geng09] Xian-Min Geng, Guang-Hui Wen, Shu-Chen Wan, Jie-Yu Xiong & Ying Wang. *The emergence of scale-free networks with a seceding mechanism*. *Physica A : Statistical Mechanics and its Applications*, vol. 388, no. 20, pages 4484–4490, Elsevier, 2009. 25
- [Georgeon12] Olivier Georgeon & Frank Ritter. *An Intrinsically-Motivated Schema Mechanism to Model and Simulate Emergent Cognition*. *Cognitive Systems Research*, vol. 15-16, pages 73–92, Elsevier, June 2012. 16
- [Gibson66] James Jerome Gibson & Leonard Carmichael. *The senses considered as perceptual systems*, volume 2. Houghton Mifflin Boston, 1966. 55
- [Goodfellow16] Ian Goodfellow, Yoshua Bengio, Aaron Courville & Yoshua Bengio. *Deep learning*, volume 1. MIT press Cambridge, 2016. 63

- [Grassé59] Pierre-P Grassé. *La reconstruction du nid et les coordinations interindividuelles chez *Bellicositermes natalensis* et *Cubitermes* sp. la théorie de la stigmergie : Essai d'interprétation du comportement des termites constructeurs*. Insectes sociaux, vol. 6, no. 1, pages 41–80, Springer, 1959. 11, 22
- [Graves14a] Alex Graves, Greg Wayne & Ivo Danihelka. *Neural Turing Machines*, 2014. cite arxiv :1410.5401. 12
- [Graves14b] Alex Graves, Greg Wayne & Ivo Danihelka. *Neural Turing Machines*. CoRR, vol. abs/1410.5401, 2014. 62, 63
- [Guclu08] Hasan Guclu & Murat Yuksel. *Limited scale-free overlay topologies for unstructured peer-to-peer networks*. IEEE Transactions on Parallel and Distributed Systems, vol. 20, no. 5, pages 667–679, IEEE, 2008. 23, 25
- [Guériau15a] Maxime Guériau, Romain Billot, Frédéric Armetta, Salima Hassas & Nour-Eddin El Faouzi. *Un simulateur multiagent de trafic coopératif*. In Vercouter, Laurent, Picard & Gauthier, editeurs, 23es Journées Francophones sur les Systèmes Multi-Agents (JFSMA'15), pages 165–174, Rennes, France, June 2015. Cepaduès. 31
- [Guériau15b] Maxime Guériau, Romain Billot, Nour-Eddin El Faouzi, Salima Hassas & Frédéric Armetta. *Multi-Agent Dynamic Coupling for Cooperative Vehicles Modeling*. In The Twenty-Ninth Conference on Artificial Intelligence AAAI'2015 - (DEMO Track), Proceedings of the Twenty-Ninth AAAI Conference on Artificial Intelligence, pages 4276–4277, Austin, United States, January 2015. 31
- [Guériau17] Maxime Guériau, Frédéric Armetta, Salima Hassas, Romain Billot & Nour Eddin El Faouzi. *Contrôle par apprentissage constructiviste pour le trafic coopératif*. In Journées Francophones sur les Systèmes Multi-Agents (JFSMA 2017), Caen, France, July 2017. 50
- [Hadsell06] Raia Hadsell, Sumit Chopra & Yann LeCun. *Dimensionality reduction by learning an invariant mapping*. In Computer vision and pattern recognition, 2006 IEEE computer society conference on, volume 2, pages 1735–1742. IEEE, 2006. 53
- [Held63] Richard Held & Alan Hein. *Movement-produced stimulation in the development of visually guided behavior*. Journal of comparative and physiological psychology, vol. 56, no. 5, American Psychological Association, 1963. 55

- [Hochreiter97] Sepp Hochreiter & Jürgen Schmidhuber. *Long short-term memory*. *Neural computation*, vol. 9, no. 8, pages 1735–1780, MIT Press, 1997. 56
- [Hoshen16] Yedid Hoshen & Shmuel Peleg. *Visual Learning of Arithmetic Operations*. In *Proceedings of the Thirtieth AAAI Conference on Artificial Intelligence, AAAI'16*, pages 3733–3739. AAAI Press, 2016. 64, 65, 66
- [Hotelling31] Harold Hotelling *et al.* *The generalization of Student's ratio*. Springer, 1931. 62
- [Jafari14] O. H. Jafari, D. Mitzel & B. Leibe. *Real-Time RGB-D based People Detection and Tracking for Mobile Robots and Head-Worn Cameras*. In *IEEE International Conference on Robotics and Automation (ICRA)*, 2014. 53
- [Kaiser15] Łukasz Kaiser & Ilya Sutskever. *Neural gpus learn algorithms*. arXiv preprint arXiv :1511.08228, 2015. 63
- [Kesting07] Arne Kesting, Martin Treiber & Dirk Helbing. *General lane-changing model MOBIL for car-following models*. *Transportation Research Record : Journal of the Transportation Research Board*, vol. 1999, no. 1, pages 86–94, 2007. 47
- [Kingma13] Diederik P Kingma & Max Welling. *Auto-encoding variational bayes*. arXiv preprint arXiv :1312.6114, 2013. 56
- [Krizhevsky12] Alex Krizhevsky, Ilya Sutskever & Geoffrey E Hinton. *Imagenet classification with deep convolutional neural networks*. In *Advances in neural information processing systems*, pages 1097–1105, 2012. 63
- [Kuipers06] B.J. Kuipers, P. Beeson, J. Modayil & J. Provost. *Bootstrap learning of foundational representations*. *Connection Science*, vol. 18, no. 2, pages 145–158, 2006. 34
- [Leal-Taixé16] Laura Leal-Taixé, Cristian Canton-Ferrer & Konrad Schindler. *Learning by tracking : Siamese CNN for robust target association*. In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pages 33–40, 2016. 53
- [Leibe07] B. Leibe, K. Schindler, N. Cornelis & L. V. Gool. *Coupled Object Detection and Tracking from Static Cameras and Moving Vehicles*. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2007. 53
- [Liang15] Xiaodan Liang, Si Liu, Yunchao Wei, Luoqi Liu, Liang Lin & Shuicheng Yan. *Towards computational baby learning : A*

- weakly-supervised approach for object detection*. In Proceedings of the IEEE International Conference on Computer Vision, pages 999–1007, 2015. 53
- [Mazac14] Sébastien Mazac, Frédéric Armetta & Salima Hassas. *Bootstrapping sensori-motor patterns for a constructivist learning system in continuous environments*. In M. I. T. Press, editeur, ALIFE 14 : the 14th International Conference on the Synthesis and Simulation of Living Systems, pages 160–167, July 2014. 37
- [Mazac15a] Sébastien Mazac. *Approche décentralisée de l'apprentissage constructiviste et modélisation multi-agent du problème d'amorçage de l'apprentissage sensorimoteur en environnement continu : application à l'intelligence ambiante*. Theses, Université Claude Bernard - Lyon I, October 2015. 32, 37
- [Mazac15b] Sébastien Mazac, Frédéric Armetta & Salima Hassas. *Approche décentralisée pour un apprentissage constructiviste en environnement continu : application à l'intelligence ambiante*. In Journées Francophones sur les Systèmes Multi-Agents (JFSMA), Rennes, France, June 2015. Cépaduès. 37
- [Medjkoune17] Nawel Medjkoune, Frédéric Armetta, Mathieu Lefort & Stefan Duffner. *Autonomous object recognition in videos using Siamese Neural Networks*. In EUCognition Meeting (European Society for Cognitive Systems) on "Learning : Beyond Deep Neural Networks", page 4, Zurich, Switzerland, November 2017. 52, 55
- [Michlmayr06] Elke Michlmayr. *Self-organization for search in peer-to-peer networks : the exploitation-exploration dilemma*. In Proceedings of the 1st international conference on Bio inspired models of network, information and computing systems, pages 29–es, 2006. 23
- [Mossio08] Matteo Mossio & Dario Taraborelli. *Action-dependent perceptual invariants : From ecological to sensorimotor approaches*. Consciousness and cognition, vol. 17, no. 4, Elsevier, 2008. 55
- [Nguyen Vu12a] Quang-Anh Nguyen Vu. *Consistency and robustness in multi-agent systems disturbed : application to a decentralized system for collecting distributed information*. Theses, Université Claude Bernard - Lyon I, Dec 2012. 29
- [Nguyen Vu12b] Quang Anh Nguyen Vu, Richard Canal, Benoit Gaudou, Salima Hassas & Frédéric Armetta. *TrustSets - Using trust*

