



**HAL**  
open science

# Les identités au centre de la mise en oeuvre de comportements dans le cadre de collectifs multi-agents : application au Web des Objets

El Mehdi Khalfi

## ► To cite this version:

El Mehdi Khalfi. Les identités au centre de la mise en oeuvre de comportements dans le cadre de collectifs multi-agents : application au Web des Objets. Systèmes embarqués. Université Grenoble Alpes, 2018. Français. NNT : 2018GREAM068 . tel-02080299

**HAL Id: tel-02080299**

**<https://theses.hal.science/tel-02080299v1>**

Submitted on 26 Mar 2019

**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.



## THÈSE

Pour obtenir le grade de

## DOCTEUR DE LA COMMUNAUTÉ UNIVERSITÉ GRENOBLE ALPES

Spécialité : Informatique

Arrêté ministériel : 25 mai 2016

Présentée par

**El mehdi KHALFI**

Thèse dirigée par **Jean-Paul JAMONT**, UGA

et codirigée par **Michel OCCELLO**, UGA

préparée au sein du **Laboratoire de conception et d'intégration  
des systèmes**

dans l'**École Doctorale Mathématiques, Sciences et  
technologies de l'information, Informatique**

**Les identités au centre de la mise en œuvre  
de comportements dans le cadre de  
collectifs multi-agents : application au Web  
des Objets**

**Towards an Identity-based Cooperation in  
Coexisting Multiagent Systems**

Thèse soutenue publiquement le **28 novembre 2018**,  
devant le jury composé de :

**Monsieur JEAN-PAUL JAMONT**

MAITRE DE CONFERENCES, UNIVERSITE GRENOBLE ALPES,  
Directeur de thèse

**Madame MARIE-PIERRE GLEIZES**

PROFESSEUR, UNIVERSITE TOULOUSE-III PAUL-SABATIER,  
Examineur

**Monsieur VINCENT CHEVRIER**

PROFESSEUR, UNIVERSITE DE LORRAINE, Président

**Monsieur MICHEL OCCELLO**

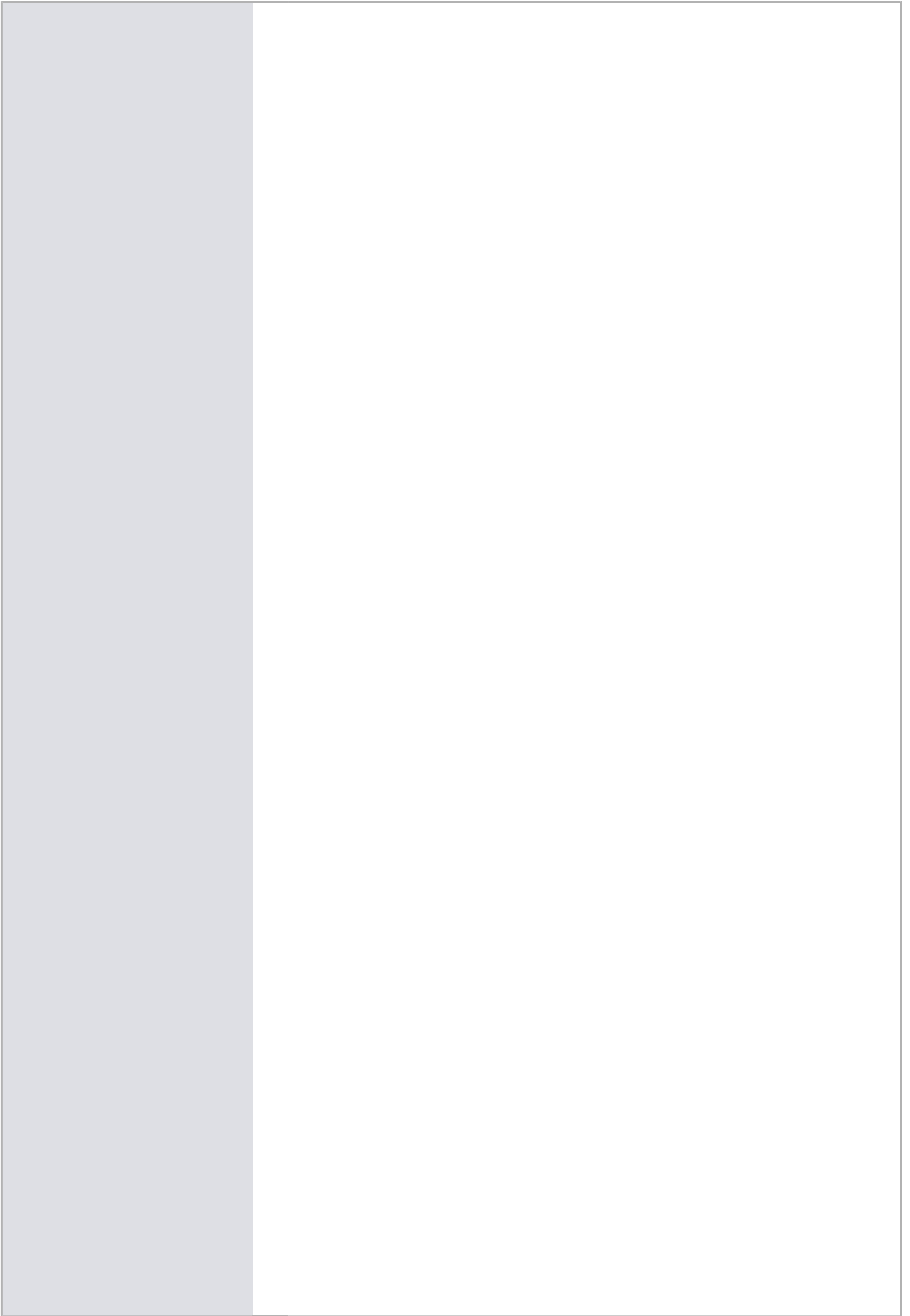
PROFESSEUR, UNIVERSITE GRENOBLE ALPES, Co-directeur de  
thèse

**Monsieur NICOLAS SABOURET**

PROFESSEUR, UNIVERSITE PARIS-SUD, Rapporteur

**Monsieur RENE MANDIAU**

PROFESSEUR, UNIV. POLYTECHNIQUE DES HAUTS-DE-FRANCE,  
Rapporteur







Thèse effectuée au sein du **Laboratoire de Conception et d'Intégration des Systèmes**

de l'Université Grenoble Alpes

50, rue Barthélémy de Laffemas

BP54 26902

VALENCE Cedex 09

France



## Résumé

Avec le développement des objets connectés, les agents embarqués déployés dans des environnements physiques et les applications multi-agents qui les impliquent deviennent de plus en plus diffus. Plusieurs systèmes multi-agents sont amenés à partager le même environnement physique. Cette cohabitation d'agents de SMA différents, qui n'ont pas nécessairement été prévus pour interagir entre eux, amène cependant ces agents à se solliciter. Un agent peut alors participer à la réalisation d'objectifs incompatibles avec les siens ou avec les objectifs des systèmes et des coalitions auxquels il appartient.

Notre objectif est de doter ces agents de la capacité de décider de leur participation ou non à des actions collectives tout en préservant leurs intérêts individuels et collectifs. Ainsi, nous proposons le modèle d'agent, IDEAS (*IDEntity Agent System*) qui permet aux agents cohabitants dans des environnements *multi-SMA* de construire une représentation des intérêts individuels et collectifs avant de décider de s'engager dans la réalisation d'actions, ils raisonnent sur ces représentations de manière à préserver leurs intérêts. Ce modèle est principalement basé sur *la théorie de l'identité*, le cadre théorique que nous avons choisi pour son intérêt à la définition des comportements des individus dans des groupes en étudiant leurs dynamiques intra-personnelle, inter-personnelle, et inter-groupe. Nous nous inspirons de cette théorie pour implémenter un processus de construction et de maintenance de l'identité d'agent, ce qui lui permettra d'avoir une représentation explicite de ses intérêts individuels et collectifs. Ensuite, pour permettre aux agents de raisonner sur les sollicitations des autres agents afin de décider de coopérer avec eux ou non, nous nous inspirons d'un concept philosophique qui est à la base de l'interaction humaine : l'*intersubjectivité*. C'est l'idée de prendre en considération la signification intentionnelle des actions des autres lors de l'interaction avec eux. Ce concept est implémenté dans le modèle IDEAS sous la forme d'un algorithme de reconnaissance d'intention multi-agent. En employant ce modèle, l'agent peut décider de son engagement de coopérer avec les autres en correspondant ce qu'il a prédit de leurs sollicitations avec la représentation de son identité qu'il a construite. Notre approche est validée par des expérimentations illustrant l'exploitation du modèle IDEAS sur un scénario de viticulture (ANR ASAWoO) et sur d'autres scénarios plus complexes qui confirment la pertinence de notre contribution.

# Abstract

Embedded agents deployed in physical environments are increasingly interoperable, and are likely to coexist with agents of others systems in a same physical space. So, an agent needs to be able to cooperate with agents from other systems and to form coalitions with unfamiliar teammates. However, before committing to cooperate with others, an agent must take into account that it may get involved in the achievement of objectives that are incompatible with its own, with the global objectives of its system, or with those derived from its previously joined coalitions. Our goal is to endow these agents with the ability to decide whether or not to participate in collective actions while preserving their individual and collective interests. Thus, we propose an agent model, which we called IDEAS (*IDEntity Agent System*), to enable cohabiting agents in *multi-MAS* environments to construct a representation of the individual and collective interests that they have to preserve, and decide what actions to commit to execute according to their interests. This model is mainly based on *the identity theory*, the theoretical framework we chose for its interest in defining behaviors of individuals in groups by studying their intra-personal, inter-personal, and inter-group dynamics. We draw inspiration from the identity theory to implement a agent identity construction process, which will enable it to have an explicit representation of its individual and collective interests. Then, to be able to reason on the solicitations of other agents in order to decide to cooperate with them or not, our contribution is inspired by a philosophical concept which is at the base of the human interaction : *intersubjectivity*. It is the idea of taking into consideration the intentional meaning of the actions of others when interacting with them. This concept is implemented in IDEAS model as a multi-agent intent recognition system. By employing this model, the agent can decide on his commitment to cooperate with others by matching what he has predicted from their solicitations with the representation of his constructed and maintained identity. Our approach is validated by experiments illustrating the exploitation of IDEAS model on a viticulture scenario and on other more complex scenarios which confirm the relevance of our contribution.



# Remerciements

Ce travail n'aurait jamais vu le jour sans l'aide de nombreuses personnes. Je profite de l'occasion pour remercier les personnes qui ont participé, de près ou de loin, à la réalisation de cette thèse.

Avant toute chose, je remercie mes directeurs de thèse Dr. Jean-Paul Jamont, et Pr Michel Occello de m'avoir donné la possibilité de mener à bien ce projet de thèse et de m'avoir fourni des conseils précieux tout au long du processus. Jean-Paul, merci de m'avoir présenté le merveilleux domaine des systèmes multi-agents, de m'avoir invité à rejoindre le laboratoire LCIS pour travailler sur ce sujet, et en particulier d'avoir cru en moi et de m'avoir soutenu dans toutes mes activités. Avoir eu la possibilité de travailler avec un superviseur aussi optimiste et encourageant a complètement modifié le cours de ma vie de la meilleure façon possible. Michel, votre point de vue ne manque jamais d'élucider les forces, mais aussi les faiblesses de mon travail.

Je remercie très chaleureusement tous les membres du jury d'avoir accepté de participer à l'évaluation de mon travail de thèse : je remercie Pr. Vincent Chevrier de me faire l'honneur de présider mon jury de thèse. Je remercie Pr. René Mandiau et Pr. Nicolas Sabouret d'avoir accepté d'être rapporteurs de mon travail de thèse et d'avoir apporté des jugements constructifs et des pistes d'amélioration de ce travail. Je remercie Pr. Marie-Pierre Gleizes d'avoir accepté d'examiner mon travail et de ses retours précieux sur le manuscrit.

Je remercie Laurent Lefèvre, chef de l'équipe MACSY-COSY lors de mon arrivée au LCIS, puis directeur du laboratoire LCIS. Je tiens à remercier aussi le corps technique et administratif du LCIS pour leur soutien technique, personnel et professionnel tout au long de mon séjour au laboratoire. Merci à vous Jennyfer, Caroline, Cédric, et Karim.

Doctorants camarades que j'ai rencontrés au cours de mon séjour au LCIS, je vous remercie pour vos discussions, votre soutien et vos encouragements. Merci à Afra, Khadija, Karem, Ayoub, Yassine, Youness, Huy, Hung, Lai, Thinh. J'espère que nous resterons en contact.

Cette thèse a été financée par l'Agence Nationale de la Recherche sous le numéro <ANR13-INFR-012>.



# Table des matières

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Vers l'Internet des objets sociaux . . . . .	1
1.2	De la coexistence des agents vers la coopération des SMA . . . . .	2
1.3	Problématique . . . . .	3
1.4	Contributions . . . . .	5
1.5	Structure du manuscrit . . . . .	6
<b>2</b>	<b>La coopération et ses potentielles interférences</b>	<b>9</b>
2.1	Spectre des interactions sociales coopératives . . . . .	10
2.1.1	Communication par environnement . . . . .	10
2.1.2	Coordination non concertée . . . . .	11
2.1.3	Communication explicite . . . . .	12
2.1.4	Coordination planifiée . . . . .	12
2.1.5	Coopération . . . . .	12
2.2	Caractéristiques de la coopération . . . . .	13
2.2.1	Classification des types d'interactions sociales coopératives . . . . .	14
2.2.2	L'intentionnalité . . . . .	15
2.2.3	L'engagement . . . . .	15
2.2.4	La compatibilité des objectifs . . . . .	16

2.3	L'interférence dans le cadre de la coopération . . . . .	17
2.3.1	L'interférence positive . . . . .	18
2.3.2	L'interférence négative . . . . .	19
2.3.2.1	Classification des types de conflits . . . . .	19
2.3.2.2	La gestion des conflits . . . . .	21
2.3.3	Interférence combinée . . . . .	22
2.4	Synthèse . . . . .	23
<b>3</b>	<b>L'identité au centre de la coopération</b>	<b>25</b>
3.1	L'identité et son rôle dans le comportement de l'individu . . . . .	26
3.1.1	Le soi et l'identité . . . . .	27
3.2	L'identité personnelle . . . . .	30
3.2.1	En philosophie/psychologie . . . . .	31
3.2.2	En sociologie . . . . .	32
3.3	L'identité sociale . . . . .	33
3.3.1	La théorie de l'identité sociale . . . . .	34
3.3.2	Types d'identités sociales . . . . .	37
3.4	L'identité collective . . . . .	38
3.5	Modèles de l'identité . . . . .	39
3.5.1	Un processus de construction d'identité sociale . . . . .	39
3.5.2	Un modèle d'identité . . . . .	40
3.6	Synthèse . . . . .	42
<b>4</b>	<b>IDEAS : un modèle d'agent centré identité</b>	<b>47</b>
4.1	Le modèle d'agent de référence . . . . .	48
4.1.1	BDI . . . . .	48
4.1.2	Cycle de vie d'un agent BDI . . . . .	49
4.2	Présentation du modèle IDEAS . . . . .	53
4.2.1	Architecture d'agent . . . . .	53
4.2.2	Construction et maintenance d'identité . . . . .	55

---

4.2.3	Processus d'engagement centré identité . . . . .	58
4.2.3.1	Modèle d'engagement . . . . .	60
4.2.3.2	Fonction de score d'identité . . . . .	60
4.2.4	Cycle de vie . . . . .	61
4.3	Synthèse . . . . .	63
<b>5</b>	<b>Implantation</b> . . . . .	<b>65</b>
5.1	Contexte applicatif . . . . .	66
5.1.1	Défis . . . . .	66
5.1.2	Architecture de l' <i>avatar</i> . . . . .	68
5.2	Étude de cas . . . . .	71
5.2.1	Scénario . . . . .	71
5.2.2	La planification par réseaux hiérarchisés de tâches . . . . .	72
5.3	La construction de l'identité d'un agent IDEAS . . . . .	77
5.3.1	Le mécanisme de reconnaissance d'intention dans IDEAS . . . . .	79
5.3.2	Le processus de construction d'identité . . . . .	81
5.4	Le processus d'engagement des agents IDEAS . . . . .	83
5.5	Synthèse . . . . .	86
<b>6</b>	<b>Évaluation</b> . . . . .	<b>89</b>
6.1	Objectifs . . . . .	89
6.2	Description du cadre expérimental . . . . .	90
6.3	Intérêt et pertinence du modèle . . . . .	92
6.3.1	Interférences négatives et positives . . . . .	92
6.3.2	Nombre d'agents partageant une même identité sociale . . . . .	94
6.4	Coût de l'exploitation du modèle . . . . .	96
6.4.1	Comparaison de la qualité de service . . . . .	96
6.4.2	Temps d'exécution . . . . .	97
6.4.3	Évaluation de la précision de l'algorithme de reconnaissance d'intention . . . . .	97
6.5	Synthèse . . . . .	98

<b>Conclusion</b>	<b>103</b>
<b>Annexe A</b>	<b>107</b>
<b>Bibliographie</b>	<b>126</b>

## Table des figures

3.1	Évolution du nombre des articles et dissertations sur le <i>soi</i> et l' <i>identité</i> . . . . .	28
3.2	Une taxonomie des points de vue sur l'identité personnelle (traduite de [Gasser and Stefan, 2012]). . . . .	30
3.3	Composants de l'identité personnelle selon [Layder, 2004]. . . . .	33
3.4	Un modèle de formation de l'identité sociale [Postmes et al., 2005]. . . . .	39
3.5	Modèle d'identité . . . . .	41
3.6	Modèle d'identité de deux acteurs en interaction [Burke, 1997] . . . . .	42
3.7	Classification des disciplines et théories des architectures d'agents. . . . .	43
3.8	Synthèse des éléments utilisés dans les architectures d'agents et leurs analogues dans la théorie de l'identité. . . . .	44
4.1	Cycle de raisonnement simplifié d'un modèle BDI. . . . .	50
4.2	Un modèle abstrait d'IDEAS . . . . .	55
5.1	L'architecture logicielle de l'avatar . . . . .	69
5.2	Illustration d'un scénario de viticulture . . . . .	71
5.3	La librairie de plan utilisée dans l'étude de cas. . . . .	72
5.4	Les classes et les propriétés de l'ontologie qui encode la librairie de plans. . . . .	75
5.5	Adaptation du protocole <i>Extended CNP</i> proposé par Akinine et al. [2004] . . . . .	84

---

6.1	Comparaison d'interférences négatives entre un nombre croissant d'actions dans des équipes 3-robustes effectuant des tâches aléatoires. . . . .	93
6.2	La variation de la somme des utilités des actions composantes entre les coalitions formées avant et après le processus d'engagement. . . . .	96
6.3	Le temps de résolution pour le processus d'engagement par rapport au nombre d'actions composantes pour une équipe 5-robuste. . . . .	97
6.4	L'évolution de la précision de l'algorithme de reconnaissance d'intention.	99



# Liste des tableaux

2.1	Classification des situations d'interaction [Ferber, 1995] . . . . .	17
2.2	Classification des types de conflits . . . . .	20
6.1	Tâches aléatoires décomposées en actions composantes. . . . .	95



# Introduction

Depuis leur origine, les technologies de l'information et de la communication ont connu un développement impressionnant et des progrès technologiques rapides. Les exemples les plus marquants sont la miniaturisation progressive des circuits intégrés et les augmentations continues de la puissance de calcul et de la bande passante des réseaux. Au cours des dernières années, cette progression a conduit à une situation où toutes les performances et la bande passante nécessaires aux développeurs sont plus ou moins disponibles, ce qui signifie que la technologie de l'information peut maintenant être utilisée dans tous les domaines de notre vie. En effet, la prolifération des objets connectés et la meilleure couverture du réseau mobile ont étendu la portée et l'interconnectivité des logiciels à des dispositifs mobiles et des cartes électroniques. Cette tendance conduit à l'émergence de systèmes ouverts, flexibles et interactifs qui connectent de façon transparente le monde physique avec le monde virtuel. De plus, avec l'intégration des données et des services librement accessibles sur le Web, les systèmes et les objets connectés sont de plus en plus transformés en services intégrés directement utilisables par les utilisateurs.