- to detect deceitful agents in a distributed information collecting system.* Journal of Ambient Intelligence and Humanized Computing, vol. 3, no. 4, pages 251–263, Springer, December 2012. 29
- [Nollet20] Bastien Nollet, Mathieu Lefort & Frédéric Armetta. *Learning Arithmetic Operations With A Multistep Deep Learning.* In The International Joint Conference on Neural Networks (IJCNN), Glasgow, United Kingdom, July 2020. 63, 64, 65
- [Nothdurft05] Hans-Christoph Nothdurft. *Saliency of Feature Contrast.* In Laurent Itti, Geraint Rees & John K. Tsotsos, editeurs, Neurobiology of Attention, pages 233 – 239. Academic Press, Burlington, 2005. 54
- [O’Regan01a] J Kevin O’Regan & Alva Noë. *A sensorimotor account of vision and visual consciousness.* Behavioral and brain sciences, vol. 24, no. 5, page 939, Cambridge University Press, 2001. 55
- [O’Regan01b] J. Kevin O’Regan & Alva Noë. *A sensorimotor account of vision and visual consciousness.* Behavioral and Brain Sciences, vol. 24, no. 5, page 939–973, Cambridge University Press, 2001. 14
- [Oudeyer07] Pierre-Yves Oudeyer & Frederic Kaplan. *What is Intrinsic Motivation? A Typology of Computational Approaches.* Frontiers in Neurorobotics, vol. 1, page 6, 2007. 8
- [Oudeyer16] P-Y Oudeyer, Jacqueline Gottlieb & Manuel Lopes. *Intrinsic motivation, curiosity, and learning : Theory and applications in educational technologies.* Progress in brain research, vol. 229, pages 257–284, Elsevier, 2016. 16
- [Page54] Ewan S Page. *Continuous inspection schemes.* Biometrika, vol. 41, no. 1/2, pages 100–115, JSTOR, 1954. 60
- [Piaget54] Jean Piaget. *The Construction of Reality in the Child.* Basic Books, 1954. 11, 13
- [Piaget55] Jean Piaget. *The construction of reality in the child.* Journal of Consulting Psychology, vol. 19, no. 1, page 77, Elsevier, 1955. 3
- [Rao95] Anand S Rao, Michael P Georgeff et al. *BDI agents : from theory to practice.* In Icmas, volume 95, pages 312–319, 1995. 10
- [Rao19] Dushyant Rao, Francesco Visin, Andrei Rusu, Razvan Pascanu, Yee Whye Teh & Raia Hadsell. *Continual Unsupervised Representation Learning.* In Advances in Neural

- Information Processing Systems, pages 7645–7655, 2019. xii, 59, 60
- [Rebuffi17] Sylvestre-Alvise Rebuffi, Alexander Kolesnikov, Georg Sperl & Christoph H Lampert. *icarl : Incremental classifier and representation learning*. In CVPR, pages 2001–2010, 2017. 59
- [Schillo99] M. Schillo, P. Funk & M. Rovatsos. *Who can you trust : Dealing with deception*. In Proceedings of Autonomous Agents '99 Workshop on "Deception, Fraud, and Trust in Agent Societies", pages 81–94, Seattle, USA, May 1999. 29
- [Searle80] John R. Searle. *Minds, brains, and programs*. Behavioral and Brain Sciences, vol. 3, no. 03, pages 417–424, 1980. 10
- [Settles09] Burr Settles. *Active learning literature survey*. University of Wisconsin-Madison Department of Computer Sciences, 2009. 16, 64
- [Siegelmann95] H.T. Siegelmann & E.D. Sontag. *On the Computational Power of Neural Nets*. Journal of Computer and System Sciences, vol. 50, no. 1, pages 132 – 150, 1995. 63
- [Sutton98] Richard S. Sutton & Andrew G. Barto. Reinforcement learning : An introduction. MIT Press, Cambridge, MA, 1998. 13
- [Sutton18] Richard S Sutton & Andrew G Barto. Reinforcement learning : An introduction. MIT press, 2018. 18
- [Taillandier10] Patrick Taillandier, Duc-An Vo, Edouard Amouroux & Alexis Drogoul. *GAMA : bringing GIS and multi-level capabilities to multi-agent simulation*. In European Workshop on Multi-Agent Systems, 2010. 31
- [Tian19] Yuandong Tian, Jerry Ma, Qucheng Gong, Shubho Sengupta, Zhuoyuan Chen, James Pinkerton & Larry Zitnick. *Elf opengo : An analysis and open reimplement of alpha-zero*. In International conference on machine learning, pages 6244–6253. PMLR, 2019. 63
- [Treiber13] Martin Treiber & Arne Kesting. Traffic flow dynamics : data, models and simulation. Springer, 2013. 47
- [Treisman80] Anne M Treisman & Garry Gelade. *A feature-integration theory of attention*. Cognitive psychology, vol. 12, no. 1, pages 97–136, Elsevier, 1980. 54
- [Tseng09] Jie-Jun Tseng, Sai-Ping Li, Shu-Heng Chen & Sun-Chong Wang. *Emergence of scale-free networks in markets*. Advances

- in *Complex Systems*, vol. 12, no. 01, pages 87–97, World Scientific, 2009. 25
- [Turing50] A. M. Turing. *Computing Machinery and Intelligence*, 1950. One of the most influential papers in the history of the cognitive sciences : <http://cogsci.umn.edu/millennium/final.html>. 1
- [Vaswani17a] A Vaswani. *Attention is all you need*. *Advances in Neural Information Processing Systems*, 2017. 12
- [Vaswani17b] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N Gomez, Łukasz Kaiser & Illia Polosukhin. *Attention is All you Need*. In *Advances in Neural Information Processing Systems 30*, pages 5998–6008. Curran Associates, Inc., 2017. 63
- [Von Glasersfeld84] E. Von Glasersfeld. *An introduction to radical constructivism*. *The invented reality*, pages 17–40, 1984. 8
- [Wang15] Xiaolong Wang & Abhinav Gupta. *Unsupervised learning of visual representations using videos*. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 2794–2802, 2015. 53
- [Wu12] Jieyu Wu, Xinyu Shao, Jinhang Li & Gang Huang. *Scale-free properties of information flux networks in genetic algorithms*. *Physica A : Statistical Mechanics and its Applications*, vol. 391, no. 4, pages 1692–1701, Elsevier, 2012. 25
- [Zhang17] Yu Zhang & Qiang Yang. *A survey on multi-task learning*. arXiv preprint arXiv :1707.08114, 2017. 65
- [Zhao06] Shanyu Zhao, Daniel Stutzbach & Reza Rejaie. *Characterizing files in the modern gnutella network : A measurement study*. In *Multimedia Computing and Networking 2006*, volume 6071, pages 239–251. SPIE, 2006. 23

Curriculum Vitae des activités de recherche et d'enseignement

1 Synthèse du parcours professionnel et contexte d'exercice

2001 .. 2002	Ingénieur informatique - Société HD Technology
2002 .. 2003	DEA
Oct. 2003.. Déc. 2006	Doctorat - Laboratoire GAMA puis LIRIS (Lyon1)
Fev. 2007 .. Août 2008	ATER - Laboratoire LIESP (Lyon1)
Sept. 2008 .. en cours	Maître de conférences - Laboratoire LIRIS (Lyon1)
Sept. 2015 .. Jan. 2016	1/2 CRCT pour élargissement thématique de recherche
Jan. 2023 .. en cours	Responsable de l'équipe SyCoSMA du LIRIS

Je débute ma carrière en tant d'informaticien, mais très intéressé par la recherche et l'enseignement je reprends mes études et réalise un doctorat dans le domaine des systèmes multi-agents, et j'obtiens par la suite un poste de maître de conférences. Le développement d'IA autonomes a alors toujours fait partie de mes centres d'intérêt scientifiques. En 2015, mon projet d'élargissement thématique de CRCT est soutenu par l'UCBL ce qui me permet d'approfondir mes connaissances dans le domaine des réseaux de neurones informatiques, et développer par la suite de nouveaux projets et de nouvelles applications et collaborations industrielles. J'ai pu m'investir au cours de ces années dans l'enseignement, et accompagner les étudiants dans leurs apprentissages. J'ai également pu participer aux tâches collectives et proposer mes services. Je suis désormais responsable de l'équipe SyCoSMA du laboratoire LIRIS à Lyon depuis le début de l'année 2023.

2 Activité scientifique

mots-clefs : *intelligence artificielle, systèmes autonomes, intelligence collective, systèmes multi-agents, modélisations cognitives, constructivisme, apprentissage, modèles de langage*

2.1 Présentation synthétique des thématiques de recherche

Je m'intéresse à la conception de systèmes intelligents et autonomes, c'est-à-dire des systèmes capables de développer leur cognition sans superviseur ou avec une supervision limitée. Cette thématique de recherche fait partie du domaine de l'IA, et met l'accent sur le fait que le système est incarné dans son environnement et doit développer ses capacités cognitives en autonomie afin d'atteindre ses objectifs.

Au cours de mes premières années de recherche, dans la continuité de mes travaux de thèse sur la proposition d'une approche par SMA du partage de ressources critiques, je me suis intéressé à ce qu'apportent les Systèmes Multi-Agents pour la conception de systèmes autonomes, à leurs capacités d'auto-organisation et à l'intelligence collective qu'ils peuvent manifester. Ces travaux mettent naturellement l'accent sur les mécanismes de contrôle et de coordination décentralisés.

Par la suite, je me suis intéressé aux possibilités offertes dans le cadre d'une modélisation centralisée de systèmes cognitifs, et j'ai pu bénéficier pour mon projet d'ouverture thématique d'une demi CRCT en 2015. Je me suis intéressé à des domaines de recherches d'inspirations divers, tels que les diverses méthodes du domaine de l'apprentissage informatique et particulièrement des réseaux de neurones, mais ai également développé une inspiration découlant des neurosciences, des sciences cognitives ou même de la philosophie de la connaissance qui peut compléter avantageusement notre compréhension du rôle de la connaissance. Lors de mon parcours, je me suis intéressé à différents travaux interdisciplinaires.

Mes recherches sont appliquées à un large éventail de domaines (intelligence ambiante, IA pour les jeux combinatoires, contrôle du trafic routier, large modèles de langue, modèles prédictifs, etc.), pour lesquels j'ai pu proposer et participer à l'élaboration de différents modèles d'apprentissage. Au travers de mon activité de recherche, je développe des partenariats pour transférer des technologies vers des applications industrielles.

2.2 Liste des encadrements scientifiques

Lorsque les encadrements font partie des projets détaillés dans le mémoire (chapitre 3), ceux-ci sont notifiés en gras..

Synthèse :

- Thèse en cours : 1 direction de thèse
- Thèses soutenues : 6
- Post-doc et ingénieur : 2 post-doc (12M et 26M), 1 ingénieur recherche (13M, en cours), 1 ingénieur recherche (3M), 1 post-doc (12M)
- Plus de 20 stages de master (liste complète en annexe)

Thèse en cours :

- Maya Medjad (janv. 2023 ..), « Apprentissage pour la reconnaissance d'intention et la gestion de conversations », CIFRE société Reecall (<https://reecall.com>), directeur de thèse, co-encadrement 60%, Bruno Yun 40% depuis septembre 2023, [34]

Thèses soutenues :

- Ruiqi Dai (2018 ..2022), « Continual class-incremental learning for autonomous object recognition in image sequences », Bourse MESR, soutenue le 14 Septembre 2022, co-encadrement 50%, Stefan Duffner (directeur) 25%, Mathieu Guillermin 25%, [5][6], [3.6.3 Détection d'objets en apprentissage continu]
- Maxime Guériau (2013..2016), «Systèmes multi-agents, auto-organisation et contrôle par apprentissage constructiviste pour la modélisation et la régulation dans les systèmes coopératifs de trafic », financement ADR ARC 7, soutenue le 12 December 2016, co-encadrement à 25%, Salima Hassas 25% (cotutelle 50% UCBL, 50% LICIT/IFSTTAR Romain Billot, Nour-Eddin El Faouzi) [13][14][17][35][37][44], [3.2.1 Application à la régulation d'un ensemble de véhicules connectés], [3.5 Apprentissage et régulation du trafic routier]
- Sébastien Mazac (2012 ..2015), « Approche décentralisée de l'apprentissage constructiviste et modélisation multi-agent du problème d'amorçage de l'apprentissage sensorimoteur en environnement continu : application à l'intelligence ambiante », CIFRE société OSLO, soutenue le 6 Octobre 2015, co-encadrement 50%, Salima Hassas 50%, [16][36], [3.3 Apprentissage de régularités en environnement continu et Intelligence Ambiante]
- André Fabbri (2011 ..2015), « Dynamique d'apprentissage pour Monte Carlo Tree Search : applications aux jeux de Go et du Clobber solitaire impartial », soutenue le 22 October 2015, Bourse MESR, encadrant principal 45%, Salima Hassas 30%, Éric Duchêne 25%, [15][21][40], 3.4 Extensions de MCTS pour les jeux combinatoires
- Q. A. NGuyen Vu (2008 ..2012) : « Resilience in complex systems, application to Multi-Robot Systems for danger mapping », Thèse IFI, soutenue le 5/12/2012, co-encadrement à 25%, Salima Hassas 25% (50% UCBL, 50% IFI Richard Canal) [22][24][28][29][30][38][43][45], [3.2.1 Application à la cartographie collaborative]
- Olivier Lefevre (2007..2010) : "Contrôle Endogène des Systèmes Multi-Agents pour la résolution de Problèmes Complexes", soutenue le 5 Octobre 2010, co-encadrement à 50%, Salima Hassas 50%, [32]

Encadrements d'ingénieurs et post-docs :

- Projet PepperMint (Sept. 2021 - Déc. 2024) [2][39][49] :
 - Applications robotiques, Julien Perier-Camby [Ingénieur Recherche 13M], co-encadrement à 90%, co-encadrants : Heike Baldaufquilliatre
 - Agent conversationnel, Brooke Stephenson [Post-Doc 12M] co-encadrement à 80%, co-encadrants : Heike Baldaufquilliatre, Mathieu Lefort
 - Analyse conversationnelle, co-encadrement à 10% de Lucien Tisserant [Post-Doc 26M], co-encadrants : Heike Baldaufquilliatre, Salima Hassas, Mathieu Lefort
- Projet Génération de recettes culinaires (Sept. 2019 - Avr. 2021) : responsable scientifique, co-encadrement à 60%, Bruno Malfoy (Pulsalys) 40% : Lucie Martinet (Post-Doc, 12M), Pierre-Elliott Thiboud (Ingénieur de recherche, 3M)

Encadrements de stages, présentés dans le mémoire

(liste complète en annexe)

- Jan. 2017..2020, Reconnaissance et suivi d'objets dans une séquence vidéo par des paradigmes d'apprentissage constructivistes, porteur du projet transverse accepté et financé par le LIRIS, stage Valentin Chaffraix dans la continuité financé en accompagnement de la thèse de Ruiqi Dai
 - stage de M2 Recherche Nawel Medjkoune, co-encadrement 33%, co-encadrants Mathieu Lefort et Stefan Duffner [12], **[3.6.1 Exploiter les régularités spacio-temporelles d'un flux vidéo]**
 - stage de M2 Recherche Louis Manhès, co-encadrement 33%, co-encadrants Mathieu Lefort et Stefan Duffner
 - Janvier 2020, stage M2 IA Valentin Chaffraix, étude d'un modèle d'attention visuelle de deep learning dans un contexte sensorimoteur encadrement 33%, co-encadrants Mathieu Lefort, Stefan Duffner [4], **[3.6.2 Rôle de l'action pour l'apprentissage de représentations visuelles]**
- Sept. 2019 .., IA et apprentissage d'algorithmes, projet financé par mes projets collaboratifs industriels, **[3.7 IA et apprentissage d'algorithmes]**
 - Sept. 2024, stage M1 Polytech, Farah Ben Slama, encadrement 100%
 - Mai 2021, stage 5ème année ENS de Pierre Marec encadrement 100%
 - Janvier 2020, stage M2 IA d'Anthony Baccuet, encadrement 100% [3][33]
 - Septembre 2019. stage 4A Bastien Nollet, encadrement 100% [7]

2.3 Publications et productions scientifiques

Les noms des étudiants encadrés sont soulignés.