## 1.1 Vers l'Internet des objets sociaux

Les environnements dans lesquels ces systèmes sont plongés connaissent l'interaction de nombreux objets distribués sur plusieurs réseaux. Dans ce monde de réseaux ouverts, les systèmes et les services sont utilisés dans une grande variété de processus et de contextes socio-techniques non prédéterminés. Les objets fournissant des services sont mis en réseau et peuvent prendre des décisions ou réaliser des actions de manière autonome et adaptée à des contextes différents. Ces objets ont une capacité de communication pour interagir et disposent des capacités d'action et de perception pour rendre

l'environnement plus convenable à ses occupants. D'autre part, leur capacité à coopérer et à négocier permet aux objets de ces systèmes de fournir des services composites et de formuler des stratégies de résolution de problèmes de manière distribuée et coordonnée. Les mécanismes sous-jacents sont typiquement étudiés dans le domaine des systèmes multi-agents. Ainsi, comme le suggère le lancement du premier *International Workshop on the Internet of Agents (IoA)* en 2016, des communautés de chercheurs commencent à s'intéresser à cette confluence de méthodes et modèles de conception d'agents intelligents capables d'agir collectivement, et de techniques et protocoles d'interopérabilité des objets physiques représentés par des agents logiciels. Cette tendance est renforcée par l'apparition de plusieurs termes montrant l'intérêt croissant pour l'utilisation d'approches sociales dans des paradigmes comme l'Internet des Objets, et le Web des Objets. Des termes comme *Social Web of Things* [Ciortea et al., 2013], *Social Internet of Things* [Atzori et al., 2012], *Internet of Social Things* [Nansen et al., 2014], *Smart-its* [Holmquist et al., 2001], *Everything is alive* [Eguchi et al., 2013], *Cyber Physical Social Systems* [Hussein et al., 2015], *Collective Cyber-Physical Systems* [Jamont, 2016], *Wisdom Web of Things* [Zhong et al., 2013] ont prospéré en essayant de rendre les objets connectés plus capables d'établir des interactions sociales entre eux et avec les utilisateurs, depuis Internet.

## 1.2 De la coexistence des agents vers la coopération des SMA

Les approches multi-agents sont de plus en plus utilisées dans la conception d'applications reposant sur l'utilisation d'objets connectés (par exemple, l'Internet des objets [Jamont et al., 2014, Mrissa et al., 2015, Khalfi et al., 2014], l'intelligence ambiante [Spanoudakis and Moraitis, 2015], et les systèmes cyber-physiques [Jamont and Ocello, 2015]). Le développement massif de ces objets amène ainsi les agents de différents SMA à coexister dans un même environnement. L'environnement d'un SMA n'appartient donc plus à lui seul, mais peut être partiellement ou complètement partagé avec d'autres systèmes.

Dans la plupart des applications multi-agents, les agents sont conçus pour interagir avec les agents du même système. Leurs interactions sont menées dans des cadres clairement prédéfinis lors de leur conception. Lorsque des SMA coexistent, ils peuvent être dans un état d'indépendance, c'est-à-dire que les actions de leurs agents sont disjointes et chacun cherche à satisfaire ses objectifs individuels et à servir l'objectif global de son système, indépendamment des agents des autres SMA.

Les *interactions sociales* entre SMA commencent dès qu'il y a une dépendance entre les actions de leurs agents. Même si ces derniers ne partagent pas nécessairement les mêmes modèles et architectures, ils sont quand même capables d'échanger et comprendre des messages, car ils sont de plus en plus interopérables du fait des normes et des standards, au niveau technique (interfaces et protocoles, comme ceux de 6LowPAN, Wifi, CoAP), syntaxique (codage et formats de données, comme JSON-LD) et sémantique (ontologies, comme SSN (Semantic Sensor Network)). Du fait de son autonomie, lorsqu'un agent n'a pas les ressources et les compétences nécessaires pour atteindre seul ses objectifs, il cherche à coopérer avec des agents de son voisinage, qu'il soient membres du même SMA ou non. La notion d'appartenance d'un agent à un SMA n'est d'ailleurs souvent pas explicite.

### 1.3 Problématique

Les agents déployés dans des contextes multi-SMA sont caractérisés par la variété des concepteurs : tous les collectifs d'agents coexistants ne partagent pas les mêmes états mentaux et attitudes. Cette hétérogénéité induit des différences dans les comportements et les intentions des agents qui affectent les relations entre les SMA coexistants. De plus, l'ouverture des systèmes fait que la nature des interactions entre les SMA est déterminée par chaque agent et non par les concepteurs des systèmes, puisque les agents sont amenés à former des coalitions spontanées, et à s'engager dans des coopérations avec des coéquipiers précédemment inconnus. Par conséquent, l'efficacité de chaque système est fortement influencée par la nature des interactions sociales *inter-SMA* des agents.

La cohabitation d'agents de systèmes différents, qui n'ont pas nécessairement été prévus pour interagir entre eux, les amène cependant à se solliciter. Dans un contexte où les interactions ont lieu entre les agents d'un même système, elles sont prédéfinies lors de la conception de ce système. Toutes les actions et tâches possibles sont déjà définies dans les programmes des agents, et la plupart des situations conflictuelles sont déjà spécifiées par le développeur [Zatelli et al., 2016]. Par contre dans un contexte multi-SMA, la coopération d'agents avec des coéquipiers inconnus [Stone et al., 2010] et la formation de coalitions d'agents appartenant à différents systèmes peuvent soulever des problèmes liés à la poursuite d'objectifs incompatibles. Un agent peut alors participer, sans qu'il en soit conscient, à la réalisation d'objectifs incompatibles avec les siens ou avec ceux de son SMA.

## La question globale de recherche

*Comment un agent peut s'engager dans une coopération avec un autre collectif d'agents en préservant ses propres intérêts individuels et collectifs ?*

De par sa définition, un agent est une entité proactive et bienveillante, c'est-à-dire possédant un ensemble d'objectifs qu'il souhaite satisfaire en fonction de ses ressources, de ses compétences, de ses perceptions, de ses représentations et de ses interactions, tout en étant coopératif. Lorsqu'un agent n'a pas les ressources et les compétences nécessaires pour atteindre ses objectifs individuellement, il peut coopérer avec d'autres agents pour accomplir certaines tâches. Ces agents peuvent être membres du même SMA ou appartenir à d'autres SMA coexistants dans le même environnement.

Quand un agent s'engage à participer à une tâche collective ou à résoudre un problème collectif, il cherche généralement à éviter les *situations non-coopératives* [Georgé et al., 2011]. Il faut qu'il s'assure que cette participation ne l'empêchera pas de réaliser ses autres missions. Pour pouvoir détecter les situations non coopératives, il considère (i) ses objectifs individuels et collectifs, (ii) ses engagements individuels et collectifs, et (iii) les motivations derrière la sollicitation de l'agent qui a initié la demande de coopération.

La question globale est divisée en sous-questions plus spécifiques qui sont introduites dans la suite. Pour chaque sous-question, nous introduisons notre motivation et la démarche pour aborder le problème.

**Question 1** *Comment représenter les intérêts individuels et collectifs qu'un agent doit préserver ?*

Étant potentiellement situé dans un environnement multi-SMA, un agent doit constituer une représentation de ce qu'il doit préserver avant de s'engager à exécuter une action ou établir une nouvelle coopération. La définition de cette représentation permettra à un agent de construire une représentation des autres agents afin de raisonner sur leurs sollicitations et de définir le cadre d'interaction avec eux.

**Question 2** *Comment un agent peut raisonner sur les sollicitations des autres collectifs ?*

Lorsque il participe à une coalition, un agent a une connaissance limitée ou inexistante des autres membres de la coalition et de la tâche qui a nécessité sa formation. Plus généralement, un agent n'a pas de connaissances sur la fonction globale de son système (qui est, dans certains cas, émergente) [Capera et al., 2003]. Par exemple, un agent contrôlant l'ouverture d'une fenêtre ne saura pas nécessairement si une demande d'ouverture/fer-

meture de la fenêtre est faite pour réaliser une tâche de régulation de température ou une tâche de sécurité, et il ne connaîtra pas nécessairement ses coéquipiers dans chacune de ces tâches. Avec une connaissance si limitée, lorsqu'il est sollicité pour rejoindre une autre coalition, un agent ne sera donc pas en mesure d'évaluer les interférences potentielles de cette demande avec les différents objectifs qu'il poursuit. Un agent doit donc appréhender l'implication des objectifs individuels et dérivés des coalitions dans lesquelles l'agent sollicité est impliqué.

## 1.4 Contributions

Pour répondre à la première question, qui s'intéresse à la représentation des intérêts d'un agent, nous plaçons tout d'abord la thèse dans le contexte des environnements *multi-SMA*. Une hypothèse que nous voudrions avancer est que les SMA sont de plus en plus amenés à partager le même environnement physique. Cette hypothèse est couramment utilisée dans le contexte du Web des Objets qui a comme principale motivation de permettre l'interopérabilité technique, syntaxique, et sémantique entre des plateformes et des objets physiques dans l'Internet des Objets. Ensuite, nous motivons l'intérêt de définir la façon dont un agent se considère, comment il considère les autres, et comment il se rapporte ou se comporte avec eux. Notre conviction est que la *théorie de l'identité* fournit un cadre fondamental pour maintenir des descriptions des coalitions auxquelles un agent appartient, d'avoir des représentations des autres agents, et par conséquent d'intégrer des processus de construction de son identité personnelle et de ses identités sociales.

Pour répondre à la deuxième question, relative à la prise en compte des implications des objectifs des autres lors de la réponse à leurs sollicitations, nous partons du constat qu'un agent ne peut souvent pas connaître les états mentaux des autres agents, il ne peut que déduire leur existence sur la base de ce qui lui est observable, à savoir leurs actions, interactions, liens sociaux, et sollicitations. Nous nous intéressons en fait à un concept qui est à la base de la communication humaine, appelé l'*intersubjectivité*. Cette intersubjectivité - entre deux ou plusieurs agents en interaction - consiste à prendre en considération la reconnaissance de la signification intentionnelle des actions des autres agents lors de l'interaction avec eux. Nous sommes convaincus que ce mécanisme permettra à un agent de construire une représentation des autres agents, ce qui fait partie intégrante de son identité. Nous traduisons ce mécanisme par un système de reconnaissance d'intention multi-agent.

## 1.5 Structure du manuscrit

Dans le Chapitre 2, nous définissons la coopération entre des agents dans un environnement multi-SMA. Nous commençons par rappeler les définitions des caractéristiques de la coopération, en particulier, l'intentionnalité, l'engagement, et l'ajustement des objectifs. La suite de ce chapitre est consacrée à l'interférence dans le cadre des situations d'interactions coopératives, qui est une conséquence de la coexistence des SMA au sein d'un même environnement physique. Nous présentons ensuite un des défis qui nous intéresse le plus dans la coopération, qui est la gestion des *interférences combinées*. Celles-ci affectent à la fois négativement et positivement la poursuite des objectifs d'un agent. Nous répondons à ce défi dans les chapitres suivants.

L'objectif du Chapitre 3 est de proposer un cadre théorique pour répondre à la première question. Nous nous focalisons sur la théorie de l'identité, en particulier, l'identité personnelle, l'identité sociale et l'identité collective, à travers un état de l'art des travaux faits dans le domaine de la philosophie, la psychologie, la sociologie, et la psychologie sociale. Le choix de ce cadre théorique se justifie par le fait que la théorie de l'identité s'intéresse à étudier la dynamique intra-personnelle, inter-personnelle et inter-groupe afin de définir le comportement des individus dans des groupes. Nous terminons ce chapitre en établissant le lien entre (i) des notions connues dans le domaine des SMA et (ii) les éléments qui définissent les identités personnelle et sociale.

Dans le Chapitre 4, nous proposons un modèle d'agent basé sur la théorie de l'identité, utilisant l'intersubjectivité comme mécanisme qu'un agent utilise pour construire et maintenir une représentation des agents cohabitants dans le même environnement physique. Le modèle que nous proposons est étendu du modèle BDI en raison de sa compatibilité avec notre cadre théorique. Dans la suite de ce chapitre, nous répondons à la deuxième question en proposant des processus de décision et de reconnaissance d'intention qui sont modélisés comme des problèmes de satisfaction et d'optimisation de contraintes.

Le Chapitre 5 sera consacré à la présentation des choix techniques faits pour l'implémentation de notre contribution de thèse. Nous commençons par présenter le projet ANR ASAWoO, le contexte applicatif de cette thèse, notamment la notion d'*avatar* qui est une représentation virtuelle d'un objet physique sur le Web. Ensuite, nous décrivons une étude de cas à base d'un scénario de viticulture, et nous montrons l'apport que nous proposons par rapport au cycle de vie d'un agent BDI, en réutilisant des composants de l'infrastructure ASAWoO et en les augmentant pour permettre à l'avatar d'être une en-



tité capable de raisonner sur son identité avant de coopérer avec les autres avatars. Dans la suite de ce chapitre, nous définissons la motivation derrière le choix de l'algorithme de reconnaissance d'intention multi-agent et le choix du mécanisme de négociation et d'allocation de tâches entre avatars.

Dans le chapitre 6, nous introduisons plusieurs métriques pour mesurer la qualité des résultats obtenus par notre approche, puis nous analysons les résultats obtenus lors des expérimentations.



# La coopération et ses potentielles interférences

---

<b>2.1</b>	<b>Spectre des interactions sociales coopératives</b>	<b>10</b>
2.1.1	Communication par environnement	10
2.1.2	Coordination non concertée	11
2.1.3	Communication explicite	12
2.1.4	Coordination planifiée	12
2.1.5	Coopération	12
<b>2.2</b>	<b>Caractéristiques de la coopération</b>	<b>13</b>
2.2.1	Classification des types d'interactions sociales coopératives	14
2.2.2	L'intentionnalité	15
2.2.3	L'engagement	15
2.2.4	La compatibilité des objectifs	16
<b>2.3</b>	<b>L'interférence dans le cadre de la coopération</b>	<b>17</b>
2.3.1	L'interférence positive	18
2.3.2	L'interférence négative	19
2.3.3	Interférence combinée	22
<b>2.4</b>	<b>Synthèse</b>	<b>23</b>

---

Le déploiement massif d'objets physiques dans notre vie quotidienne amène les agents de différents SMA à coexister dans le même environnement. Dans ce chapitre, nous présentons le spectre des interactions sociales dans lesquelles peuvent être impliqués ces agents. Notre objectif est de définir les caractéristiques de base d'une situation d'interaction coopérative. Nous nous baserons sur ces caractéristiques pour permettre à un

agent de mieux coopérer avec d'autres agents appartenant à des systèmes différents cohabitants dans le même environnement physique. Une telle coexistence peut montrer un large éventail de comportements qui peuvent correspondre à l'altruisme, au mutualisme, à l'antagonisme, à la coopération, au parasitisme, etc [Cesta et al., 1996]. Cette coexistence implique forcément des interférences que nous classifions dans la dernière partie de ce chapitre.

## 2.1 Spectre des interactions sociales coopératives

La coexistence au sein d'un même espace physique (ou virtuel) est la condition minimale de l'établissement des interactions entre des agents.

**Définition 2.1** (Interaction). *Une interaction est une mise en relation dynamique de deux ou plusieurs agents par le biais d'un ensemble d'actions réciproques. Les interactions s'expriment à partir d'une série d'actions dont les conséquences exercent en retour une influence sur le comportement futur des agents* [Ferber, 1999].