51 Publications depuis la dernière promotion :

- Conférences internationales à comité de lecture : 32 (1 A*, 5 A, 6 B)
- Conférences nationales à comité de lecture : 6
- Articles dans revues internationales à comité de lecture : 5 (1 Q1, 3 Q2, 1 Q3)
- Article dans revue nationale à comité de lecture : 1
- Chapitres d'ouvrages : 3
- logiciels : 2
- Brevets : 2

Conférences internationales à comité de lecture

- 1 Elliot Faugier, **Frédéric Armetta**, Angela Bonifati & Bruno Yun, Assisted Debate Builder with Large Language Models, European Conference on Artificial Intelligence (ECAI), demo paper, 2024, **CORE A**
- 2 Lucien Tisserand, **Frédéric Armetta**, Heike Baldauf-Quilliatre, Antoine Bouquin, Salima Hassas & Mathieu Lefort (2023). « Sequential annotations for naturally-occurring HRI : first insights ». Workshop Human-Robot Conversational Interaction (HRCI) in ACM/IEEE International Conference on Human-Robot Interaction, Stockholm (Suède). ArXiv : 2308.15097, 2023
- 3 **Frédéric Armetta**, Anthony Baccuet & Mathieu Lefort (2023). « Algorithmic learning, a next step for AI. An application to arithmetic operations ». 9th International Workshop on Artificial Intelligence and Cognition (AIC), Bremen (Allemagne), 2023
- 4 Alexandre Devillers, Valentin Chaffraix, **Frédéric Armetta**, Stefan Duffner & Mathieu Lefort. « The Impact of Action in Visual Representation Learning ». International Conference on Development and Learning (ICDL), London (Royaume-Uni). doi : 10.1109/ICDL53763.2022.9962210, 2022
- 5 Ruiqi Dai, Mathieu Lefort, **Frédéric Armetta**, Mathieu Guillermin & Stefan Duffner. « Novelty detection for unsupervised continual learning in image sequences ». IEEE International Conference on Tools with Artificial Intelligence (ICTAI), Washington DC (virtual) (États-Unis). doi : 10.1109/ICTAI52525.2021.00080, 2021, **CORE B**
- 6 Ruiqi Dai, Mathieu Lefort, **Frédéric Armetta**, Mathieu Guillermin & Stefan Duffner. « Self-supervised continual learning for object recognition in image sequences ». International Conference on Neural Information Processing (ICONIP), Bali (Indonésie), 2021, **CORE B**
- 7 Bastien Nollet, Mathieu Lefort & **Frédéric Armetta**. « Learning Arithmetic Operations With A Multistep Deep Learning ». The International Joint Conference on Neural Networks (IJCNN), Glasgow (Royaume-Uni). doi : 10.1109/IJCNN48605.2020.9206963, 2020, **CORE 2020 A**

- 8 Marc Bertin, Pierre Jonin, **Frédéric Armetta** & Iana Atanassova. « Identifying the Conceptual Space of Citation Contexts using Coreferences ». 4th Joint Workshop on Bibliometric-enhanced Information Retrieval and Natural Language Processing for Digital Libraries (BIRNDL) at the 42nd International ACM SIGIR Conference on Research and Development in Information Retrieval, Paris (France), pp. 138-144, 2019
- 9 Julien Perier-Camby, Marc Bertin, Iana Atanassova & **Frédéric Armetta** (2019). « A preliminary study to compare deep learning with rule-based approaches for citation classification ». 8th International Workshop on Bibliometric-enhanced Information Retrieval (BIR) co-located with the 41st European Conference on Information Retrieval (ECIR), Cologne (Allemagne), pp. 125-131, 2019
- 10 Marc Bertin, Pierre Jonin, **Frédéric Armetta** & Iana Atanassova. « Determining Citation Blocks using End-to-end Neural Coreference Resolution Model for Citation Context Analysis ». 17th International Conference on Scientometrics & Informetrics, Rome (Italie), p. 2720, 2019
- 11 Feras Al Kassar & **Frédéric Armetta**. « Extracting tags from large raw texts using End-to-End memory networks ». 2nd Workshop on Semantic Deep Learning (SemDeep-2) in the 12th International Conference on Computational Semantics (IWCS), Montpellier (France), pp. 33-47, 2017
- 12 Nawel Medjkoune, **Frédéric Armetta**, Mathieu Lefort & Stefan Duffner. « Autonomous object recognition in videos using Siamese Neural Networks ». EUCognition Meeting (European Society for Cognitive Systems) on "Learning : Beyond Deep Neural Networks", Zurich (Suisse), p. 4, 2017
- 13 Maxime Guériau, **Frédéric Armetta**, Salima Hassas, Romain Billot & Nour-Eddin El Faouzi. « A constructivist approach for a self-adaptive decision-making system : application to road traffic control ». 28th IEEE International Conference on Tools with Artificial Intelligence (ICTAI), San Jose (États-Unis), pp. 670-677. doi : 10.1109/ICTAI.2016.0107, 2016, **CORE B**
- 14 Maxime Guériau, Romain Billot, Nour-Eddin El Faouzi, Salima Hassas & **Frédéric Armetta** (2015). « Multi-Agent Dynamic Coupling for Cooperative Vehicles Modeling ». The Twenty-Ninth Conference on Artificial Intelligence AAI/2015 - (DEMO Track), Austin (États-Unis), pp. 4276-4277, 2015, **CORE A***
- 15 André Fabbri, **Frédéric Armetta**, Eric Duchêne & Salima Hassas. « Knowledge complement for Monte Carlo Tree Search : an application to combinatorial games ». 2014 IEEE 26th International Conference on Tools with Artificial Intelligence, Limassol (Chypre), pp. 997-1003, 2014, **CORE B**
- 16 Sébastien Mazac, **Frédéric Armetta** & Salima Hassas. « On bootstrapping sensori-motor patterns for a constructivist learning system in continuous environments ». Alife 14 : Fourteenth International Conference on the Synthesis and Simulation of Living Systems, New York (États-Unis). doi : 10.7551/978-0-262-32621-6-ch028, 2014, **CORE A**

- 17 Maxime Guériau, Romain Billot, Salima Hassas, **Frédéric Armetta** & Nour-Eddin El Faouzi (2014). « An extension of MovSim for Multi-Agent Cooperative Vehicles Modeling ». IEEE International Conference on Connected Vehicles & Expo, Vienna (Autriche), pp. 859-860. doi : 10.1109/ICCVE.2014.7297673, 2014
- 18 Julien Monteil, Romain Billot, **Frédéric Armetta**, Salima Hassas & Nour-Eddin El Faouzi. « Cooperative highway traffic : multi-agent modeling and robustness assessment to local perturbations ». the 92nd Annual Meeting of the Transportation Research Board, reviewed by TRB's Traffic Flow Theory and Characteristics Committee (AHB45), Washington, D.C. (États-Unis), pp. 1-19, 2013, **CORE A**
- 19 Gaël Clair, **Frédéric Armetta** & Salima Hassas. « EC4MAS : A multiagent model with endogenous control for combinatorial optimization problem solving ». Conference on Adaptive and Self-Adaptive Systems and Applications, Nice (France), pp. 38-43, 2012
- 20 **Frédéric Armetta**, Mohammed Haddad, Salima Hassas & Hamamache Kheddouci. « A self-organized system improving inner topology for data sharing efficiency ». IEEE Conference on Evolving and Adaptive Intelligent Systems (EAIS), Madrid (Espagne), pp. 145-150, . doi : 10.1109/EAIS.2012.6232820, 2012
- 21 André Fabbri, **Frédéric Armetta**, Eric Duchene & Salima Hassas. « A new self-acquired knowledge process for Monte Carlo Tree Search ». European Conference on Artificial Intelligence, Montpellier (France), 2012, **CORE A**
- 22 Quang Anh Nguyen Vu, Salima Hassas, **Frédéric Armetta**, Benoit Gaudou & Richard Canal. « Towards self-organizing approach for maintaining coherence and robustness in disturbed MAS », the Fifth IEEE International Conference on Self-Adaptive and Self-Organizing Systems (SASO), 2011
- 23 Gaël Clair, **Frédéric Armetta** & Salima Hassas. « Self-adaptive tuning of dynamic changing problem solving : a first step to endogenous control in Multi-Agents Based Problem Solvers » , The Seventh International Conference on Autonomous and Autonomous Systems (ICAS), 2011
- 24 Quang Anh Nguyen Vu, Benoit Gaudou, Richard Canal, Salima Hassas & **Frédéric Armetta**. « TrustSets - Using trust to detect deceitful agents in a distributed information collecting system », Best Paper Award, IEEE-RIVF International Conference on Computing and Telecommunication Technologies, 2010
- 25 Gaël Clair, Leonardo Lana de Carvalho, **Frédéric Armetta** & Salima Hassas. « Social representations for endogenous control ». poster track, in the proceedings of European Conference on Complex Systems (ECCS), 2011
- 26 **Frédéric Armetta**, Mohammed Haddad, Salima Hassas & Hamamache Kheddouci. « Self-organized Routing for Unstructured Peer-to-peer Networks ». poster track, the Fourth IEEE International Conference on Self-Adaptive and Self-Organizing Systems (SASO), 2010
- 27 Mohammed Haddad, **Frédéric Armetta**, Hamamache Kheddouci, Salima Hassas, « Self-organized Query Routing for Peer-to-peer Networks », 4th International workshop on Adaptive and Dependable Mobile Ubiquitous Systems (UPC),

- in The 7th International Conference on Pervasive Services (ACM ICPS), 2010
- 28 Quang Anh Nguyen Vu, Benoit Gaudou, Richard Canal, Salima Hassas & **Frédéric Armetta**. «A cluster-based approach for disturbed, spatialized, distributed information gathering systems». in the first Pacific Rim workshop on Agent-based modeling and simulation of Complex Systems (PRACSYS), included into the The 13th International Conference on Principles and Practice of Multi-Agent Systems (PRIMA), 2010, **CORE B**
- 29 Quang Anh Nguyen Vu, Benoit Gaudou, Richard Canal, Salima Hassas, **Frédéric Armetta** & Serge Stinckwich, «Using trust and cluster organisation to improve robot swarm mapping». Workshop on Robots and Sensors integration in future rescue INformation system (ROSIN) in IEEE/RSJ International Conference on Intelligent Robots and Systems (IROS), 2010
- 30 Quang Anh Nguyen Vu, Benoit Gaudou, Richard Canal, Salima Hassas & **Frédéric Armetta**, «A Trust-based approach for maintaining coherence and robustness in a distributed information collecting system», in the proceedings of the third International Conference on Theories and Applications of Computer Science (ICTACS), 2010, **CORE B**
- 31 Gaël Clair, Leonardo Lana de Carvalho, **Frédéric Armetta**, Salima Hassas & Ederaldo José Lopes, «Controle endogeno a base de representação social». poster track, XL Reunião Anual da Sociedade Brasileira de Psicologia (SBP), 2010
- 32 Olivier Lefevre, **Frédéric Armetta**, Gaël Clair & Salima Hassas. «MANA : A new multi-agent approach for complex assignment problems». The First International Conference on Adaptive and Self-adaptive Systems and Applications (ADAPTIVE 2009), Athens/Glyfada, Greece, 2009

Conférences nationales à comité de lecture

- 33 **Frédéric Armetta**, Anthony Baccuet & Mathieu Lefort. « L'apprentissage algorithmique, une nouvelle étape pour l'IA. Une application aux opérations arithmétiques ». Conférence Nationale en Intelligence Artificielle (CNIA), PFIA, Strasbourg (France), pp. 31-39, 2023
- 34 Maya Medjad, Mathieu Buonomo, Raphaël Szymocha & **Frédéric Armetta**. « Un modèle pour la généricité des agents conversationnels vocaux multi-domaines ». RJCIA 2023, Strasbourg (France), 2023
- 35 Maxime Guériau, **Frédéric Armetta**, Salima Hassas, Romain Billot & Nour Eddin El Faouzi (2017). « Contrôle par apprentissage constructiviste pour le trafic coopératif ». Journées Francophones sur les Systèmes Multi-Agents (JFSMA), 6 juillet 2017, Caen (France), 2017
- 36 Sébastien Mazac, **Frédéric Armetta** & Salima Hassas. « Approche décentralisée pour un apprentissage constructiviste en environnement continu : application à l'intelligence ambiante. ». Journées Francophones sur les Systèmes Multi-Agents (JFSMA), 1 juillet 2015, Rennes (France), 2015

- 37 Maxime Guériau, Romain Billot, **Frédéric Armetta**, Salima Hassas & Nour-Eddin El Faouzi (2015). « Un simulateur multiagent de trafic coopératif ». 23es Journées Francophones sur les Systèmes Multi-Agents (JFSMA), Rennes (France), pp. 165-174, 2015
- 38 Quang Anh Nguyen, Benoit Gaudou, Richard Canal, Salima Hassas & **Frédéric Armetta**. « Stratégie de communication dans un système de collecte d'information à base d'agents perturbés », in actes des Journées Francophones des Systèmes Multi-Agents (JFSMA), Zahia Guessoum and Salima Hassas, eds., Cépaduès, Lyon, France, 2009

Articles dans revues internationales à comité de lecture

- 39 Lucien Tisserand, Brooke Stephenson, Heike Baldauf-Quilliatre, Mathieu Lefort, Frédéric Armetta, Unraveling the Thread : Understanding and Addressing Sequential Failures in Human-Robot Interaction, volume 11, *Frontiers in Robotics and AI*, 2024, **SJR Q2**
- 40 André Fabbri, **Frédéric Armetta**, Eric Duchene & Salima Hassas. « A Self-Acquiring Knowledge Process for MCTS ». *International Journal on Artificial Intelligence Tools*, vol. 25, n°01, p. 20. doi : 10.1142/S0218213016600071, 2016, **SJR Q3**
- 41 **Frédéric Armetta**, Mohammed Haddad, Salima Hassas & Hamamache Kheddouci. « A self-organized system improving inner topology for data sharing efficiency ». *Evolving Systems*, vol. 5, pp. 109-119. doi : 10.1007/s12530-013-9087-4, 2014, **SJR Q1**
- 42 Julien Monteil, Romain Billot, **Frédéric Armetta**, Salima Hassas & Nour-Eddin El Faouzi (2013). « Cooperative Highway Traffic : Multiagent Modeling and Robustness Assessment to Local Perturbations ». *Transportation Research Record*, vol. 2, pp. 1-10, **SJR Q2**
- 43 Quang Anh Nguyen Vu, Richard Canal, Benoit Gaudou, Salima Hassas & **Frédéric Armetta**. « TrustSets - Using trust to detect deceitful agents in a distributed information collecting system ». *Journal of Ambient Intelligence and Humanized Computing*, vol. 3, n°4, pp. 251-263, 2012, **SJR Q2**

Articles dans revue nationale à comité de lecture

- 44 Maxime Guériau, **Frédéric Armetta**, Salima Hassas, Romain Billot & Nour-Eddin El Faouzi. « Apprentissage constructiviste à base de systèmes multiagents. Une application au problème complexe de la régulation coopérative du trafic ». *Revue des Sciences et Technologies de l'Information - Série RIA : Revue d'Intelligence Artificielle*, vol. 32/2, pp. 249-277. doi : 10.3166/ria.32.249-277, 2018

Chapitres d'ouvrages

- 45 Quang-Anh Nguyen Vu, Benoit Gaudou, Richard Canal, Salima Hassas & **Frédéric Armetta**. « A Cluster-Based Approach for Disturbed, Spatialized, Distributed Information Gathering Systems ». *Principles and Practice of Multi-Agent Systems*, Desai, Nirmal and Liu, Alan and Winikoff, Michael, Lecture Notes in Computer Science, Springer, pp. 588-603. doi : 10.1007/978-3-642-25920-3_43, 2012
- 46 Leonardo Lana De Carvalho, Gaël Clair, **Frédéric Armetta**, Salima Hassas & Elayne De Moura Braga. « Représentations : Du Contrôle Endogène pour des Systèmes Complexes Adaptatifs ». *Temas em Ciências Cognitivas & Representação Mental*, Lopes, E. J., Porto Alegre : Sinopsys Editora, pp. 239-259, 2012
- 47 Patrick Albert, **Frédéric Armetta** & Salima Hassas. « Agents situés : une nouvelle voie pour le développement d'applications industrielles, chapter in "Technologies Multi-Agents pour l'Industrie" ». under the direction of Amal El Fallah-Seghrouchni and Jean-Pierre Briot, *Traités IC2*, édition Hermès-Lavoisier , p101-146, 2009

Logiciel

- 48 Elliot Faugier, **Frédéric Armetta**, Angela Bonifati & Bruno Yun (2024). « ADBL2 : Assisted Debate Builder with Large Language Models ». hal-04452246 , 2024
- 49 Julien Perier-Camby, **Frédéric Armetta**. Peppermint Conversational Framework. hal-04452246 <https://github.com/jperier/ppm-conv>, 2024