Dans des situations coopératives, les interactions des agents sont déterminées dans le cadre d'une réalisation efficace des objectifs. Nous identifions différents mécanismes d'interaction qui définissent la relation entre les agents cohabitant le même environnement physique.

### 2.1.1 Communication par environnement

La communication via l'environnement est la forme d'interaction la plus simple et la plus limitée. Elle se produit lorsque l'environnement lui-même est le moyen de communication (une mémoire partagée, par exemple), et qu'il n'y a pas de communication ou d'interaction directe entre les agents. Par exemple, un agent contrôlant le thermostat a été déployé dans une maison intelligente pendant plusieurs mois. Ensuite, il détecte la présence d'autres agents, tels qu'un agent contrôleur de l'ouverture/fermeture de la fenêtre, un agent contrôleur de la télévision, et un agent contrôleur des lampes, ces trois derniers objets physiques sont installés dans le salon. Dans un premier temps l'agent thermostat communique avec le reste des agents de son voisinage en alimentant avec des mesures et son état (en marche/éteint) une *base de données contextuelles* commune entre les objets de la maison.

Une fois que les agents ont détecté la présence d'autres agents dans l'environnement, il est pertinent de surveiller leurs actions afin d'en tirer des hypothèses pour pouvoir

construire une présentation sur leurs états mentaux [Sindlar et al., 2008]. Cette représentation permettra à chaque agent d'identifier les activités et les répercussions des actions de ses voisins, qui pourront être des futurs coéquipiers avec lesquels il peut échanger des services.

**Définition 2.2** (État mental). L'état mental est une agrégation des croyances, désirs, intentions des agents. Ce sont les concepts qui font qu'un agent raisonne avant d'agir, et cherche à maximiser les bénéfices ou avantages et à réduire les coûts ou inconvénients. Les croyances représentent l'ensemble des connaissances de l'agent sur l'environnement et les autres agents. Ces connaissances peuvent être incomplètes, incorrectes, obsolètes ou incertaines. Les désirs représentent les motivations de l'agent, c'est-à-dire l'ensemble des objectifs et des situations que l'agent veut atteindre. Les intentions représentent l'état délibératif de l'agent. Il s'agit des désirs auxquels l'agent s'est engagé à réaliser.

### 2.1.2 Coordination non concertée

**Définition 2.3** (Coordination). La coordination est un agencement non-accidentel de l'activité de plusieurs systèmes couplés dans le présent ou qui étaient couplés dans le passé, entre eux ou avec un troisième système. Cet agencement est en plus de ce qui est attendu de leur comportement normal dans l'absence de tels couplages [De Jaegher et al., 2010].

Dans la coordination non concertée, un comportement coordonné se produit en raison de couplages perception-action qui font que plusieurs individus agissent d'une façon coordonnée ; cette coordination est indépendante de tout plan conjoint ou connaissance commune (qui peuvent être absents). La coordination peut se produire pour échapper (ou réduire) les interférences négatives entre des agents coexistants, comme elle peut aussi se produire pour adapter le comportement d'un agent tel qu'il profite des circonstances qui conditionnent l'environnement partagé. Dans le cas d'indépendance entre des agents, ils ont seulement besoin d'une faible coordination, c'est-à-dire s'adapter aux changements de la dimension physique de l'environnement. Par exemple, l'agent contrôleur de la fenêtre peut coordonner ses actions avec l'agent contrôleur du thermostat en arrêtant le thermostat lorsque la fenêtre est ouverte d'une manière non planifiée.

### 2.1.3 Communication explicite

En fonctionnant dans un environnement partagé avec d'autres systèmes, les agents ont besoin de coordonner leurs activités et coopérer quand ils peuvent en tirer bénéfice mutuellement. Lorsque plusieurs SMA cohabitent dans un même environnement, leurs agents ne sont pas conçus avec des connaissances pré-compilées sur les états mentaux et les capacités des autres. Pour ce faire, ils ont besoin d'échanger des informations, établir des relations interpersonnelles, expliciter leurs intentions, influencer les autres, et partager des informations sur des activités collectives. Ce type de communication est intentionnel par échange de messages directs lorsque les destinataires des messages sont connus ou diffusés lorsqu'ils ne le sont pas. Par exemple, l'agent contrôleur du thermostat peut explicitement demander à l'agent contrôleur de la fenêtre si cette dernière est ouverte ou fermée, pour pouvoir allumer ou éteindre le thermostat en conséquence.

### 2.1.4 Coordination planifiée

À partir du moment où les actions des agents s'exécutent dans un cadre coopératif, le besoin de coordination devient plus fort. Dans un premier temps, elle se produit pour éviter les redondances entre les actions des différents agents en coopération, et ensuite est envisagée pour optimiser l'exécution de ces actions pour atteindre le but global (objectif derrière la coopération) d'une manière efficace. Les situations de compétition ou de conflit exigent aussi un certain degré de coordination. Un agent situé dans un environnement hostile ne peut pas ignorer les actions des autres agents et doit coordonner ses actions convenablement. En coordination planifiée, le comportement des agents est déterminé par des représentations qui spécifient les résultats attendus et la contribution de l'agent dans la réalisation de l'objectif commun. Un agent peut prendre en considération les motivations, les pensées ou les perspectives des autres agents, ou simplement attendre qu'une action particulière se produise. Par exemple, l'agent thermostat et l'agent contrôleur de la fenêtre peuvent coordonner les états des objets physiques qu'ils contrôlent pour faire en sorte qu'il n'y ait pas une situation où la fenêtre est ouverte en même temps que le thermostat est allumé.

### 2.1.5 Coopération

La coopération peut être généralement définie comme une attitude d'un agent qui cherche à aider sans entraver les autres agents [Georgé et al., 2011]. C'est une est une forme d'in-

teraction, généralement basée sur la communication. Certains types de comportements coopératifs nécessitent de la communication explicite. Plus précisément, tous les comportements coopératifs qui nécessitent la négociation entre les agents dépendent de la communication explicite afin d'attribuer des tâches particulières aux agents participants. De manière analogue à la communication, la coopération explicite est définie comme un ensemble d'interactions qui impliquent l'échange d'informations ou d'effectuer des actions afin de bénéficier un autre agent. D'autre part, la coopération implicite consiste en des actions réalisées dans le cadre de la satisfaction des objectifs d'un agent, mais ont des effets sur l'environnement qui aident d'autres agents à atteindre leurs objectifs. Par exemple, l'agent contrôleur du thermostat et l'agent contrôleur de la fenêtre peuvent coopérer dans le cadre d'un plan ayant comme objectif la régulation de la température du salon en fonction des préférences de l'occupant de cet espace.

## 2.2 Caractéristiques de la coopération

La coopération peut être étudiée à partir de nombreuses perspectives différentes. Les travaux dans les *systèmes multi-agents en réseau* [Olfati-Saber et al., 2007] considèrent que la coopération consiste à donner son consentement à fournir son état et suivre un protocole commun qui sert l'objectif d'un groupe [Dias et al., 2014]. Par exemple, dans les applications de transport intelligent, lorsqu'un agent décide de ne pas coopérer avec ses coéquipiers et refuse de changer son état, la tâche globale ne peut pas être exécutée même si le reste de l'équipe accepte de coopérer. Dans les travaux de planification multi-agent [Torreño et al., 2017, De Weerd et al., 2005], les tâches coopératives impliquent des agents capables de planifier partageant le même environnement. Ils doivent se mettre d'accord sur des plans conjoints pour atteindre un ensemble d'objectifs globaux. Les actions à exécuter sont distribuées sur l'ensemble des agents, et la différence en terme de capacités, d'interaction avec l'environnement (action/perception), et de consommation d'énergie donne lieu à des agents "spécialisés" possédant des connaissances et des capacités différentes des autres. Cette spécialisation implique le besoin de coopération, car les agents sont obligés de s'entraider pour résoudre les différents tâches. Il existe d'autres perspectives selon lesquels on peut définir la coopération, mais toutes partagent une caractéristique principale qui est la présence d'un objectif partagé, objet de l'établissement de l'interaction sociale coopérative. Dans le contexte de cette thèse, nous nous intéressons à d'autres caractéristiques supplémentaires pour étudier comment permettre à un agent de mieux coopérer avec d'autres agents de différents SMA coexistants.

### 2.2.1 Classification des types d'interactions sociales coopératives

Deux classes d'approches majeures sont identifiées lors de la classification des interactions coopératives [Conte and Castelfranchi, 1992]. Les types d'interaction présentées dans la section précédente correspondent à la classe d'approches *top-down*, c'est-à-dire, que les agents qui composent un SMA ont un objectif commun, et agissent dans le cadre d'une tâche collective tels qu'ils sont conçus pour agir ensemble, avec des structures et schémas d'interaction préétablis qui guident leurs interactions dans la satisfaction de leur tâche collective. Les approches qui nous intéressent le plus dans le contexte de cette thèse sont les approches *bottom-up*. Dans cette classe d'approches, les agents n'ont pas a priori d'objectif commun. Leurs interactions sociales sont motivées par les efforts que chacun déploie pour atteindre ses propres objectifs. Les structures et schémas d'interaction ne sont pas pré-établis mais ils proviennent des interactions entre les agents et de la manière dont ils tiennent compte des interférences des autres dans la réalisation de leurs objectifs. Dans cette dernière classe, nous nous intéressons particulièrement aux coalitions, qui sont des structures sociales créées dynamiquement par des agents partageant un même objectif.

**Définition 2.4** (Coalition). Une coalition est définie comme un groupe d'agents qui ont décidé de coopérer pour accomplir une tâche commune. Une coalition peut travailler sur une seule tâche à la fois et les agents peuvent être membres de plus d'une coalition [Shehory and Kraus, 1998].

Deux grandes classes de formation de coalition existent. Lorsque chaque agent cherche à maximiser ses gains, on dit qu'il s'agit d'une approche de *formation de coalition fondée sur l'utilité*. La présence d'autres agents limite souvent les possibilités de chacun. Les agents développent donc des conventions et des contraintes qui leur permettent de rendre cohérent le comportement de l'ensemble afin de minimiser les pertes individuelles. Dans cette thèse, nous adoptons une approche de *formation de coalition fondée sur la complémentarité*. Dans cette approche, les agents ont des capacités complémentaires qui peuvent permettre la satisfaction des objectifs individuels de chacun. Ainsi, la présence des autres agents augmente les capacités de chacun à satisfaire des objectifs.

**Définition 2.5** (Coopération). On dit que des agents coopèrent s'ils s'engagent dans une action commune après avoir identifié et adopté un but commun. [Ferber, 1995]

Nous considérons que la nature d'une interaction sociale coopérative est co-déterminée par :



- l'intention et la volonté des agents de s'impliquer dans l'interaction,
- les stratégies d'engagement qu'ils suivent qui sont pertinentes à la nature de ces interactions,
- la compatibilité des objectifs poursuivis par les agents.

### 2.2.2 L'intentionnalité

L'intentionnalité est une condition nécessaire à la coopération. Cela signifie que deux agents ne peuvent pas être considérés comme ayant coopéré s'ils ne le font pas volontairement, même si leurs actions s'exécutent pour atteindre fortuitement le même objectif. De même, pour qu'un agent  $a_1$  puisse maintenir une interaction sociale avec un agent  $a_2$ , il faut qu'il reconnaisse les actions de  $a_2$  comme intentionnelles et motivées par des objectifs. Un agent peut intentionnellement exécuter une action, mais peut ne pas connaître l'intention de tous ses effets. Il y a donc deux notions d'intentionnalité, une intentionnalité de résultat (visant l'effet d'une action) et une intentionnalité d'acte (visant l'action, elle-même).

**Définition 2.6 (Intention).** L'intention est considérée comme l'engagement à réaliser un objectif (selon [Cohen and Levesque, 1990]) ou un plan partiel (selon [Rao and Georgeff, 1991]) en fonction de ses perceptions et de ses connaissances du monde.

Dans un contexte multi-agent, une équipe d'agents a une intention conjointe (*joint intention* [Cohen and Levesque, 1991]) lorsque ses membres s'engagent à agir conjointement pour atteindre un but de l'équipe. Cette notion d'intention conjointe implique une croyance mutuelle sur ce que doivent faire les agents. Lorsque les agents  $a_1$  et  $a_2$  s'engagent à réaliser une tâche  $t$  coopérativement :

- $a_1$  s'engage à réaliser sa partie de  $t$ ,
- $a_1$  suppose que  $a_2$  fera sa partie de  $t$ ,
- $a_1$  suppose que  $a_2$  suppose que  $a_1$  fera sa partie de  $t$ .

### 2.2.3 L'engagement

La coopération fonctionne grâce à la délégation d'une tâche (ou sous-tâche) par un agent à un autre agent, par l'intermédiaire d'une certaine demande (offre, proposition, annonce, etc.) répondant à un certain engagement (enchère, aide, contrat, adoption, etc.) [Castelfranchi and Falcone, 2002]. Ainsi, l'engagement provient de la correspondance d'une forte délégation d'un agent et d'une forte adoption d'un autre agent. Il s'agit d'une

relation réciproque entre les agents qui constitue la structure des groupes et des organisations sociales [Baldoni et al., 2018]. Lorsque des agents coopèrent, chacun suppose que l'autre est capable et a l'intention d'exécuter l'action demandée. L'engagement devient donc une garantie implicite ou formelle à réaliser un objectif commun identifié. De plus, l'engagement ne peut pas être accepté pour réaliser des actions dont l'agent ne connaît pas les résultats. Cela fera en sorte que l'agent ne s'engage pas à exécuter toute action sollicitée, mais seulement à la réalisation de celles qui sont pertinentes pour lui. Lorsqu'un agent s'engage à exécuter une ou plusieurs actions, elles sont ajoutées à ses intentions.

Au niveau de l'agent, l'engagement individuel (ou interne) se réfère à la relation entre l'agent et une action à exécuter. Lorsqu'il décide d'atteindre un objectif, l'agent se prête à réaliser des tâches et à conserver ses intentions tant qu'il croit que l'objectif est possible à atteindre et qu'il n'est pas encore atteint. On identifie trois stratégies d'engagement différentes [Rao and Georgeff, 1991] :

- l'agent s'engage à atteindre un objectif et n'abandonne son intention que lorsqu'il croit qu'elle est atteinte (*blind commitment*). C'est la stratégie d'engagement la plus forte, la plus durable.
- L'agent n'abandonne son intention que s'il croit que son objectif est atteint ou qu'il est impossible à atteindre (*single-minded commitment*).
- L'agent n'abandonne pas ses intentions tant qu'elles sont compatibles avec ses propres buts (*open-minded commitment*).

Un *engagement social* est un concept relationnel qui lie plusieurs agents à une tâche, tel qu'un agent  $a_1$  s'engage à réaliser une tâche  $t$  pour le compte d'un agent  $a_2$ , en prenant en compte son engagement précédent avec un agent  $a_3$  [Castelfranchi, 1995]. L'engagement social établit des droits et des devoirs entre les agents et permet un certain degré de prévisibilité sur leur comportement futur. Non seulement cela fournit une preuve d'intentionnalité ou d'objectif commun, mais cela permet aussi à des agents de remplir leurs parts respectives tout en supposant que les autres feront de même.

#### 2.2.4 La compatibilité des objectifs

Un agent contrôle son état, son comportement, et donc sa propension à coopérer avec un autre agent. Selon [Ferber, 1995], hormis les situations d'*indépendance* pour lesquelles les buts sont compatibles et les ressources et compétences sont suffisantes, les catégories d'interaction sont la coopération ou la compétition.

Buts	Ressources	Compétences	Catégorie d'interaction	Type d'interaction
Compatible	Suffisant	Suffisant	Indifférence	Indépendance
		Insuffisant	Coopération	Collaboration simple
	Insuffisant	Suffisant		Encombrement
		Insuffisant	Collaboration coordonnée	
Incompatible	Suffisant	Suffisant	Antagonisme	Compétition individuelle
		Insuffisant		Compétition collective
	Insuffisant	Suffisant		Conflits individuels sur des ressources
		Insuffisant		Conflits collectifs pour des ressources

TABLE 2.1 : Classification des situations d'interaction [Ferber, 1995]

La *collaboration simple* est l'addition des compétences (sans coordination). L'*encombrement* est la situation où les agents se gênent mutuellement dans l'accomplissement de leurs tâches alors qu'ils n'ont pas besoin les uns des autres. La *collaboration coordonnée* consiste à coordonner les actions des agents pour réaliser une tâche commune. La *compétition individuelle pure* amène les agents à négocier ou se concourir pour atteindre leurs buts lorsque les ressources ne sont pas la source du conflit. La *compétition collective pure* a lieu lorsque les agents n'ont pas les compétences suffisantes et doivent former des coalitions pour atteindre des objectifs. Le *conflit individuel pour des ressources* caractérise les situations où les ressources ne peuvent pas être partagées entre des agents ayant des buts incompatibles. Les *conflits collectifs pour des ressources* abordent la formation de coalitions luttant les unes contre les autres pour accéder à des ressources et ainsi accomplir leurs objectifs.

## 2.3 L'interférence dans le cadre de la coopération

Les interférences entre des agents des systèmes coexistants sont inhérentes lorsque chaque système est déployé pour des tâches globales dédiées. L'interférence se produit lorsqu'un agent est incapable d'atteindre ses objectifs de manière *indépendante*, et dépend donc des actions ou des non-actions des autres agents pour atteindre ses objectifs. Ce point fait référence à l'interdépendance entre les agents des SMA coexistants et apporte une dimension collective au conflit ou à la satisfaction des objectifs. Dans la littérature, on identifie plusieurs types d'interférences [Hoc, 2001], comme l'interférence de précondition, l'interférence d'interaction, l'interférence de contrôle mutuel, et l'interférence de redondance. Nous nous intéressons essentiellement à trois possibilités d'interférence entre des agents de SMA coexistants [Castelfranchi, 1998] : l'interférence positive, l'interférence négative, et l'interférence combinée .

### 2.3.1 L'interférence positive

**Définition 2.7** (Interférence positive). Une interférence est dite *positive* pour un agent lorsque l'exécution des actions d'un autre agent favorise la réalisation ou la préservation de ses propres objectifs. Ceci peut se produire lorsqu'il participe à exécuter des actions faisant partie du plan d'un autre agent, ou lorsqu'il adopte un de ses objectifs.

Une interférence positive peut prendre plusieurs formes :

**Action conjointe** Une action conjointe est une interaction sociale par laquelle deux agents ou plus coordonnent leurs actions dans l'espace et le temps pour provoquer un changement dans l'environnement pour arriver à un résultat commun [Knoblich et al., 2011]. Les premières approches de l'action conjointe trouvent leur origine dans l'intérêt des philosophes pour la nature de l'intentionnalité conjointe. Ces approches considèrent qu'une action conjointe est une action d'un ensemble d'agents ayant des intentions partagées. Ce qui distingue les actions individuelles des actions conjointes, c'est que ces dernières impliquent une intention partagée qui fournit les représentations permettant la planification et la coordination des actions conjointes.

**Facilitation** La facilitation va plus loin que la coordination en ce sens qu'elle ne concerne pas seulement la gestion de l'interférence. Un agent peut faciliter les actions ou la réalisation des objectifs des autres agents par de nombreux moyens, comme l'assistance, par exemple. Le mécanisme antagoniste de la facilitation est le blocage, qui consiste à effectuer des actions visant à empêcher un agent d'atteindre ses objectifs, ou à répondre négativement à ses sollicitations d'assistance. Ce sont les deux mécanismes de base qui sont respectivement utilisés dans les interactions coopératives et conflictuelles.

**Adoption d'objectif** L'adoption d'objectif permet aux agents de coopérer sur les mêmes tâches et activités, créant ainsi une relation de coopération plus intégrée. Les agents altruistes, et véritablement coopératifs peuvent adopter les objectifs des autres agents. Les agents précautionneux reconsidèrent constamment leurs engagements, ne font aucun choix arbitraire, et ont tendance à se méfier des autres agents. Dans l'adoption d'objectifs coopératifs, les agents peuvent adopter un objectif commun, où l'un peut adopter l'objectif de l'autre. L'assistance est un exemple d'adoption d'objectifs. Cependant, pour être considéré comme coopérative, une interaction doit impliquer une intention bilatérale, c'est-à-dire que l'agent destinataire doit être conscient d'être assisté.

### 2.3.2 L'interférence négative

Les interférences négatives entre agents surviennent dans de nombreux cas, et elles sont résolues de différentes stratégies selon leur type. Il y a un très grand intérêt dans les travaux en intelligence artificielle pour les interférences négatives entre des agents, communément nommées, conflits.

**Définition 2.8** (Conflit). Le conflit est défini comme une incompatibilité d'objectifs entre deux agents ou plus dans une interaction, combinée avec des tentatives de contrôle mutuel, des attitudes antagonistes, et des interférences négatives de l'un envers l'autre [Fisher, 2012].

Par exemple, Tessier et al. [2006] classent les conflits en deux classes principales : les conflits physiques et les conflits de connaissances. Les conflits physiques sont des conséquences de conflits externes et de ressources. Les conflits de connaissances (ou conflits épistémiques) se produisent lorsque chaque agent a ses propres informations qui sont différentes des autres agents. Inspiré par les conflits humains, ils identifient cinq stratégies de gestion conflits : le forçage, la soumission, l'ignorance, la négociation, et le consensus.

Le contexte de cette thèse s'applique aux environnements hébergeant des agents appartenant à différents SMA, tels que les applications Web des Objets. Pour cette raison, nous trouvons la taxonomie de Resendes et al. [2014] plus adaptée à ce contexte, car elle s'intéresse aux conflits dans les systèmes d'intelligence ambiante, qui représentent une application parfaite de notre thèse.

#### 2.3.2.1 Classification des types de conflits

Le tableau 2.2 résume les types de conflits qui sont en lien avec le sujet de notre thèse. Cette classification est inspirée de la taxonomie de Resendes et al. [2014] qui distingue les conflits selon quatre dimensions (*source*, *intervenants*, *moment de détection*, et *résolubilité*). Nous ne nous intéressons pas à la dimension *intervenants* car elle classe les conflits selon les préférences des utilisateurs. Les types de conflits dans cette dimension sont résolus par une intervention humaine, alors que dans notre travail, nous focalisons sur les interférences négatives pouvant être résolus par des agents autonomes.

**Source** Les conflits peuvent être classifiés selon leurs *sources* : un conflit de *ressource* se produit lorsque les agents doivent utiliser une même ressource en même temps comme, par exemple, l'ascenseur d'un bâtiment ; un conflit d'*application* se produit

Dimension de classification	Types de conflits	Références
Source	Ressource	[Thangarajah et al., 2002]
	Application	[Yagita et al., 2015]
	Contexte	[Camacho et al., 2014]
	Politique	[Maternaghan and Turner, 2013, Syukur et al., 2005]
Niveau structurel	Croyance	[Castelfranchi, 2015]
	But	[van Riemsdijk et al., 2009]
	Plan	
	Action	[Zatelli et al., 2016]
Moment de détection	Lors de la conception	[Dastani and van der Torre, 2004]
	A priori (pendant l'exécution)	[Tuttles et al., 2007]
	Au moment de son déclenchement (pendant l'exécution)	[Shin and Woo, 2009]
	Après son déclenchement (pendant l'exécution)	[Sycara, 1993]
Résolubilité	Évitement du conflit	[Thangarajah et al., 2002]
	Résolution du conflit	[Vasconcelos et al., 2009]
	Reconnaissance de l'incapacité	
	Reconnaissance de l'occurrence	[Resendes et al., 2014]
Niveau conceptuel	Individuel	[Zatelli et al., 2016]
	Social	[Martinho et al., 2015]
	Collectif	[Khalif et al., 2017]

TABLE 2.2 : Classification des types de conflits

lorsque des applications sont en concurrence sur l'état de l'environnement, e.g., *HomeAutomationSystem* et *EnergyHomeSystem* contrôlant la température d'une maison ; un conflit de *politique* se produit lorsque des systèmes coexistants poursuivent des politiques conflictuelles dans un contexte donné, e.g., lorsque le confort des habitants est en conflit avec la politique de sécurité ou d'énergie du bâtiment.

**Moment de détection** Relativement au *moment de détection*, un conflit peut être détecté *a priori*, c'est-à-dire qu'il est anticipé avant qu'il se produise en utilisant le raisonnement abductif ; un conflit peut être détecté *au moment de son déclenchement*, c'est-à-dire pendant qu'il se produit ; un conflit peut être détecté *a posteriori*, c'est-à-dire après qu'il se produit.

**Résolubilité** La *Résolubilité* est aussi un moyen de distinguer les conflits : l'*évitement de conflit* est utilisé comme une technique de gestion des conflits où les agents résolvent le conflit avant son apparition ; la *résolution de conflit* est la technique la plus reconnue dans la gestion des conflits. Dans d'autres conditions contre-productives, un agent peut gérer les conflits en *reconnaissant son incapacité* à les résoudre, ou *en reconnaissant son occurrence* sans tenter de le détecter ou de le résoudre.

Pour prendre en considération le contexte d'agents de systèmes coexistants, nous étendons cette classification en considérant les dimensions *niveau structurel* et *niveau conceptuel*.

**Niveau structurel** En ce qui concerne la dimension *structurelle*, les conflits peuvent provenir des *buts* incompatibles, e.g. satisfaire des préférences logiquement incohérentes de différents utilisateurs ; ils peuvent être entre des *plans*, e.g. certains

sous-objectifs des plans peuvent être incohérents ; ou ils peuvent être entre des *actions*, e.g. certaines préconditions d'une action d'agent deviennent invalides en raison des postconditions d'une autre action d'un agent.

**Niveau conceptuel** Quant à la dimension *niveau conceptuel*, les conflits *individuels* se produisent lorsque le conflit est interne à l'agent, e.g., la recharge des batteries par rapport à l'exécution d'une patrouille ; Les conflits *sociaux* se produisent entre les agents d'un même système, e.g., lorsqu'un agent est sollicité par un autre agent du même système pour effectuer une action ou une tâche conflictuelle ; Les conflits *collectifs* surviennent entre des agents de différents systèmes tout en poursuivant des tâches *trans-SMA*, e.g., en acceptant des requêtes de deux agents de systèmes différents avec des objectifs incompatibles, un agent est sur le point d'être engagé dans un conflit *collectif*.

### 2.3.2.2 La gestion des conflits

La grande partie des travaux qui s'intéressent à la gestion des conflits se concentre particulièrement sur la résolution [Kuchar and Yang, 2000]. Plusieurs stratégies sont proposées telles que le choix d'objectifs ou plans non conflictuels [Liu and Barber, 1998], l'accord de priorité, la négociation, l'arbitrage [Adler et al., 1998], le vote [Barber et al., 2000], et l'argumentation [Jung and Tambe, 2002]. La sélection d'une stratégie appropriée dépend par exemple de la nature du conflit, du niveau d'autonomie de l'agent, de ses préférences, de la qualité du réseau de communication, du coût, et du temps requis. D'autres approches plus complètes de gestion de conflits existent dans la programmation orientée-agent [Bordini et al., 2009]. Une première approche simple et assez souvent adoptée pour gérer les intentions conflictuelles consiste à permettre au développeur de spécifier explicitement les plans comme étant atomiques (c'est-à-dire ne peuvent pas être exécutés simultanément avec d'autres plans), comme c'est supporté dans des langages de programmation d'agents. Par exemple, dans Jason [Bordini et al., 2007], lorsqu'une intention contient un plan atomique (devenant une intention atomique), son exécution désactive le mécanisme d'ordonnancement circulaire du cycle de raisonnement, et l'intention atomique reste la seule en exécution jusqu'à la fin du plan atomique. Dans d'autres langages comme AgentSpeak(RT) [Rao, 1996], les plans non-atomiques peuvent être exécutés simultanément avec un plan atomique. Par exemple, si les intentions  $i_1$  et  $i_2$  sont atomiques et que  $i_3$  n'est pas atomique, alors  $i_3$  peut être exécutée simultanément avec  $i_1$  ou  $i_2$ . Dans d'autres travaux comme [Dastani and van der Torre, 2004, Pokahr et al., 2005], le développeur doit explicitement spécifier quels objectifs ou plans sont

en conflit. Les conflits sont détectés automatiquement si les objectifs sont logiquement incompatibles ( $\neg g$  et  $g$ ). Un mécanisme similaire pour détecter les conflits entre les objectifs logiquement incompatibles est implémenté dans GOAL [Hindriks et al., 2000]. Les auteurs soulignent également que bien que les objectifs  $\neg g$  et  $g$  ne peuvent pas être atteints en même temps, ils peuvent parfaitement coexister dans l'ensemble des intentions d'un agent, une fois que ces objectifs peuvent être atteints à des moments différents. Un autre moyen de gérer les conflits consiste à envisager des plans alternatifs pour atteindre les mêmes objectifs [Shapiro et al., 2012]. Si la première option du plan pour atteindre l'objectif est en conflit avec une autre intention active, l'agent peut essayer une seconde option, et ainsi de suite. Le principal avantage est que l'agent peut poursuivre et atteindre plusieurs objectifs, une fois que l'agent peut toujours chercher un plan non conflictuel pour les atteindre.

Des revues plus approfondies de la littérature, axées sur la systématisation des approches existantes de détection et de résolution automatique des conflits, sont présentées dans [Alshabi et al., 2007, Basheer et al., 2013, Campos et al., 2013, Resendes et al., 2014, Santos et al., 2017]. La plupart des approches examinées proposent des techniques pour traiter les dimensions *source*, *moment de détection* et *résolubilité* des conflits. En outre, beaucoup de travaux partent sur la résolution des conflits dans la dimension *structurelle* dans la programmation orientée agent [van Riemsdijk et al., 2009, Zatelli et al., 2016].

### 2.3.3 Interférence combinée

**Définition 2.9** (Interférence combinée). L'interférence combinée se produit lorsqu'une même action a des effets, à la fois, positifs et négatifs pour un même agent.

L'interférence *combinée* est la plus complexe à appréhender pour un agent (ou à approcher par un concepteur) car les agents ont généralement plusieurs buts de différents types et niveaux d'abstraction. La réalisation d'une action coopérative peut alors être une interférence à la fois négative et positive. Si un but est, par exemple, compatible au niveau individuel (interférence positive), il peut être incompatible au niveau collectif (interférence négative). Ainsi, un agent peut approuver une interférence négative, comme il peut désavouer une interférence positive. Par exemple, un agent contrôlant l'ouverture d'une fenêtre considérerait la demande d'ouverture pour aérer une pièce comme une interférence positive vis-à-vis l'objectif de réguler la température d'une chambre, alors qu'il la considérerait comme négative vis-à-vis l'objectif d'assurer la sécurité de toute la maison.



## 2.4 Synthèse

Dans ce chapitre, nous avons défini le type d'interférence auquel nous nous intéressons dans la suite de ce manuscrit. Cette définition est basée sur la description d'un spectre de différentes situations d'interactions sociales coopératives qu'un agent puisse traverser, et des caractéristiques de la coopération, comme étant la dernière situation de ce spectre. Ensuite, nous avons présenté une classification des conflits (que nous considérons comme interférence négative) qui nous paraît la mieux adaptée au contexte applicatif de notre travail.

Considérant la classification présentée dans le Tableau 2.2, nous nous intéressons aux conflits d'*applications* (c'est-à-dire, des tâches de haut niveau) en supposant que les conflits proviennent des applications *trans-SMA* impliquant des agents de différents systèmes coexistants. Nous considérons les conflits qui se produisent entre les *actions* des agents car un agent n'est pas nécessairement capable d'observer des *buts* ou des *plans* de haut niveau, il ne peut que déduire leur existence sur la base d'actions observables.

De nombreux langages de programmation agent prennent en charge la spécification des situations conflictuelles *lors de la conception*, mais dans notre contexte, l'interaction avec des agents d'autres systèmes nécessite de gérer des situations inattendues lors de l'exécution. Par conséquent, il nous paraît plus pertinent de considérer le type de conflit *a priori*, c'est-à-dire que le conflit est anticipé avant qu'il se produise en utilisant le raisonnement abductif.

Quant à la résolubilité, nous proposons qu'un agent suit, au pire des cas, une stratégie prudente d'*évitement de conflit* (un agent agira prudemment en rejetant les demandes conduisant à une situation d'interférence), et une stratégie proactive de *résolution de conflit* au mieux (c'est-à-dire qu'un agent négociera pour participer à l'accomplissement d'une tâche tout en minimisant les interférences négatives).

La dernière dimension qui distingue le type de conflit auquel nous nous intéressons est le niveau conceptuel. Comme les travaux dans la littérature propose de nombreuses techniques utiles pour la résolution des conflits aux niveaux *agent* et *système*, nous abordons le problème des agents des SMA coexistants tout en gérant les interférences négatives d'un *niveau collectif*.

Pour que les équipes d'agents puissent coopérer dans des SMA ouverts et hétérogènes, ils sont confrontés au défi d'établir et de maintenir des relations bénéfiques avec des systèmes les uns inconnus aux autres. Afin de prendre des décisions concernant la coopération avec d'autres équipes et s'engager à participer à des actions conjointes, les

agents ont besoin d'un moyen d'estimer la probabilité que leurs partenaires potentiels n'interféreront pas négativement avec la réalisation de leurs propres objectifs. Une hypothèse essentielle est que les SMA sont conçus ensemble comme une combinaison, étant donné qu'ils sont susceptibles de coopérer d'une manière synergétique.

Dans le cadre de ces interférences, un agent a intérêt à favoriser les interférences positives, à éviter les négatives, et à gérer le compromis qu'offrent des interférences combinées afin de décider au mieux de son engagement à coopérer avec les agents de son voisinage. Un agent doit avant tout vérifier l'ajustement de l'objectif caché derrière la sollicitation d'une coopération, avec ses propres objectifs (individuels et collectifs). Les objectifs collectifs et les coéquipiers n'étant pas explicitement identifiés lors d'une sollicitation, il est intéressant de définir et de maintenir des descriptions de ces entités. Pour ce faire, nous proposons donc de s'inspirer de la théorie de l'identité.

## L'identité au centre de la coopération

---

3.1	L'identité et son rôle dans le comportement de l'individu . . . . .	26
3.1.1	Le soi et l'identité . . . . .	27
3.2	L'identité personnelle . . . . .	30
3.2.1	En philosophie/psychologie . . . . .	31
3.2.2	En sociologie . . . . .	32
3.3	L'identité sociale . . . . .	33
3.3.1	La théorie de l'identité sociale . . . . .	34
3.3.2	Types d'identités sociales . . . . .	37
3.4	L'identité collective . . . . .	38
3.5	Modèles de l'identité . . . . .	39
3.5.1	Un processus de construction d'identité sociale . . . . .	39
3.5.2	Un modèle d'identité . . . . .	40
3.6	Synthèse . . . . .	42

---

L'objectif du présent chapitre est d'exposer le cadre théorique qui nous a permis de répondre à la première question de recherche que nous nous sommes posée : *comment représenter les intérêts individuels et collectifs qu'un agent doit préserver?* Nous proposons ici de s'inspirer de la *théorie de l'identité*. Pour motiver ce choix, nous commençons le chapitre par définir comment cette théorie considère un individu dans une société. Nous avons remarqué que les caractéristiques auxquelles s'intéresse la théorie de l'identité, ou la manière dont elle considère un individu correspondent parfaitement aux propriétés d'un agent. Ainsi, la théorie de l'identité peut permettre aux agents étudiés dans cette thèse de construire une représentation de leurs intérêts individuels et collectifs. Ces derniers font partie intégrante de leurs identités. Ensuite, nous appuyons notre motivation

par des statistiques sur l'utilisation des termes *soi* et *identité* dans la psychologie sociale, qui montrent comment la recherche méthodologique et empirique dans la théorie de l'identité est en forte progression, ce qui peut être exploité et mérite l'attention de la communauté SMA. Après, nous définissons les différents types d'identité depuis plusieurs perspectives, et nous présentons les modèles d'identité sur lesquels nous basons notre modèle d'agent. Nous concluons ce chapitre par la discussion d'une relecture axée agent de la théorie de l'identité.