Brevets déposés

- 50 Simulation of dynamic city environments and a related fabrication method, Bernard Legaut , Lionel Barret , **Frédéric Armetta**, Saida Bouakaz , Salima Hassas. United States, Patent n° : US2009051686. 4905., <https://hal.science/hal-01499336>, 2009
- 51 Method of assigning a set of resources to multiple agents, Inventor : **Frédéric Armetta**, Publication number : US2008040178 (A1), Classification : - international : G05B19/418; G05B19/418 – European : G05B19/418P, Owner : Oslo, Publication date : 2008-02-14

2.4 Diffusion et rayonnement**Organisation de colloques, conférences, journées d'étude****Comités de programme**

- JFSMA 2014-2018 (Journées Francophones sur les Systèmes Multi-Agents)
- BICA'15 (Biologically Inspired Cognitive Architectures Society)

- DEMO and CONTEST, membre du jury, IEEE SASO'12 (IEEE International Conference on Self-Adaptive and Self-Organizing Systems)
- SIS (Smart Intelligent Industrial Systems)@HoloMAS'2011 (Industrial Applications of Holonic and Multi-Agent Systems)
- Poster Track, SASO'2009
- ASCWKM (Agent Supported Collaborative Work for Knowledge Management) @ IEEE
- ICDIM'2009 (International Conference on Digital Information Management)
- SARC (Self-Adaptation for Robustness and Cooperation in Holonic MultiAgent Systems and in MultiAgent Systems, International Workshop @ IEEE SASO'2009
- MA4CS (MultiAgents for modelling Complex Systems) @ ECCS'2007 (European Conference on Complex Systems), Dresden, Germany 2007

Comité de lecture

- AAMAS'18 (International Conference on Autonomous Agents and Multiagent Systems), classification CORE : rang A*
- IJCAI'17 (International Joint Conference on Artificial Intelligence), classification CORE : rang A*
- ACM TAAS'17 (Transactions on Autonomous and Adaptive Systems)
- RIA'14 (Revue d'Intelligence Artificielle)
- MISC'12 (International Symposium on Modelling and Implementation of Complex Systems)
- Journal of Intelligent Manufacturing'10'11'12
- AAMAS'10 (International Conference on Autonomous Agents and Multiagent Systems)
- ACM TAAS'08 (Transactions on Autonomous and Adaptive Systems)
- SASO'08 (International Conference on Self-Adaptive and Self-Organizing Systems)
- ESAW'08 (The 9th Annual International Workshop "Engineering Societies in the Agents World")
- JFSMA'07'08 (Journées Francophones des Systèmes MultiAgents)

Comité d'organisation

- **Publicity Chair**, IEEE SASO'2015, <http://saso2015.mit.edu/>
 - conférence internationale de référence dans le domaine des systèmes auto-organisés
 - lancement des différentes communications et synthèses (conférence principale et 6 workshops), participation à la constitution d'une liste des chercheurs de la communauté, recherche, enregistrement et utilisation de listes de diffusions de communautés proches
- **Finance Chair** de la conférence IEEE SASO'2012 (<http://saso2012.univ-lyon1.fr/>)

La conférence internationale de référence dans le domaine des systèmes auto-organisés. 120 chercheurs accueillis, 10 Workshops. Événement organisé à Lyon pendant une semaine. Correspondant Computer Science IEEE et organisation sur place.

— **WebMaster**

- JFSMA'09 <http://liris.cnrs.fr/sycosma/JFSMA09/>, et organisation sur place
- MA4CS'07

Expertises

- Reviewer pour un projet de l'appel à projets générique 2024 de l'ANR (phase 2)
- Membre du comité d'expertise des appels à propositions 2019 et 2020 de l'Initiative d'Excellence Paris Seine
- Expertise internationale d'un dossier de bourse de doctorat de l'AUF 2009

Participation à des jurys de thèse et de HDR (hors établissement)

- Membre du jury de la thèse de de Thilbault Fourez en tant qu'examineur, laboratoire IRIT (Toulouse), soutenue le 18 juin 2024.

Diffusion du savoir (vulgarisation), responsabilités et activités au sein de sociétés savantes ou associations

- CES Innovation Award 2015-2016 attribué à notre partenaire UBIANT
UBIANT remporte un CES Innovation Award dans la catégorie "Tech For A Better World" (technologie pour un monde meilleur) décerné par le CES (Consumer Electronics Show). Ce prix international très compétitif récompense aussi les travaux de thèse de Sébastien Mazac et la coopération fructueuse avec l'équipe SMA du LIRIS (CIFRE). Ce prix a motivé l'Interview BFM business du président de UBIANT (présentation des thématiques de recherche en intelligence ambiante) : Ubiant, la start-up qui équilibre confort et énergie dans la maison.
- Présentation de l'approche Guided-CESNA pour le projet ALF (Aires de Livraison du Futur), au Colloque "La logistique à l'ère de la ville numérique" organisé par le PREDIT (Programme de Recherche et d'Innovation dans les Transports Terrestres) à l'hémicycle de l'Hôtel de Région Rhône-Alpes, le 11 Avril 2013. Le projet ALF (Aires de Livraison du Futur) a été promu Lauréat du PREDIT4. J'ai représenté ensuite l'équipe SMA lors de la cérémonie de remise de prix, Palais Brongniart, Paris, 7 et 8 Octobre 2013.

Participation à un réseau de recherche

- Séminaire invité sur les Self-*Systems, à la 6ème édition du Workshop Franco-Japonais, Paris, 11-12 juin 2013
- Présentation au stand LIRIS du Big Bang de la Robotique <http://www.coboteam.fr/evenement/19-mars-lancement-de-la-demarche-robotique-rhone-a> Lyon, 2015

- GT : IMS2-Intelligent Manufacturing & Services Systems groupe du GDR MACS (Modélisation, Analyse, Conduite des Systèmes dynamiques), Strasbourg, 2009.
- Participation au projet de recherche « GET - Gouvernance des Systèmes Complexes : une vision Enactive, Application au domaine des Transports ». Projet soutenu par l'Institut Rhône-Alpin des Systèmes Complexes (IXXI) - (2008-2010), en partenariat avec le LORIA-Nancy , le LIG-Grenoble , le LET-Lyon2 et Communauté Urbaine du Grand Nancy (CUGN).
- Intervenant invité au colloque du numérique 2009, Eurexpo, Lyon.
- Intervenant invité à la session " Speed Gaming ", GDC Paris (Game Developers Conference), Cœur Défense, Paris, 2008

2.5 Responsabilités scientifiques

Responsable de projets financés

- Projet collaboratif pluridisciplinaire PepperMint (230 K€, Sept. 2021 - Déc. 2024)
 - Porteur avec Heike Baldauf-Quilliatre (laboratoire ICAR, Analyse Conversationnelle) depuis janvier 2023
 - Étude des pratiques incarnées du turn-taking dans l'interaction humain-robot LabEx ASLAN.
 - Partenaires : LIRIS (équipes SyCoSMA, SAARA), ICAR, Université de Oulu-Finland (GenZ)
 - 1 Post Doc LIRIS (12 M), 1 Post Doc ICAR (24 M), 1 Ingénieur de recherche (IR) LIRIS (12 M) + 1 ingénieur de recherche (IR) LIRIS (10 M), 2 stages M2 (complément enveloppe de l'équipe SyCoSMA du LabEx), co-encadrement LIRIS Salima Hassas, Mathieu Lefort, Erwan Guillou
- Apprentissage pour la reconnaissance d'intention et la gestion de conversations (45 K€, Jan. 2023 - Déc. 2025)
 - Direction
 - Contrat de collaboration CIFRE, ANRT
 - Société Reecall (<https://reecall.com>)
 - Thèse Maya Medjad <https://liris.cnrs.fr/these/these-maya-medjad> (co-encadrement LIRIS Bruno Yun)
- Génération de recettes culinaires (97K€, Sept. 2019 - Avr. 2021)
 - Porteur <https://www.pulsalys.fr/startup/popotte-duck>
 - Contrat de maturation Pulsalys, transfert
 - Partenaires : Laboratoire LIRIS, Pulsalys, Entreprise PopotteDuck
 - Post-Doc (12M), Ingénieur Recherche (3M), Stage M2, co-encadrement Pulsalys Bruno Malfoy (Chargé de développement technologique)

- IA décisionnelle pour les CRM (Sept. 2017 - Jan. 2019, 30 K€)
 - Porteur, 2 contrats de collaboration
 - Module SICAIA de la société GCR Contact
 - "Développé en collaboration avec le CNRS (laboratoire d'excellence du LIRIS), SICAIA bénéficie des dernières avancées dans le domaine du "Machine learning" et du "Process Mining" : extrait du site de GRC Contact <https://www.grc-contact.fr/intelligence-artificielle-crm-sicaia>
 - 3 stages financés, co-encadrement LIRIS Haytham Elghazel