### 3.1 L'identité et son rôle dans le comportement de l'individu

Bien qu'un individu ne puisse jamais se tenir complètement à l'écart de la société, il conserve néanmoins une certaine indépendance vis-à-vis de celle-ci. Il est capable de choisir comment on se comporte envers les autres d'une manière qui, pour lui, est "appropriée" et qui satisfait ses propres besoins, souhaits et désirs. La société ne peut lui présenter qu'un ensemble de choix, elle ne peut jamais déterminer complètement pour lui quels choix il adopte réellement. Les pressions sociales le contraignent toujours dans une certaine mesure pour se conformer ou s'inscrire dans des modèles de comportement établis. Cependant, il y a toujours un espace privé et personnel dans lequel il est libre de faire lui-même ses choix et d'être responsable, s'il le souhaite. En ce sens, il tient compte de l'influence de la société, sous la forme de règles, de normes, de lois, et d'attentes sur la manière dont les autres se comportent envers lui. Celles-ci influencent inévitablement ses choix et ses décisions, mais il est libre de décider lui-même de leur applicabilité ou de leur pertinence pour lui et pour les situations dans lesquelles il se trouve. Très souvent, il les utilise simplement comme lignes directrices et invente ses propres versions adaptées. Ceci est à cause de deux caractéristiques :

- Premièrement, il est autonome capable de penser et agir indépendamment. Il a les connaissances et les compétences qui lui permettent de se comporter avec d'autres personnes et situations selon ses propres termes.
- Deuxièmement, chaque individu est unique parce qu'il a eu un ensemble unique d'expériences. Chacun fait l'expérience du monde d'une manière différente des autres, et la façon dont il répond aux événements est aussi différente.

Un individu ne peut s'absoudre de son environnement, car il est influencé par sa famille, ses amis, son éducation, son travail, sa classe sociale, son sexe, la politique (notamment de son pays), son histoire, etc. À chaque instant de sa vie, il est à la fois dépendant et

contributeur de son environnement social. D'un point de vue social, il doit définir prudemment le cadre des interactions avec les membres de son réseau social. Et d'un point de vue personnel, son comportement et ses émotions ne sont pas simplement déterminés par la société, car il a toujours la liberté de choisir quoi faire et comment le faire. Ainsi, un individu doit à la fois, (i) interagir *dans* un milieu social et (ii) garder le contrôle de son état interne *à l'écart de* ce même milieu. Ces deux aspects peuvent apparaître incompatibles mais ils vont naturellement ensemble dans notre vie sociale, ce qu'explique la théorie de l'identité [Burke and Stets, 2009].

La théorie de l'identité est utilisée en psychologie sociale pour étudier la dynamique intra-personnelle, inter-personnelle et inter-groupe afin de définir le comportement des individus en groupes dans, pratiquement, le contexte de cette thèse. Nous considérons cette théorie comme un cadre fondamental pour :

- maintenir les descriptions des groupes auxquels un agent appartient,
- identifier les objectifs par les membres de chaque groupe,
- garder des représentations d'autres agents, et par conséquent,
- intégrer processus de construction de l'identité d'un agent.

### 3.1.1 Le soi et l'identité

L'importance du *soi* et de l'*identité* en sociologie et en psychologie remonte d'abord aux travaux originaux de [James, 1890], suivis de certains efforts importants des sociologues [Cooley, 1902] et des philosophes [Mead, 1934]. Pourtant, la recherche approfondie et empirique sur le soi et l'identité n'a prospéré comme une science méthodologiquement rigoureuse qu'avec la renaissance de la sociologie après la seconde guerre mondiale, et la révolution cognitive de la psychologie dans les années 1950. Selon l'étude du sociologue Owens [2006], et en se basant sur des statistiques de PsycINFO (base de données bibliographique produite par l'association américaine de psychologie), la théorie et la recherche sur le soi et l'identité, bien qu'elle soit née dans les années 1950, n'est pas très ancienne d'après le nombre de publications. Cette remarque est particulièrement vraie pour l'identité. Comme le montrent les chiffres qui suivent, les articles de revue et les dissertations mentionnant *soi* ou *identité* dans leurs titres ou résumés ont connu une croissance exponentielle des années 1950 à aujourd'hui.

Ces chiffres montrent la popularité croissante de ces concepts, et indiquent que la théorie et la recherche sur le soi et l'identité continueront leur vigoureuse représentation dans la littérature.

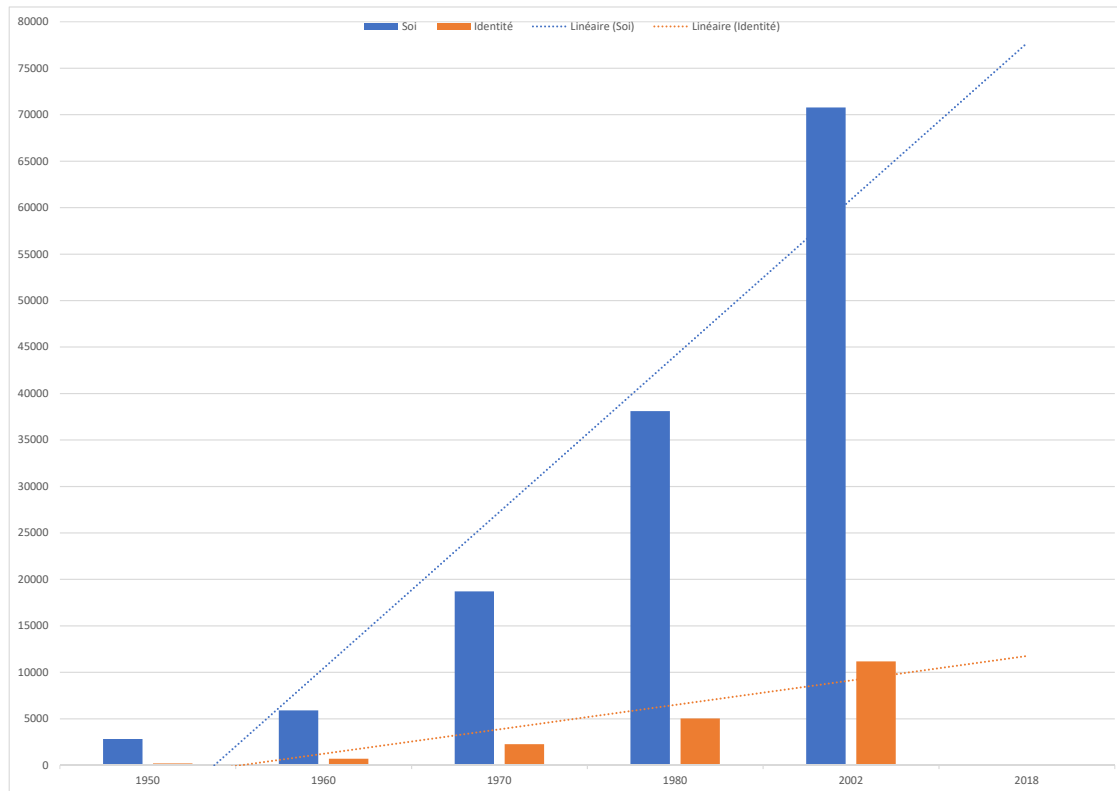


FIGURE 3.1 : Évolution du nombre des articles et dissertations sur le *soi* et l'*identité*.

**Soi ou identité?** Le soi et l'identité sont des concepts complémentaires ayant des traits en commun, mais ils sont tout de même des concepts distincts. Cette complémentarité survient parfois à cause de l'imprécision et de la confusion, en particulier dans la façon dont ils sont similaires mais distincts. Owens [2006] a différencié le soi de l'identité en considérant que le soi est un *processus et une organisation née de l'autoréflexion* alors que l'identité est un *outil* ou stratagème avec lequel les individus ou les groupes se catégorisent et se présentent à la société.

**Définition 3.1 (Soi).** Le soi est *un système organisé et interactif de pensées, de sentiments, d'identités et de motivations* [Owens, 2006] qui

- est né de l'auto-réflexivité et du langage,
- que les individus s'attribuent,
- qui caractérise des êtres humains spécifiques.

Chaque individu développe une compréhension de lui-même comme physiquement distinct et séparé des autres. Au-delà d'un soi physique, les individus ont un soi privé qui se définit par la conscience des pensées et des sentiments internes qui sont privés dans la mesure où ils ne peuvent pas être directement connus par les autres. Pourtant, ces états

internes sont en grande partie façonnés par la compréhension partagée au sein d'une société particulière de ce que doit être l'être humain. Le soi est une vision composite d'un individu qui est formée par l'expérience directe et les évaluations d'autres personnes significatives. Il est considéré comme une structure d'interprétation dynamique qui intervient dans les processus intra-personnels (le traitement cognitif de l'information, l'affection et la motivation) et inter-personnels (l'interaction avec le milieu social : perception sociale, réaction, choix d'une situation d'interaction) les plus importants. Le soi est la représentation mentale d'une personne de sa propre personnalité, de son identité sociale et de ses rôles sociaux. De plus en plus de travaux en psychologie considèrent que le soi doit fonctionner dans un environnement, et qu'il n'y a pas de façon significative de parler d'individu abstrait de sa société. En fait, l'être humain est incomplet et donc incapable de fonctionner correctement à moins d'être intégré dans une société spécifique. Les relations durables et les engagements envers l'environnement social aident à définir qui est une personne. Dans les travaux en sociologie, il y a une tendance similaire pour définir le concept de soi comme enraciné dans le cadre social et comme un reflet des réactions et des évaluations des autres. Les normes de référence d'un groupe social deviennent les normes internalisées par rapport auxquelles les individus se jugent. Ainsi, le soi est modifié dans une large mesure par les processus d'apprentissage social.

L'identité est subsumée dans le concept plus large de soi, et est un nouveau venu à la psychologie sociale [Owens, 2006].

**Définition 3.2 (Identité).** L'identité consiste en la catégorisation du soi en tant qu'occupant d'un rôle, et l'incorporation, dans le soi, des significations et des attentes associées à ce rôle et à son fonctionnement. Ces attentes et significations forment un ensemble de normes qui guident le comportement [Stets and Burke, 2000].

L'identité est un concept qui définit ce qu'est une personne en particulier. Ce concept est devenu un terrain fertile pour comprendre le comportement collectif, l'expérience personnelle et la relation entre soi et la société. DeLamater and Ward [2006] définit quatre prémisses qui sous-tendent la théorie de l'identité :

- les individus sont des acteurs et des réacteurs ;
- l'action et les interactions humaines sont façonnées en grande partie par les définitions que les acteurs tirent de la situation, et ces définitions sont basées sur des significations partagées qui se posent lorsque les individus interagissent les uns avec les autres ;
- les significations que les individus attribuent à eux-mêmes, et donc leur auto-concepts, sont cruciaux pour le processus qui produit leurs actions et leurs in-

teractions ;

- comme d'autres significations, les auto-conceptions sont moulées dans le cours de l'interaction avec les autres et sont en grande partie les résultats des réponses des autres à la personne.

Les divers éléments qui composent l'identité d'un individu peuvent être divisés en deux catégories générales. L'une est la perception privée de soi et des sentiments de continuité et d'unicité, que l'on appelle *identité personnelle* ; l'autre concerne les rôles et les relations sociales d'une personne, que l'on peut appeler *identité sociale*.

## 3.2 L'identité personnelle

Plusieurs études dans différentes disciplines s'efforcent à définir les termes autour de la théorie de l'identité, même si [James, 1890] a, il y a plus d'un siècle, fait la remarque que le soi (y compris l'identité) est "la question la plus énigmatique que la psychologie doit traiter". Cette perplexité se manifeste par exemple par l'utilisation imprécise de soi et identité (comme dans [Layder, 2004]). De même, le concept d'identité personnelle est utilisé et défini de différentes manières.

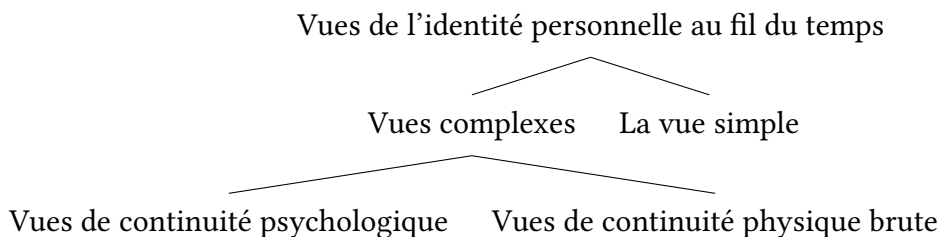


FIGURE 3.2 : Une taxonomie des points de vue sur l'identité personnelle (traduite de [Gasser and Stefan, 2012]).

Il existe deux manières pour aborder le concept d'identité personnelle : une *vue simple*, et une *vue complexe* [Gasser and Stefan, 2012].

La vue simple de l'identité personnelle soutient que l'identité personnelle n'est pas constituée par des continuités de propriétés physiques ou mentales ou de la substance physique qui les compose, mais elle est une caractéristique distincte du monde, bien qu'elle soit compatible avec l'identité personnelle étant causée par de telles continuités. Par exemple,



une personne  $p_2$  à un moment  $t_2$  peut être la même personne qu'une personne  $p_1$ , dans un temps plus tôt  $t_1$ , quelles que soient les propriétés physiques ou mentales et quel que soit le corps possédé par chaque personne.  $P_2$  ne peut pas se souvenir ce qu'a fait ou vécu  $p_1$  au moment  $t_1$  ou avant, et peut avoir un caractère totalement différent et un corps totalement différent de  $p_1$  à  $t_1$ . Les principaux arguments en faveur de la vue simple consistent à proposer des expériences de pensée dans lesquelles les personnes subissent des changements radicaux de la vie mentale et de la constitution corporelle, et dans lesquelles il est revendiqué qu'il est possible qu'elles continuent d'exister. On déduit ainsi que les continuités mentionnées ne sont pas nécessaires à l'identité personnelle.

La vue complexe considère que l'identité personnelle est constituée par un certain degré particulier de continuité des propriétés physiques et mentales et de la matière qui forme le corps d'une personne. Les principaux arguments en faveur de ce point de vue sont que les exemples de l'identité personnelle sont tous ceux dans lesquels il existe une certaine continuité, et que la *vue simple* conduit à des contradictions. Il existe plusieurs variétés de la vue complexe selon lesquels les degrés de continuité assurent l'identité d'une personne au fil du temps. Une première variété est l'opinion que le concept d'identité personnelle n'a aucune application en dehors des situations normales de fortes continuités physiques et mentales. Une autre variété de la vue complexe soutient que l'identité personnelle est nécessairement une question de degré. Plus les continuités sont faibles, moins le degré auquel la personne à  $t_2$  est la même que la personne à  $t_1$ .

### 3.2.1 En philosophie/psychologie

[Li et al. \[2011\]](#) considèrent que l'identité personnelle englobe l'ensemble de conditions nécessaires et suffisantes dans lesquelles le soi persiste dans le temps. Une identité personnelle peut inclure des informations personnelles (nom, date et lieu de naissance), des identifiants personnels (numéro de sécurité sociale, numéro de passeport, etc.), des descriptions physiques (taille, poids, etc.) et des informations biométriques (par ex. empreinte digitale, ADN).

Pour définir l'identité personnelle, [Garrett \[2002\]](#) se pose d'abord des questions sur ce qui fait qu'une même personne persiste au fil du temps, et sur les conditions qui doivent être remplies pour qu'une même personne continue d'exister. Pour répondre à ces questions, il a défini des critères (physique et psychologique) qui ne prétendent pas simplement offrir des façons de définir assez généralement qui est une personne, mais prétendent préciser en quoi consiste l'identité personnelle dans le temps. Selon le *critère physique*, l'identité d'une personne au cours du temps consiste au maintien d'une certaine rela-

tion de continuité physique entre la même personne à des moments différents. Selon cette vision, une identité personnelle, c'est être le même objet biologique au fil du temps. Plusieurs versions du critère physique diffèrent en fonction de quel élément biologique fonde l'identité personnelle : le corps, le cerveau, le système nerveux, etc. Toutes ces versions soutiennent que l'identité personnelle consiste en la continuité d'un élément biologique. De même, selon le *critère psychologique*, l'identité d'une personne au cours du temps consiste dans la tenue de la relation de continuité psychologique entre une personne à des moments différents. Cette relation est composée d'un certain nombre de chaînes de connexions psychologiques enclenchées, telles que celles qui tiennent au fil du temps entre l'expérience-mémoire et l'expérience-souvenir, entre une intention et l'action qui la manifeste, ou la chaîne de rétention des croyances, des désirs, des souvenirs, et du caractère. D'autres travaux de Noonan [2003] étendent la liste des critères en considérant le *critère corporel*, le *critère cérébral*, et le *critère de mémoire*.

Strawson [2014] a proposé une interprétation du chapitre "Of Identity and Diversity" de l'œuvre philosophique "An Essay Concerning Human Understanding" de Locke [1836] publié initialement en 1689, qui est devenu l'un des ouvrages fondateurs de l'empirisme (l'un des courants majeurs de la théorie de la connaissance). Strawson [2014] propose qu'avant d'étudier l'identité personnelle, il faut définir ce qu'est une personne. Selon lui, une personne est un être intelligent capable de penser, de raisonner et de réfléchir. Elle peut se considérer comme soi-même, la même chose pensante, dans des temps et des lieux différents, elle ne le fait que par cette conscience, inséparable de la pensée. Il considère qu'une personne *Lockeéenne* [P] est composée de trois éléments  $[P] = [M] \pm [I] + [A]$  tels que : [M] est le corps matériel humain, [I] est l'âme immatérielle, et [A] est l'ensemble des actions, pensées, expériences, présentes et passées. Ainsi, l'identité personnelle est la réponse philosophique à la question suivante : quelles sont les conditions nécessaires et suffisantes de la vérité de la réclamation qu'une personne [P] considérée maintenant à l'instant  $t_2$ , est la même personne qu'une personne  $[P_X]$  considérée à un autre moment  $t_1$ ? Qu'est-ce qui doit être vrai si  $[P_X]$  est la même personne que [P] ?

### 3.2.2 En sociologie

Layder [2004] considère l'identité personnelle comme la façon dont une personne se considère, comment elle considère les autres, et comment elle se rapporte ou se comporte avec eux. Il s'agit du centre de la conscience, des besoins émotionnels, des désirs et du contrôle de soi, en fonction desquels l'individu réfléchit et agit sur ses interactions sociales. Pour mieux comprendre l'identité personnelle, il distingue entre trois aspects :

les attributs généraux, l'organisation interne, et les besoins/fonctions/capacités. Selon

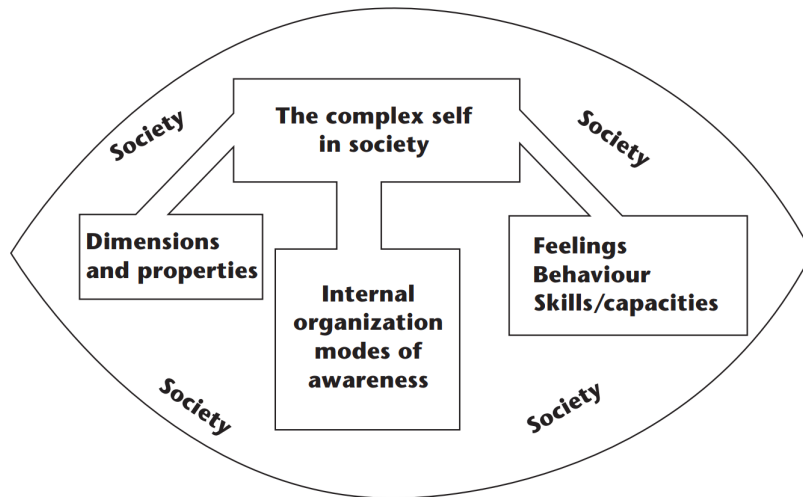


FIGURE 3.3 : Composants de l'identité personnelle selon [Layder, 2004].

lui, l'identité personnelle a cinq propriétés et/ou dimensions, l'identité personnelle :

- est de nature psychologique et sociale,
- a une dimension émotionnelle,
- est capable de différentes intensités de contrôle de soi et des autres,
- est flexible, évolue, change, et se développe,
- a un aspect spirituel dormant et souvent non développé (*soi supérieur*)

Les identités personnelles servent de base sur laquelle les identités sociales et les biographies personnelles peuvent s'appuyer [McCall and Simmons, 1966]. Si un individu ne pouvait être reconnu d'une interaction à l'autre comme la même personne, aucune relation sociale stable ne pouvait être construite et, par conséquent, il n'y aurait aucune identité sociale.

### 3.3 L'identité sociale

L'histoire de l'humanité a connu toutes sortes de préjudices et discriminations. Il y a des comportements inter-groupes extrêmement horribles comme l'annihilation de races entières par exemple. Il y a aussi d'autres sortes d'inhumanité des individus envers leur propre genre : la haine, la domination, la subjugation, l'exploitation, la dégradation, l'oppression et l'extermination, sont toutes des caractéristiques de notre histoire. Et pourtant, les individus humains font parties des plus sociables de tous les êtres vivants. Ils se réjouissent et prospèrent en compagnie des autres. Ils ne passent pas nécessairement

toutes leurs vies en présence des autres, mais plus fondamentalement ils sont des produits de l'histoire, de la culture et de la société. On peut dire qu'ils sont construits socialement. Leurs points de vue, opinions, valeurs, activités et moyens de communication sont appris ou acquis des autres. Leur comportement est en grande partie régi par des normes, ou des accords entre personnes, pour convenir des manières appropriées et acceptables de se comporter, et des opinions à tenir dans des circonstances particulières. Sans un tel accord, la coopération, qui est au cœur de l'existence humaine, serait impossible. On en vient alors à ce paradoxe : comment expliquer la coexistence de cohésion nécessaire à l'existence sociale avec des divisions dans la société ? Ce paradoxe est remis en lumière dans la psychologie sociale [Abrams and Hogg, 2006], qui s'intéresse aux comportements des individus dans des groupes.

### 3.3.1 La théorie de l'identité sociale

L'identité sociale est une théorie de la psychologie sociale qui s'intéresse aux relations inter-groupes, aux processus de groupes, et au soi social. Elle trouve son origine dans les travaux de Henri Tajfel sur les facteurs sociaux de la perception [Tajfel, 1969], et sur les aspects de croyances sociales et cognitives du racisme, préjudice, et discrimination [Tajfel, 1970]. L'identité selon cette théorie peut donc être définie de manière générale comme : les catégories que les individus utilisent pour spécifier qui ils sont, et pour se situer par rapport à d'autres personnes. La définition la plus répandue de l'identité sociale est celle de Tajfel [1978].

**Définition 3.3** (Identité sociale). *La part de la conception de soi qui découle de sa connaissance, de son appartenance à un groupe social, ainsi que les valeurs et la signification émotionnelle attachés à cette appartenance*”.

Durant les années 1980, John Turner a initié un développement théorique très important de la théorie de l'identité sociale en concevant la théorie de l'*auto-catégorisation* [Turner, 1985]. Même si cette théorie est distincte de la théorie de l'identité sociale, elle est assez proche pour être considérée comme faisant partie de la même perspective métathéorique. En ce sens, l'identité implique à la fois un caractère distinctif (je ne suis pas comme eux, *out-group*) et une similitude comme les autres (je suis comme eux, *in-group*).

**La théorie de l'auto-catégorisation** Les individus sont essentiellement dépersonnalisés : ils sont perçus comme étant des réactions au comportement du groupe et non pas

comme des individus uniques. La dépersonnalisation de soi est le processus fondamental qui sous-tend les phénomènes de groupe (stéréotypes sociaux, cohésion de groupe, coopération et altruisme, contagion émotionnelle et empathie, comportement collectif, normes partagées et processus d'influence mutuelle) [Hogg et al., 1995]. Il n'a pas d'implications négatives dans la mesure où il se réfère simplement à un changement contextuel du niveau d'identité (de l'individu unique au membre du groupe), et non à une perte d'identité. Grâce à la dépersonnalisation, l'auto-catégorisation met en adéquation la perception de soi et le comportement avec le cadre contextuel pertinent au sein du groupe, transformant ainsi les individus en membres du groupe et l'individualité en comportement de groupe. La théorie de l'auto-catégorisation [Turner, 1985] est un développement récent qui élabore en détail le fonctionnement du processus de catégorisation en tant que base cognitive du comportement des individus dans un groupe. Ce dernier accentue à la fois les similitudes et les différences perçues entre personnes appartenant à la même catégorie ou à différentes catégories sociales. La catégorisation de soi et des autres en *in-group* et *out-group* définit l'identité sociale des individus et accentue leur similitude perçue avec la représentation cognitive des personnes et des caractéristiques déterminantes du groupe. Selon la théorie de l'auto-catégorisation, l'identité peut être considérée comme une série d'auto-catégorisations à différents niveaux d'inclusion, allant des plus spécifiques aux plus larges. Une personne utilisera un niveau spécifique de l'auto-catégorisation en fonction de la *saillance* d'une catégorie dans un contexte social donné. La saillance est déterminée par le degré d'accessibilité à une catégorie sociale et par son adéquation contextuelle, c'est-à-dire par son adéquation aux similarités intra-catégories et aux différences inter-catégories.

Selon la perspective de l'identité sociale, les individus se définissent et s'évaluent en fonction des groupes auxquels ils appartiennent, dans le sens où les groupes fournissent aux individus un concept de soi social, et les individus ont autant d'identités sociales que les groupes auxquels ils se sentent appartenir. Comme les identités sociales définissent, prescrivent, et évaluent qui on est et comment on doit penser, sentir, et agir en fonction des cadres sociaux, les individus ont un fort désir d'établir ou de maintenir la supériorité de leur propre groupe par rapport aux autres groupes concernées. Il y a une lutte entre les groupes pour un caractère distinctif positif de chacun. Cette lutte est tempérée par la compréhension des individus de la nature des relations inter-groupes, qui est à la base du système de croyances sociales d'un groupe. En particulier, les individus prêtent attention à plusieurs aspects selon Burke [2006] : les croyances sur le statut social de son groupe par rapport à un *out-group*, les croyances sur la façon dont cette relation est stable, les croyances sur la façon dont elle est légitime, les croyances sur la perméabi-

lité de la frontière entre son *in-group*, et les *out-groups* et donc la possibilité de passer psychologiquement d'un groupe à l'autre.

En plus des travaux de Tajfel et ses collaborateurs, la théorie de l'identité sociale est un sujet de grand intérêt pour d'autres philosophes, psychologues, et sociologues. [Bagozzi \[2000\]](#) considère que l'identité sociale consiste en trois éléments : les processus d'auto-catégorisation sociale (c.-à-d., une personne se voit comme un membre d'une catégorie sociale, ce qui a des implications sur le sens de soi, et conduit à la *dépersonnalisation* et à différentes perceptions et actions envers son *in-group* et *out-group*) ; les sentiments en conséquence de l'appartenance à un groupe (e.g., la fierté, la solidarité, le pouvoir) ; l'estimation de soi fondée sur le groupe (les perceptions de similarité, d'appartenance). Un groupe existe psychologiquement si trois personnes ou plus se comportent et s'évaluent en termes d'attributs partagés qui les distinguent collectivement des autres personnes [[Burke, 2006](#)]. On a besoin d'au moins trois personnes pour déduire les normes de groupes et le comportement des autres (dans une dyade, il n'y a qu'une seule "autre personne"). D'une autre part, plusieurs processus de groupes ne peuvent pas avoir lieu dans une dyade (e.g., l'influence sociale, la formation de coalition, la déviance). [Li et al. \[2011\]](#) définissent l'identité sociale d'une personne comme l'histoire biographique qui s'accumule avec le temps. Elle reflète la façon dont une personne interagit avec la société à laquelle elle appartient. D'un point de vue psychologique, l'identité sociale est la perception de soi en tant que membre de certains groupes sociaux (les membres d'un groupe social peuvent partager certaines caractéristiques communes qui peuvent se distinguer des autres groupes). Du point de vue sociologique, l'identité sociale est caractérisée par les relations sociales entre des acteurs qui jouent des rôles complémentaires (Identité de rôle, Section 3.3.2).

Selon [Seddone \[2014\]](#), l'identité sociale se construit autour de trois aspects fondamentaux :

- l'appartenance d'un individu à un groupe social (i.e. son rôle dans ce groupe, ses tâches, et l'avantage derrière cette appartenance),
- l'appartenance des autres membres du groupe social,
- la conscience de l'existence d'un groupe social par les individus et de l'objectif global partagé du groupe.

Ces trois aspects sont nécessaires pour la construction d'une identité sociale, car ils proviennent de l'individu et deviennent communs entre les individus partageant la même identité sociale.

### 3.3.2 Types d'identités sociales

On identifie plusieurs types d'identité sociale dans la littérature. Brewer [2001] distingue quatre types :

- les *identités sociales basées sur la personne*, qui insistent sur la façon dont les propriétés des groupes sont internalisées par les membres individuels du groupe dans le cadre du concept de soi ;
- les *identités sociales relationnelles*, qui définissent le soi par rapport à d'autres personnes spécifiques avec lesquelles on interagit dans un contexte de groupe ;
- les *identités sociales de groupe*, qui se réfère à la perception de soi en tant que partie intégrante ou interchangeable d'un groupe plus large ou d'unité sociale (comme définie dans *la théorie de l'auto-catégorisation*) ;
- les *identités collectives*, qui désignent un processus par lequel les membres du groupe ne partagent pas seulement des attributs auto-définis, mais s'engagent également dans une action sociale pour forger une image de ce que le groupe représente et comment il est représenté et vu par les autres. Nous allons définir ce type d'identité dans la Section 3.4.

Brewer and Gardner [1996] distinguent trois aspects du soi dans la définition de l'identité sociale :

- le soi individuel, qui est défini par les traits personnels qui différencie le soi des autres ;
- le soi relationnel, qui définit les relations dyadiques entre le soi et certaines personnes ;
- le soi collectif, qui est défini par l'appartenance à un groupe qui différencie entre le "nous" du "eux".

**Identité de rôle** Chaque identité de rôle inclut toutes les significations qu'une personne attache à lui-même en jouant un rôle. Ces significations sont, en partie, dérivées de la culture et de la structure sociale (les individus sont socialisés dans ce que signifie être professeur, chef d'entreprise, etc.). Cependant, les personnes apportent aussi dans leur identité de rôle un peu de leurs propres compréhensions et interprétations quant à ce que l'identité signifie pour eux. Ainsi, les significations associées aux identités de rôles sont à la fois partagées et subjectives. Les individus doivent donc négocier ces dernières avec d'autres qui peuvent avoir un ensemble différent de compréhensions sur les significations d'une certaine identité de rôle. Quelles que soient les significations de

l'identité, elles sont liées aux significations implicites de son comportement de rôle. En d'autres termes, il existe une correspondance entre ces significations personnelles dans un rôle et un comportement de rôle. Par exemple, l'identité de rôle de l'enseignant peut impliquer des significations comme aider et encourager les élèves à exprimer des idées, guider l'action, organiser la communication, répondre aux questions. Le rôle d'un chef d'équipe peut inclure des significations de la puissance du contrôle, de la transmission des instructions, de la formation des salariés, de l'affectation des postes de travail, de la coordination du travail de l'équipe. Dans les identités basées sur les rôles, il existe différentes perceptions et actions entre individus, car les rôles interagissent avec les rôles complémentaires. Toute identité de rôle qu'une personne réclame dans une interaction a toujours un rôle complémentaire connexe revendiqué par une autre. Ainsi, toute identité de rôle est toujours liée à une identité complémentaire correspondante [Burke and Reitzes, 1981]. Par exemple, l'identité de rôle de la fille est mise en œuvre en ce qui concerne l'identité du rôle complémentaire correspondant de la mère ou du père. L'identité de rôle de l'élève se joue par rapport à l'identité du rôle complémentaire de l'enseignant. De plus, les rôles activent un sentiment d'efficacité plutôt qu'un sentiment d'estime de soi comme dans les identités sociales. C'est ce que l'on fait dans son rôle qui est important par rapport à ce qu'on est à base de notre identité sociale.

### 3.4 L'identité collective

L'identité personnelle regroupe les attributs et les significations qu'un individu attribue à lui-même, tandis que les identités sociales peuvent s'intégrer dans l'identité personnelle, et naître des appartenances sociales ou des rôles que joue un individu. Or, l'identité collective est distincte de l'identité personnelle et sociale de plusieurs façons.

- Premièrement, l'identité collective n'est pas nécessairement intégrée dans l'identité sociale existante parce qu'elle est souvent émergente plutôt qu'ancrée dans les groupes sociaux [DeLamater and Ward, 2006].
- Deuxièmement, l'identité collective tend à être plus éphémère que l'identité sociale ou personnelle [Snow, 2001].
- Troisièmement, contrairement à l'identité personnelle et sociale, l'observation d'une identité collective se fait en considérant les relations entre groupes.

Le concept d'identité collective implique une représentation partagée d'un groupe, fondée sur des expériences et des intérêts communs [Brewer, 2001], mais il fait également référence à un processus actif de façonnage et de construction d'une image de ce que le



groupe représente et comment il souhaite être perçu par les autres. Ainsi, l'identité collective représente une réalisation d'efforts collectifs, au-delà de ce que les membres ont en commun au départ (identité sociale). En tant que tel, le concept d'identité collective constitue un lien critique entre l'identité sociale et l'action collective.

### 3.5 Modèles de l'identité

Il y a plusieurs travaux dans la littérature de la théorie de l'identité qui s'intéressent à modéliser le comportement d'un individu et les relations inter-groupes. Dans cette section, nous présentons deux modèles qui ont le plus influencé notre réflexion sur la contribution de cette thèse car ils s'intéressent à la construction de l'identité d'un individu. Le premier est un modèle de construction d'identité sociale, et le deuxième est un modèle d'identité.