Membre associé de projets financés

- Multimodal deep SensoriMotor Representation learning (MeSMRise) (Avr. 2024 - Oct. 2028, 511k€)
 - Projet ANR (PRC) financé
 - Impliqué dans les workpackages *Object representation learning* et *Action decision learning*
 - Partenaires : Laboratoire LIRIS (porteur Mathieu Lefort, équipe SyCoSMA), Laboratoire Jean Kuntzman, Institut Pascal
- Agents conversationnels orientés but (22.5K€, Mai 2021 – Oct. 2022)
 - Contrat de collaboration société Recall (<https://recall.com>)
 - Associé puis Porteur, l'expertise en traitement naturel du langage est devenue centrale au projet dont j'ai pris la direction
- SMARTER : Structured Multi-sourced AI Reasoning for Enhanced Decision-Making (6K€, Jan. 2024 - Dec. 2025)
 - Projet interne financé par le laboratoire LIRIS
 - porteur Bruno Yun
 - Équipes : SyCoSMA (porteur Bruno Yun), BD
 - Expertise en modèles de langue et mécanismes de co-construction cognitifs
- HAGARICE : Harnessing Argumentation Graphs in Augmented Reality for Immersive Co-Creation and Exploration (15K€, 2024)
 - AAP Accueil EC 2024 de l'UCBL
 - porteur Bruno Yun
 - Expertise en modèles de langue et mécanismes de co-construction dans un environnement de réalité augmentée pour l'argumentation. Ce projet s'appuie sur le projet SMARTER, pour une application dans un environnement interactif de réalité augmenté

3 Investissement pédagogique

3.1 Présentation synthétique de l'activité d'enseignement

Mes enseignements portent principalement sur l'apprentissage de la programmation en troisième année de licence (paradigme objet, analyse et conception), sur différents approfondissements et mises en application (projets, design pattern, UML, programmation modulaire, graphique, événementielle, architecture MVC Strict, réseau, etc.), et sur l'enseignement de l'intelligence artificielle (réseaux de neurones, modèles de langue, apprentissage par renforcement, graphes) en master 2. J'encadre également régulièrement des stages de 4ème et 5ème année sur mes différents projets, et suis tuteur d'étudiants du master 2 IA (Lyon1).

Je privilégie l'interaction avec les étudiants lors de mes enseignements (variation des supports, alternance entre acquisition théorique et pratique, sujets ludiques adaptés) afin d'intéresser les étudiants dans leurs apprentissages. Ces dernières années, le plagiat de code informatique a considérablement augmenté ce qui m'a amené à adapter les travaux de programmation évalués donnés aux étudiants. D'une part, je propose de nouveaux sujets de projets systématiquement (chaque semestre pour les niveaux L3 (Lyon1 ou Polytech), je mets à jour des sujets pour les M2, avec un accompagnement dédié aux projets (analyse guidée, diagrammes, code spécifique que je fournis qui permet d'accompagner l'analyse guidée et d'amorcer rapidement le travail pour les étudiants). Nous analysons les projets soumis par les étudiants avec un logiciel de détection de plagiat dédié aux codes informatiques. Bien que les étudiants soient encouragés à échanger durant les séances, et travailler en binômes, je tiens à ce que les évaluations soient considérées individuellement pour plus d'équité. Je me suis depuis longtemps investi en tant que responsable de modules, ce qui nécessite un fort investissement, de nombreux ajustements, particulièrement concernant le module de programmation orienté objet de L3 qui accueille annuellement 5 à 6 groupes TD d'étudiants (hors dépassements), et pour lequel je compte un nombre réduit d'enseignants qui peuvent intervenir durablement au sein de l'UE. L'encadrement se compose ainsi majoritairement de vacataires, d'ACEs et d'ATERs ce qui nécessite un suivi et un effort de coordination prononcés, mais permet aussi de dynamiser le module.

3.2 Présentation des enseignements

J'interviens pour des CM, TD et TP en formation initiale troisième année de licence informatique et en master 2 IA à l'université Lyon1. En troisième année, les effectifs correspondent en moyenne à 5 groupes de 35 étudiants (mais suivant les années, a atteint ponctuellement jusqu'à 270 étudiants). Le master 2 IA est quant à lui limité à 1 groupe de TD. J'interviens également en CM, TD et TP en formation initiale (1 groupe) et en formation par apprentissage (1 groupe) en 3ème année (équivalent L3) à l'école

d'ingénieur Polytech Lyon (1/3 de mon service statutaire ces dernières années) tel que détaillé en annexe 5.1.

3.3 Responsabilités pédagogiques

- Responsable de la partie **traitement naturel du langage du module Bio-Inspired Machine Learning du Master 2 IA de Lyon** (depuis 2021). Au sein du module, encadrement de travaux pratiques en apprentissage par renforcement (2021-2022, 2022-2023) et apprentissage profonds appliqué aux graphes (2023-2024). 30 à 40 étudiants par an.
- **Responsabilité du module de programmation orientée objet - Troisième année de licence informatique**, 170 à 200 étudiants par an
 - 2010 - 2020 : co-responsable (60%)
 - 2020 - en cours : responsable (100 %)
 - Organisation des enseignements, cours magistraux, recrutement et formation des enseignants TD et TP (doctorants, ATER, peu de permanents), planification, organisation des soutenances de projets, surveillances, etc.
- **Responsable de 2 modules de Projets informatiques encadrés** (3ème année Info Polytech depuis 2010 (40 étudiants par an) et 3ème année Info Apprentis Polytech depuis 2012 (15 à 20 étudiants))

4 Responsabilités collectives et d'intérêt général

4.1 Responsabilités administratives

Responsable de l'équipe SyCoSMA du LIRIS

- Prise de responsabilités début 2023
- SyCoSMA compte 23 membres, dont 1Pr., 6 Mcf. <https://liris.cnrs.fr/equipe/sycosma>
- L'équipe SyCoSMA est une des douze équipes du LIRIS (330 membres, <https://liris.cnrs.fr/liris>)
- L'activité de recherche de l'équipe SyCoSMA s'intéresse à l'étude de systèmes intelligents autonomes, capables d'auto-développements cognitifs en interactions avec l'environnement impliquant ou non une dynamique distribuée. Les modèles développés s'appuient et mènent à des propositions dans les domaines du *deep learning*, de l'apprentissage par renforcement, des théories sensori-motrices, des systèmes multi-agents, des agents conversationnels et des méthodes d'argumentation.
- Implications :
 - Positionnement scientifique de l'équipe en concertation avec les membres permanents, présentation du bilan mi-parcours à la direction du LIRIS en janvier 2024, projet d'évolution de l'équipe pour la prochaine évaluation HCERES, avis/participation à l'arbitrage pour certains dépôts de projets
 - Coordination avec la direction (diffusion des opportunités, remonté d'informations), participation à la préparation des profils et à leur diffusion, à la constitution des COS (recrutement d'un Mcf (2023, Bruno Yun), d'un Pr (2024, Maxime Morge), entretiens avec les candidats, administration, budget équipe

4.2 Responsabilités et mandats locaux ou régionaux

Membre de conseils

- Membre du conseil scientifique du LIRIS qui se réunit mensuellement, depuis début 2023. Ce conseil discute des orientations scientifiques à donner pour le laboratoire.
- Membre élu du conseil du laboratoire GAMA : 2011-2012
- Membre élu du conseil du département informatique Lyon1 : 2009-2013

Membre de comités

- Membre élu du comité consultatif du département informatique de la FST 2013-2016

- Membre invité du comité de spécialistes : MCF61, "Modélisation et commande décentralisée de systèmes complexes", MCF Grenoble INP-ESISAR Valence, 2014
- Membre local des comités de spécialistes : MCF27, "Graph", Lyon, 2010 - MCF27, "Image", Lyon, 2011 - MCF27, "SMA", Lyon, 2012, MCF27, "SMA", Lyon, 2015
- Membre local des comités de suivi de thèse de Méghane Decroocq et Thomas Ranvier (thèses soutenues en 2023), Aymar Tchagoué, Ichrak Ennaceur, chair d'une session de la journée des thèses du LIRIS (7 mars 2024), Interaction & cognition, Simulation et Sciences du vivant

4.3 Autres responsabilités

Administration informatique

- Administration d'une plateforme collaborative MatterMost (2018..2023)
 - Les plateformes de communication collaborative et de gestion de projets jouent un rôle important dans nos activités, particulièrement depuis la période de Covid et le développement du travail à distance.
 - Dans le cadre de la transition numérique, en tant d'informaticien, j'ai installé et administré une machine virtuelle ouverte sur l'extérieur du réseau Lyon1, hébergeant la plateforme Mattermost <https://teamwork.liris.cnrs.fr>. Celle-ci a répondu a un besoin des membres du LIRIS, notamment suite aux modifications des conditions d'utilisation des plateformes non institutionnelles équivalentes telles que Slack <https://slack.com/> (protection des données, confidentialité, etc.).
 - Cette plateforme a accueilli depuis 2018 plus de 85 utilisateurs actifs (30000 posts), la cellule Informatique du LIRIS a décidé fin 2023 de valider ce service et de le pérenniser, avec la mise en place d'une version intégrée a GitLab et à son système d'information.
- Administration et mise à disposition d'une machine de calculs scientifique (2018..2023) :
 - Mise a disposition pour les doctorants et stagiaires de l'équipe d'une machines dédiée au machine learning.
 - Carte de calcul GPGPU Quadro P6000 (6 K€) de mon projet de recherche accepté par le programme de collaborations académiques de NVIDIA. Point d'accès au calculateur Jean Zay.

5 Annexes

5.1 Tableau des enseignements

Année	Niveau	Diplôme	Intitulé	Type (1)	Nature (2)	Effectifs	Volume (eq. TD)
2023-2024	M2	IA	Bio-Inspired machine learning	I	3CM, 12TP	25	16.5
	M2	IA	Suivi stage	I	6S	2	6
	L3	Lyon1	Algorithmique et programmation orientée objet	I	48CM, 57TP	171	129
	L3	Polytech	Projet informatique encadré	I	4.5CM, 9TD, 19.5TP	39	35.25
	L3	Polytech App	Projet informatique encadré	A	4.5CM, 6TD, 9TP	17	21.75
2022-2023	M2	IA	Bio-Inspired machine learning	I	3CM, 12TP	41	16.5
	M2	IA	Suivi stage	I	6S	2	6
	L3	Lyon1	Algorithmique et programmation orientée objet	I	48CM, 57TP	161	129
	L3	Polytech	Projet informatique encadré	I	4.5CM, 9TD, 19.5TP	33	35.25
	L3	Polytech App	Projet informatique encadré	A	4.5CM, 6TD, 9TP	19	21.75
2021-2022	M2	IA	Bio-Inspired Intelligence	I	3CM, 3TD, 9TP	39	16.5
	L3	Lyon1	Algorithmique et programmation orientée objet	I	34CM, 72TP	167	123
	L3	Polytech	Projet informatique encadré	I	4.5CM, 9TD, 19.5TP	41	35.25
	L3	Polytech	Réseaux d'ordinateurs	I	9TD, 6TP	41	15
	L3	Polytech App	Projet informatique encadré	A	10.5TD, 6TP	19	16.5
	L3	Polytech App	Réseaux et Systèmes de Communication	A	4.5TD, 3TP	19	7.5

Année	Niveau	Diplôme	Intitulé	Type (1)	Nature (2)	Effectifs	Volume (eq. TD)
2020-2021	L3	Lyon1	Algorithmique et programmation orientée objet	I	32CM, 72TP	191	120
	L3	Lyon1	Suivi stage	S	2S	2	2
	L3	Polytech	Projet informatique encadré	I	4.5CM, 9TD, 19.5TP	40	35.25
	L3	Polytech	Réseaux d'ordinateurs	I	9TD, 6TP	40	15
	L3	Polytech App	Projet informatique encadré	A	9TD, 10.5TP	20	19.5
	L3	Polytech App	Réseaux et Systèmes de Communication	A	4.5TD, 3TP	20	7.5
2019-2020	L3	Lyon1	Algorithmique et programmation orientée objet	I	32CM, 74TP	253	124
	L3	Polytech	Projet informatique encadré	I	4CM, 8TD, 24TP	40	38
	L3	Polytech	Réseaux d'ordinateurs	I	8TD, 8TP	40	16
	L3	Polytech App	Projet informatique encadré	A	8TD, 12TP	1 Gr.	20
	L3	Polytech App	Réseaux et Systèmes de Communication	A	4TD, 4TP	1 Gr.	8

(1) formation initiale (I) / alternance (A)

(2) cours magistraux, TP, TD, encadrement de travaux de fin d'étude et de stages (S)

(3) Intelligence Artificielle (IA), Département informatique de Lyon 1 (Lyon 1), filière informatique de Polytech Lyon (Polytech), filière informatique de Polytech Lyon Apprentissage (Polytech App)

5.2 Liste des encadrements scientifiques de master

Encadrements stages de master :

- Sept. 2024, stage M1 Polytech, Farah Ben Slama, encadrement 100%
- Jan. 2024, stage de M2 IA de Elliot Faugier, projet SMARTER, co-encadrement 25%, co-encadrants Bruno Yun et Angela Bonifati [1][48]
- Mai 2021, stage 5ème année ENS de Pierre Marec encadrement 100%
- Janvier 2020, stage M2 IA d'Anthony Baccuet, encadrement 100% [3][33]

- Janvier 2020, stage M2 IA Valentin Chaffraix, étude d'un modèle d'attention visuelle de deep learning dans un contexte sensorimoteur encadrement 33%, co-encadrants Mathieu Lefort, Stefan Duffner [4]
- Septembre 2019. stage 4A Bastien Nollet, encadrement 100% [7]
- Fev. 2018, projet IA décisionelle pour les CRM, stage 5A INSA de Maxime Le Duc, encadrement 100%
- Jan. 2018, Reconnaissance et suivi d'objets dans une séquence vidéo par des paradigmes d'apprentissage constructivistes, porteur du projet transverse accepté et financé par le LIRIS, stage de M2 Recherche Louis Manhès, co-encadrement 33%, co-encadrants Mathieu Lefort et Stefan Duffner
- 2016-2018, Apprentissage profond et animation basée squelette pour le transfert de style, Projet transverse financé par le LIRIS [équipe SMA/équipe SAARA], 3 stages de M2 Recherche, Florent Gattoni, Pierre Couy, Etienne Desbois co-encadrement 33%, co-encadrants Alexandre Meyer et Mathieu lefort
- 2018, Sujet d'ouverture sur la génération et la co-construction musicale à base de deeplearning, projet étudiant de M1 Guilhem Marion, encadrement 100%
- Sept. 2017, projet IA décisionelle pour les CRM, collaboration avec l'entreprise GRC Contact, stages de 4ème année Polytech MAM et INFO de Philomène Bobichon et Mathieu Regnard, co-encadrement 70%, co-encadrant Haytham Elgazel
- Jan. 2017, Reconnaissance et suivi d'objets dans une séquence vidéo par des paradigmes d'apprentissage constructivistes, porteur du projet transverse accepté et financé par le LIRIS, stage de M2 Recherche Nawel Medjkoune, co-encadrement 33%, co-encadrants Mathieu Lefort et Stefan Duffner [12]
- Jan. 2017, Accompagnement Startup DealOn, porteur, 1 stage M2 international, Feras AL Kassar, encadrement 100% [11]
- Jan. 2017, collaboration interdisciplinaire avec Marc Bertin (laboratoire ELICO), 2 stages de M1, co-encadrement 50% [8][9][10]
- Jan. 2017, Apprentissage par curiosité, stage de M2 de Tom Dusséaux, co-encadrant Mathieu Lefort co-encadrement 50%
- Jan. 2013, « Emergence de langage : une approche à base d'agents pour des systèmes intelligibles », stage de M2 IA de Kevin Sanlaville, co-encadrement à 50%, co-encadrante Salima Hassas
- Jan. 2011, "Graphes et Auto-organisation pour des échanges pair-à-pair plus performants", stage de M2 IA de M. E. Gazbour, co-encadrement à 50%, co-encadrant Mohammed Haddad
- Jan. 2011 : "Négociations mixtes dans un système multi-agents pour le développement de stratégies évolutives dans les jeux vidéos", stage de M2 de Cyril Moron, co-encadrement à 33%, co-encadrante Salima Hassas
- Jan. 2008, "Etude de la relation local/global dans les approches émergentes par système multi-agents : une application à la K-coloration", stage de M2 IA de B. Smati, co-encadrement à 50%, co-encadrante Salima Hassas