#### 3.5.1 Un processus de construction d'identité sociale

Postmes et al. [2005] proposent un modèle (voir la Figure 3.4) pour refléter et orienter la recherche dans la formation de l'identité sociale. Dans ce sens, l'identité sociale est cruciale pour déterminer la forme et la direction de l'influence sociale d'un individu sur les membres du groupe. Le modèle proposé consiste en deux processus. Le premier per-

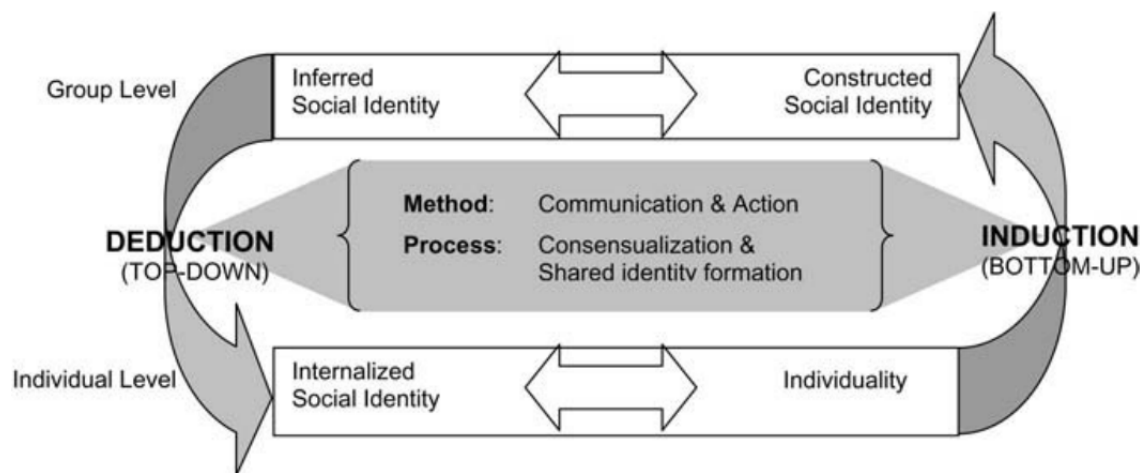


FIGURE 3.4 : Un modèle de formation de l'identité sociale [Postmes et al., 2005].

met aux membres d'un groupe de déduire l'identité sociale à partir du contexte social plus large (*déduction* de l'identité sociale). Cela peut être vu comme un processus *top-down* de formation de l'identité sur la base de compréhensions perçues à partir de réa-

lités existantes dans la structure sociale. Il existe un deuxième processus, moins étudié, par lequel les contributions des membres du groupe servent comme entrées de données pour l'*induction* de leur identité sociale. Ce processus peut être vu comme un approche *bottom-up* par laquelle le comportement des individus informe les attributs de leur identité sociale.

### 3.5.2 Un modèle d'identité

Dans la théorie de l'identité, la modèle d'identité est un modèle de processus. Il s'agit d'un système dynamique et continu qui répond aux changements de l'environnement (ou ceux produits par d'autres acteurs sur le même environnement) qui perturbent les perceptions du système. Ce qui différencie le modèle d'identité de la plupart des modèles comportementaux est que le comportement n'est pas seulement une fonction d'identité ou une réponse à une stimulation, il est considéré comme résultant de la relation entre les perceptions de l'environnement et la *norme d'identité* interne à l'individu. Par exemple, même si la norme d'identité d'un professeur est définie à un niveau élevé d'intellectualité, son comportement ne reflète pas toujours ce niveau. Il peut être plus ou moins élevé selon ses perceptions actuelles de soi-même et d'une situation spécifique. S'il se considère comme se comportant à un niveau d'intellectualité supérieur à celui qui est défini dans sa norme d'identité, il agirait moins intellectuellement pour compenser, et vice-versa. Son comportement est donc une fonction de la relation entre sa norme d'identité et sa propre perception des significations auto-pertinentes. Ses actions sont menées dans le cadre de neutraliser les perturbations qui contrebalancent ses perceptions par rapport à celles contenues dans sa norme d'identité. [Burke \[2006\]](#) propose un modèle d'identité qui est sous forme de boucle de rétroaction qui s'établit une fois qu'une identité est activée (voir Figure 3.5).

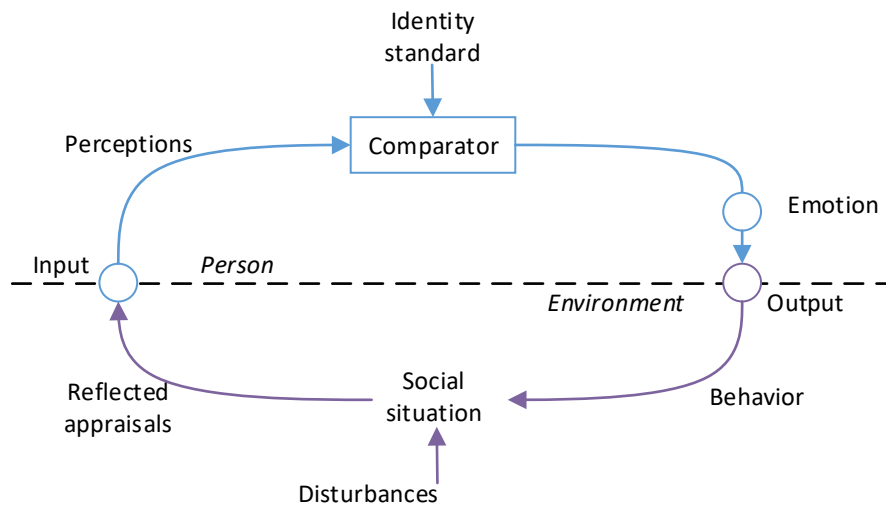


FIGURE 3.5 : Modèle d'identité

Cette boucle a quatre composants principaux :

- *identity standard*, c.-à-d., les significations de soi de son identité,
- *perceptions*, c.-à-d., les entrées perceptuelles de la situation, y compris comment on se voit, et le retour significatif que le soi obtient des autres (*reflected appraisals*),
- *identity comparator*, c.-à-d., un processus qui compare l'entrée perceptive avec l'identité que l'individu a actuellement,
- *meaningful behavior*, c.-à-d., le résultat de la comparaison entre les *perceptions* des significations de soi de la situation avec la vraie signification de soi-même contenue dans la *norme d'identité*.

Ainsi, le comportement d'un individu contrôle ses perceptions et sa norme d'identité lorsqu'ils sont perturbés par les événements que génère chaque situation d'interaction sociale, telle que les perceptions correspondent à la norme d'identité interne qui définit ce que l'individu est. Ce contrôle n'est pas direct, il se fait à travers la situation sociale, en modifiant les perceptions de la situation qui sont saisies par les identités de tous les individus impliqués. Les perceptions sont modifiées, puis sont comparées à la norme d'identité, et la boucle continue ininterrompue, en ajustant les perceptions pour être en conformité avec la norme d'identité. Lorsque deux ou plusieurs individus sont impliqués, comme c'est illustré dans la Figure 3.6, le comportement de chacun sert à perturber ou à modifier les perceptions des autres, appelant ainsi à un comportement neutralisant de leurs parts. De plus, un individu a plusieurs identités sociales et identités de rôle en fonction des groupes auxquels il appartient, et des rôles qu'il joue dans ces groupes. Par exemple, pour le même individu, il doit maintenir un certain niveau de sociabilité, d'affirmation de soi, d'intellectualité, de responsabilité, étant donné qu'il

est "étudiant". Il doit maintenir un certain degré de féminité ou masculinité, selon son "identité de genre". Dans une interaction sociale, ces facettes de l'identité d'un individu entrent en jeu via leurs processus de comparateur d'identité, tels que l'identité communiquée lors de l'interaction est le résultat du contrebalancement de ces comparateurs.

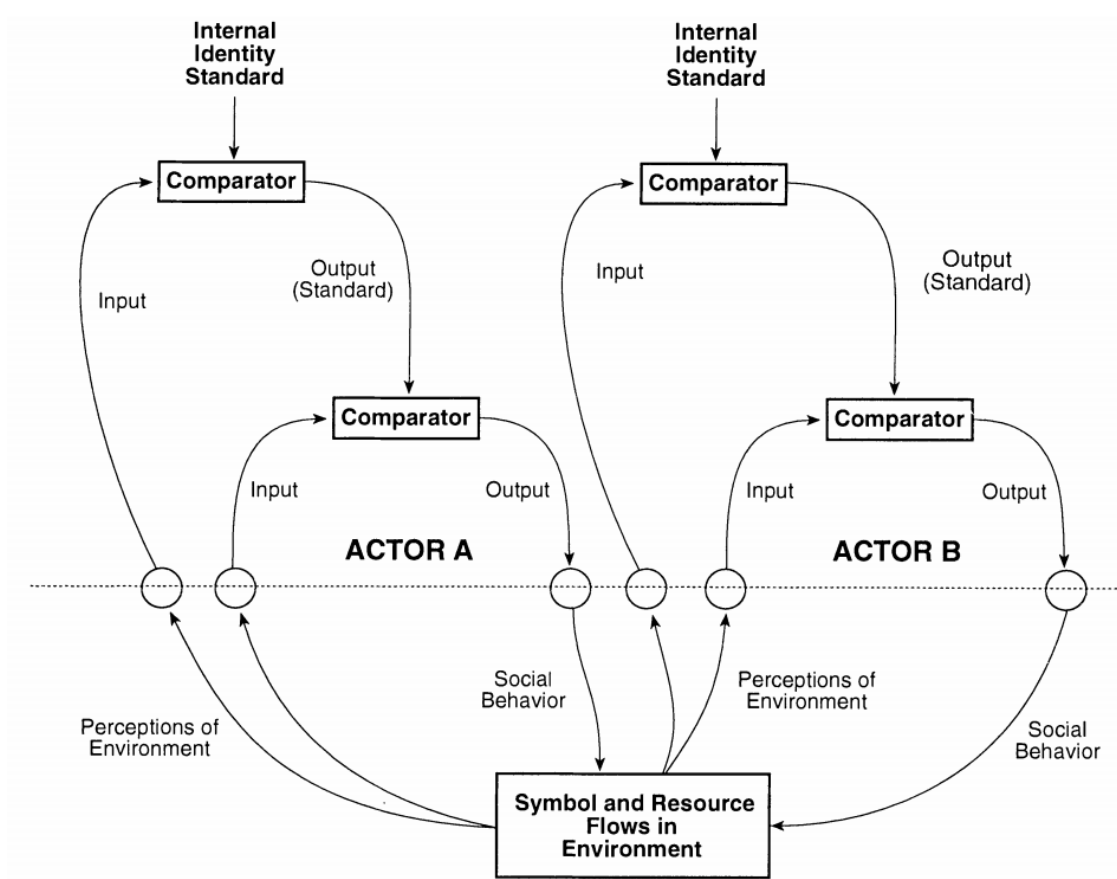


FIGURE 3.6 : Modèle d'identité de deux acteurs en interaction [Burke, 1997]

### 3.6 Synthèse

Les théories de l'identité personnelle et sociale fournissent une perspective théorique à partir de laquelle on peut examiner les relations entre des individus participant à plusieurs groupes sociaux [Abrams and Hogg, 2006]. Dans cette thèse, et plus généralement, les agents sont déployés pour pouvoir agir collectivement sur un environnement. L'état interne de chaque agent est à la fois influencé par ses interactions sociales, qui sont menées et motivées par les états internes des agents. Ainsi, notre définition de l'identité est à la fois de nature sociale et psychologique, et ne peut pas être exclusivement sociale ou psychologique. Le soi conditionne les interactions sociales de l'individu, et en même

temps, le soi ne peut exister que dans un contexte social. En ce qui concerne l'identité sociale, elle intègre à la fois des aspects psychologiques, lorsque le contexte social est un groupe de grande échelle, et des aspects sociaux, pour les relations entre groupes dans lesquels les membres interagissent entre eux. L'identité sociale devient ainsi un concept multi-niveau qui comprend la définition de ses groupes sociaux à différentes échelles.

Nous proposons ici une relecture de ces travaux axés sur l'identité dans un contexte multi-agent. Nous nous sommes intéressés aux liens entre les architectures et les plateformes multi-agents et les théories qui les soutiennent. Dans un premier temps, nous avons complété la classification proposée par Braubach et al. [2008] avec des travaux plus récents issus de différentes revues [Bordini et al., 2006, Gupta and Kansal, 2011, Balke and Gilbert, 2014, Kravari and Bassiliades, 2015], afin de mettre en évidence les contributions multi-agents et leurs fondements théoriques (Figure 3.7).

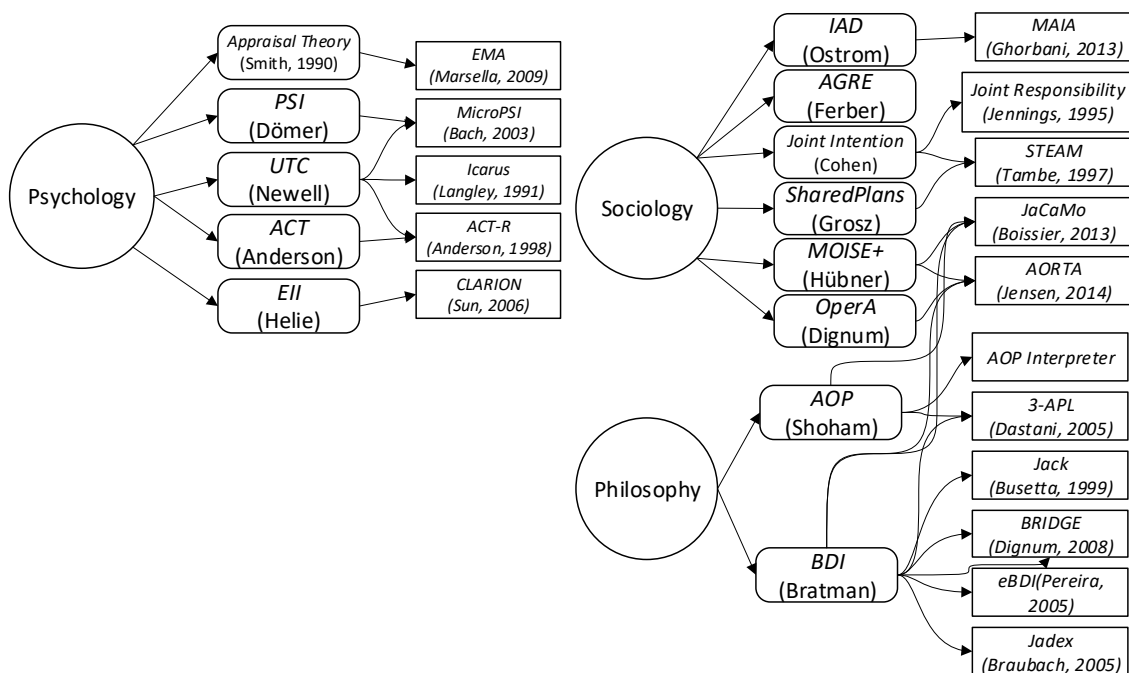


FIGURE 3.7 : Classification des disciplines et théories des architectures d'agents.

Ainsi, nous avons dans un second temps proposé une classification des éléments issus des théories de l'identité afférente utilisés dans ces mêmes architectures (Figure 3.8). On remarque ainsi que les caractéristiques essentielles de l'identité personnelle et de l'identité sociale sont intrinsèquement contenues dans plusieurs architectures, ces dernières partageant les mêmes fondements théoriques que ceux de la théorie de l'identité. Néanmoins, peu d'attention est accordée à la construction de ces identités qu'elles soient personnelles ou sociales.

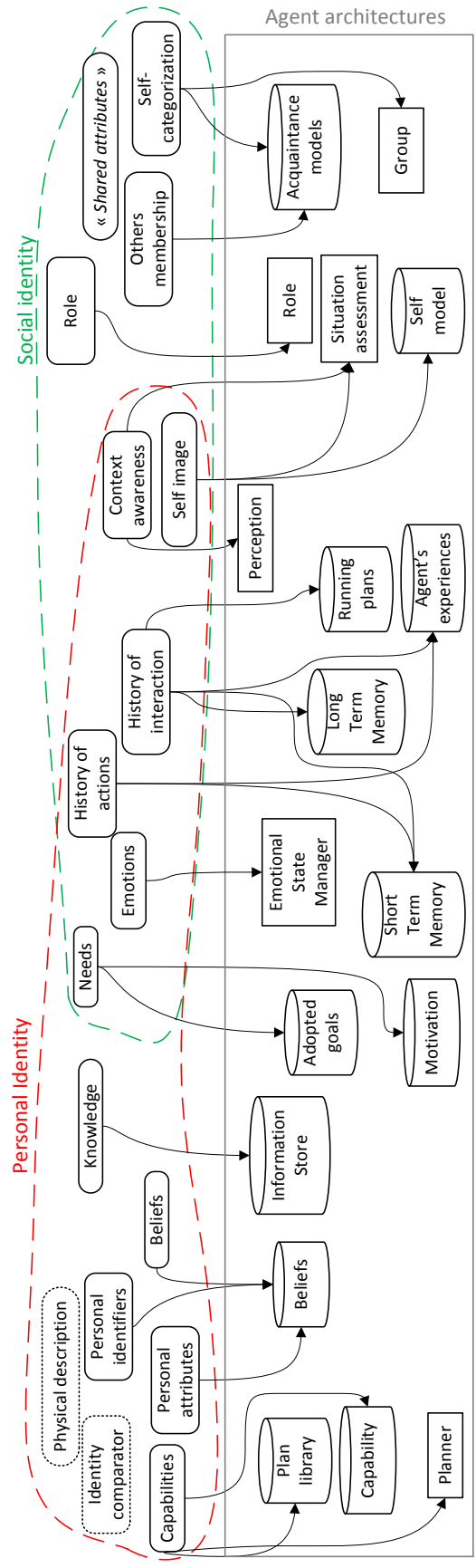


FIGURE 3.8 : Synthèse des éléments utilisés dans les architectures d'agents et leurs analogues dans la théorie de l'identité.

À partir de cette étude illustrée, nous avons notamment constaté dans les architectures d'agent le manque d'un processus homologue au concept philosophique de *comparateur d'identité* [Burke, 2006] dans les architectures d'agents examinées.

**Comparateur d'identité** Dans une situation d'interaction, le processus de comparaison d'identité d'un individu évalue les significations pertinentes tirées de son environnement, et celles tirées des retours d'interaction avec autrui, et les compare avec sa propre identité de référence pour la préserver ou contrôler. Cela nous a amenés à introduire un *processus de construction d'identité* pour permettre la réalisation de cette fonction. Nous avons identifié deux processus qui permettront à un agent d'appréhender ses interactions sociales dans un environnement multi-SMA. Le premier est un processus d'engagement dans des actions collectives, qui s'appuie sur le deuxième processus de construction des identités personnelles et sociales.

**Intersubjectivité** L'auto-conception du soi ou de l'identité d'une personne dépend en partie de ce que les autres pensent de lui, c.-à-d. son identité publique subjective, qui est à son tour basée principalement sur les perceptions que les autres ont de lui (son identité publique objective), que les autres lui communiquent (identité publique objective communiquée). Selon Stern [2004], l'intersubjectivité contribue à la continuité du groupe. Elle favorise la formation de groupe et sa cohérence. Elle permet un fonctionnement collectif plus efficace, rapide, flexible et coordonné. Plus un groupe partage des significations intersubjectives, plus grand est le maintien, la cohésion, et la communication du groupe. Dans le contexte SMA, un agent doit avoir la capacité à prendre en compte les états mentaux des autres agents dans sa perception de l'environnement. Pourtant, il ne peut souvent pas les connaître, il ne peut que déduire leur existence sur la base de ce qui lui est observable, à savoir les actions, les messages échangés, les liens sociaux, et les sollicitations. Nous nous intéressons en fait au concept de intersubjectivité qui - entre deux ou plusieurs agents en interaction - consiste à prendre en considération la reconnaissance de la signification intentionnelle des actions et sollicitations. L'intersubjectivité se réfère à l'interconnexion du *soi*, des *autres* et de l'*environnement*.

Les quelques préconisations établies dans ce chapitre vont fournir les fondements du modèle d'agent centré identité que nous proposons dans le chapitre suivant.





# IDEAS : un modèle d'agent centré identité

---

<b>4.1</b>	<b>Le modèle d'agent de référence</b>	<b>48</b>
4.1.1	BDI	48
4.1.2	Cycle de vie d'un agent BDI	49
<b>4.2</b>	<b>Présentation du modèle IDEAS</b>	<b>53</b>
4.2.1	Architecture d'agent	53
4.2.2	Construction et maintenance d'identité	55
4.2.3	Processus d'engagement centré identité	58
4.2.4	Cycle de vie	61
<b>4.3</b>	<b>Synthèse</b>	<b>63</b>

---

L'objectif de ce chapitre est de présenter le modèle IDEAS (IDentity Agent System) qui intègre des propriétés et des éléments issus de la théorie de l'identité, pour doter l'agent de la capacité de raisonner sur ses potentiels engagements à coopérer dans un contexte multi-SMA, afin d'éviter les situations d'interaction interférant négativement dans la réalisation de ses objectifs.

Nous avons choisi de présenter le modèle IDEAS comme une extension au cycle de décision d'un agent classique tel un agent BDI que nous prendrons comme modèle de référence.

## 4.1 Le modèle d'agent de référence

Au fil du temps, les SMA se sont révélés être une approche très puissante pour modéliser et comprendre les phénomènes collectifs dans des domaines tels que l'économie, le commerce, la santé, l'urbanisme et les sciences sociales. Ils peuvent être utilisés pour étudier ces phénomènes ou résoudre des problèmes difficiles qu'ils posent. À cette fin, la communauté de recherche SMA a développé une variété d'architectures d'agents (génériques ou orientées vers un domaine d'utilisation spécifique) au cours des trois dernières décennies. Toutes ces architectures sont aussi diverses que les communautés des chercheurs qui les produisent. Avec un nombre si grand, le choix de celle qui convient le mieux pour un contexte applicatif fait appel à l'expérience du concepteur du SMA, à la documentation ou aux publications. Il y a plusieurs architectures d'agents qui peuvent nous servir comme base pour exploiter notre modèle IDEAS. Quelques exemples sont illustrés dans l'étude bibliographique que nous avons menée sur les disciplines et les théories des architectures d'agents (voir Figure 3.7, page 43), et d'autres ont fait l'objet d'étude de plusieurs *surveys* [Bordini et al., 2006, Gupta and Kansal, 2011, Balke and Gilbert, 2014, Kravari and Bassiliades, 2015]. Toutes les architectures étudiées contiennent au moins trois éléments en lien direct avec la théorie de l'identité, mais notre choix porte sur un modèle BDI. Ce choix est motivé par ces caractéristiques :

- généralité : il ne dépend pas d'un domaine d'application, et peut être utilisé dans plusieurs domaines et types de problèmes,
- extensibilité : il peut être étendu pour prendre en compte un contexte applicatif donné,
- popularité : disponibilité de plusieurs implémentations, et les documentations qui les accompagnent,
- compatibilité avec notre cadre théorique, au niveau des fondements théoriques (psychologie sociale), et notamment les concepts BDI tels que croyance, désir, intention, objectif, action, historique, etc, qui sont d'une façon ou d'une autre étudiés dans la théorie de l'identité.

### 4.1.1 BDI

Un agent BDI est situé dans un environnement potentiellement changeant, qui reçoit des entrées perceptuelles continues, et qui agit sur cet environnement en fonction de son état mental interne. Les croyances, les désirs et les intentions sont les trois principales attitudes mentales qui capturent les composantes d'information, de motivation

et de décision d'un agent BDI. En plus de ces attitudes, plusieurs extensions du modèle BDI intègrent d'autres notions comme l'émotion [Pereira et al., 2005], les engagements [Winikoff, 2007], les capacités [Rönquist, 2007], la planification [Sardina et al., 2006], etc. L'implémentation BDI que nous utilisons pour construire notre modèle IDEAS est celle utilisée dans l'interprète *Jason* [Bordini et al., 2007] du langage AgentSpeak(L) [Rao, 1996].

Le modèle BDI de base que nous considérons est modélisé par un tuple [Rao, 1996] :

$$A_{bdi} = \langle \mathcal{E}, \mathcal{B}, \mathcal{P}_{lib}, ctx, plan, \mathcal{I}, buf, \mathcal{S}_M, \mathcal{S}_E, \mathcal{S}_O, \mathcal{S}_I \rangle,$$

avec :

- $\mathcal{E}$  un ensemble d'évènements,
- $\mathcal{B}$  un ensemble de croyances,
- $\mathcal{P}_{lib}$  une librairie de plans,
- $\mathcal{I}$  un ensemble d'intentions,
- La fonction *buf* (*belief update function*) sert à mettre à jour l'ensemble  $\mathcal{B}$ ;
- La fonction de sélection  $\mathcal{S}_E$  extrait un évènement de l'ensemble  $\mathcal{E}$  afin de chercher des plans dans  $\mathcal{P}_{lib}$  qui peuvent être déclenchés pour traiter l'évènement sélectionné.
- La fonction  $\mathcal{S}_O$  sélectionne un de ces plans afin de l'associer à une nouvelle intention ajoutée à l'ensemble  $\mathcal{I}$ ,
- $\mathcal{S}_I$  sélectionne une intention de  $\mathcal{I}$  qui déclenche le plan qui lui est associé.

### 4.1.2 Cycle de vie d'un agent BDI

Le cycle de raisonnement simplifié est illustré dans la Figure 4.1. Conformément à ce cycle, un agent BDI perçoit l'environnement où il est situé, pour saisir son évolution de manière à mettre à jour l'ensemble des croyances  $\mathcal{B}$ . L'architecture de l'agent BDI doit avoir un composant capable de percevoir l'environnement, sous une forme symbolique sous la forme d'une liste de littéraux. Chaque littéral est un percept, qui est une représentation symbolique d'une propriété particulière de l'état actuel de l'environnement

**Mise à jour des croyances (*belief update*)** Une fois que la liste des percepts obtenue, la base de croyances  $\mathcal{B}$  doit être mise à jour pour refléter les changements perçus de l'environnement. Ceci est fait par la fonction de mise à jour des croyances *buf* (*belief update function*). Chaque modification de l'ensemble  $\mathcal{B}$  lors de l'exé-

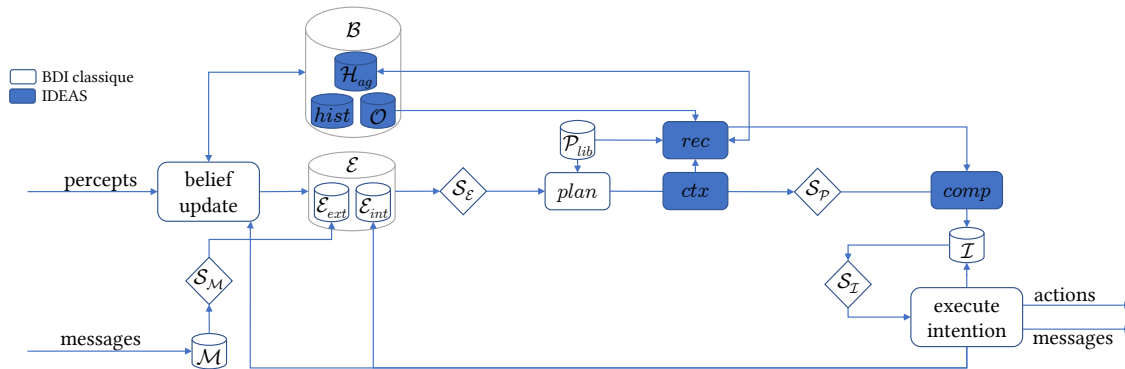


FIGURE 4.1 : Cycle de raisonnement simplifié d'un modèle BDI.

cution de la fonction `buf` génère un évènement. Les évènements générés à cause de l'ajout/suppression des croyances à partir de la perception de l'environnement sont appelés des *évènements externes*  $\mathcal{E}_{ext} \subseteq \mathcal{E}$ . Les *évènements internes*  $\mathcal{E}_{int} \subseteq \mathcal{E}$  ( $\mathcal{E} = \mathcal{E}_{int} \cup \mathcal{E}_{ext}$ ) ont une intention qui leur est associée (l'intention qui a provoqué la génération de l'évènement). Pour donner un exemple, considérons un agent qui perçoit l'état d'une porte (identifiée comme `door_2`). Il va acquérir une nouvelle croyance, et l'évènement généré serait :

$$\langle +\text{closed}(\text{door\_2}), [\text{source}(\text{percept})], \emptyset \rangle$$

La syntaxe utilisée ici est fortement inspirée de *Jason*, et utilisée pour illustration.

**Gestion des messages ( $\mathcal{M}, \mathcal{S}_M$ )** Les agents du même système, ou les autres cohabitants du même environnement physique sont aussi des sources d'information importantes pour un agent. L'architecture globale de l'agent doit avoir un composant qui récupère les messages reçus (l'ensemble  $\mathcal{M}$ ). De plus, étant donné qu'un seul message peut être traité dans chaque cycle de raisonnement, dans de nombreux scénarios, les agents voudront donner des priorités à certains messages. L'implémentation BDI de référence traite (par défaut) les messages dans l'ordre dans lequel ils ont été reçus, et accepte simplement tous les messages de tous les agents. De plus, elle donne la possibilité au développeur de personnaliser la politique de réception des messages, et déterminer quels agents peuvent fournir des informations, un savoir-faire, ou envoyer des sollicitations pour la réalisation de tâches, à un agent

particulier. Cette fonctionnalité est implémentée par la fonction de sélection de messages .

Les étapes précédentes concernaient la mise à jour par l'agent de ses croyances sur le monde et sur les autres agents. Les étapes qui suivent décrivent l'interprétation des messages et des perceptions de l'agent.

**Sélection d'évènement ( $\mathcal{E}$ ,  $\mathcal{S}_{\mathcal{E}}$ )** Les agents BDI gèrent continuellement des évènements, qui représentent soit des changements perçus dans l'environnement, soit des changements dans les objectifs de l'agent. Dans chaque cycle de raisonnement, un seul évènement en attente sera traité. Il peut y avoir plusieurs évènements en attente dans l'ensemble  $\mathcal{E}$ , par exemple, lorsque divers aspects de l'environnement ont récemment changé et que l'agent n'a pas passé suffisamment de cycles de raisonnement pour les gérer tous. Par conséquent, il faut permettre à l'agent de sélectionner un évènement particulier à traiter dans un cycle de raisonnement. Ceci est fait par une fonction de sélection d'évènement spécifique à l'agent ( $\mathcal{S}_{\mathcal{E}}$ ), qui peut également être personnalisée pour tenir compte des priorités spécifiques à chaque agent en fonction de l'application. Dans certaines applications, tous les évènements ont la même importance pour un agent. L'implémentation par défaut de la fonction  $\mathcal{S}_{\mathcal{E}}$  consiste arbitrairement à sélectionner le premier évènement de la liste. Si le développeur choisit de ne pas personnaliser la fonction de sélection d'évènement, le composant  $\mathcal{E}$  peut donc être considéré comme une simple file d'évènements.

Les prochaines étapes du cycle de raisonnement considèrent qu'il existe un évènement sélectionné, qui va être supprimé de l'ensemble des évènements.

**Unification des évènements avec des plans ( $\mathcal{P}_{lib}$ , *plan*)** Maintenant que l'agent a un évènement sélectionné, il doit trouver un plan qui lui permettra d'agir pour traiter cet évènement. La première chose à faire est de trouver dans la librairie de plans ( $\mathcal{P}_{lib}$ ) tous les plans qui sont pertinents pour l'évènement donné. Ceci est fait par la fonction de planification  $\mathcal{P}$  qui récupère tous les plans de  $\mathcal{P}_{lib}$  qui ont une précondition pouvant être unifiée avec l'évènement sélectionné. Par exemple, supposons que l'évènement sélectionné était

$\langle +closed(door\_2) \ [source(percept)], \ \emptyset \rangle$

L'agent peut alors passer à l'étape suivante où il continuera à filtrer l'ensemble des plans applicables parmi l'ensemble des plans pouvant être unifiés avec cet évènement.

**Choix des plans applicables (*ctx*)** Dans l'étape précédente, l'agent a sélectionné les plans

ayant des préconditions qui peuvent être unifiées avec l'évènement généré. Ainsi, même s'il a sélectionné les plans les plus pertinents, une partie de ces derniers peut ne pas être *applicable*, c.-à-d. qu'il doit utiliser un plan qui, compte-tenu de son savoir-faire et de ses croyances actuelles, semble avoir une chance de réussir. Pour ce faire, il doit vérifier, à l'aide de la fonction *ctx*, si le contexte de chacun des plans pertinents est considéré comme propice.

**Sélection d'un plan applicable ( $\mathcal{S}_O$ )** Dans cette étape, tous les plans sélectionnés sont appropriés dans le sens où l'exécution de l'un ou l'autre suffira pour traiter l'évènement sélectionné. Par conséquent, l'agent doit choisir un de ces plans applicables et s'engager à l'exécuter. S'engager à exécuter un plan signifie effectivement que l'agent a l'intention de poursuivre la séquence d'actions déterminée par ce plan. La sélection du plan particulier parmi l'ensemble des plans applicables qui sera inclus dans l'ensemble  $\mathcal{I}$  est implémentée par la fonction  $\mathcal{S}_O$ . Comme c'est le cas pour les autres fonctions de sélection,  $\mathcal{S}_O$  est aussi personnalisable. Son comportement prédéfini consiste à choisir un plan applicable en fonction de l'ordre dans lequel ils apparaissent dans  $\mathcal{P}_{lib}$ . Pour étendre ce comportement, le développeur peut, par exemple, sélectionner un plan qui satisfait d'autres évènements à la fois dans l'ensemble  $\mathcal{E}$ .

**Sélection des intentions ( $\mathcal{S}_I$ )** Dans cette étape, l'agent a ajouté dans l'ensemble de ses intentions  $\mathcal{I}$  la séquence d'actions du plan sélectionné par  $\mathcal{S}_O$  pour traiter l'évènement. En outre, comme nous l'avons mentionné ci-dessus, un agent a généralement plus d'une intention dans l'ensemble  $\mathcal{I}$ , chacune représentant une mission différente. Toutes sont en compétition pour l'attention de l'agent, ce qui signifie que chacune d'entre elles peut potentiellement être exécutée. Cependant, dans un même cycle de raisonnement, au plus une intention est exécutée. Par conséquent, la prochaine décision à faire avant que l'agent puisse agir est de choisir une intention particulière parmi celles qui sont actuellement prêtes à être exécutées. Nous utiliserons à nouveau une fonction de sélection d'intention  $\mathcal{S}_I$  pour faire ce choix. Comme dans les autres cas, l'idée est que cette décision est typiquement propre à l'agent, et les détails de la personnalisation sont délégués au développeur. En fonction de l'application, atteindre certains objectifs sera plus urgent que d'autres, donc le choix de la prochaine intention est très important pour l'efficacité de l'agent dans son environnement.

Dans un cycle de raisonnement, il y a trois actions principales qu'un agent fait :

- mettre à jour ses informations sur le monde et les autres agents,

- gérer un des évènements,
- agir sur l’environnement en *exécutant* une de ses intentions.

**Exécution des intentions** Dans cette dernière étape, l’agent agit sur l’environnement à travers l’exécution d’actions, l’envoi de messages, la mise à jour des intentions, ou la mise à jour des évènements internes. Il faut rappeler que les évènements internes ont une intention qui leur est associée, c.-à-d. résultent d’un changement des objectifs de l’agent, alors que les évènements externes sont générés à partir de la perception de l’environnement.

## 4.2 Présentation du modèle IDEAS

Le modèle présenté ici est une extension de l’architecture BDI décrite ci-dessus, basée sur le modèle d’identité de [Burke \[2006\]](#) illustré dans la Figure 3.5 (page 41). L’essence de cette extension consiste à ajouter deux processus :

- un processus de construction et maintenance de l’identité de l’agent, qui se base sur le *comparateur d’identité*. Selon [Burke \[2006\]](#), dans une situation d’interaction, le processus de comparaison d’identité d’un agent évalue les perceptions de son environnement, et celles tirées des retours d’interaction avec les autres agents, et les compare avec sa propre identité de référence pour la préserver, la contrôler, ou l’adapter ;
- un processus d’engagement qui prend des décisions visant à préserver l’identité sociale d’un agent. Ce processus est principalement utilisé lors de l’interaction avec les autres agents cohabitants dans l’environnement.

Nous présentons dans ce qui suit les adaptations apportées à l’architecture et au cycle de raisonnement du modèle BDI de base, telles qu’elles permettent à un agent d’avoir une représentation de son identité, et de s’engager à des choix sociaux raisonnables.

### 4.2.1 Architecture d’agent

Le modèle IDEAS que nous proposons est modélisé par un tuple (illustré dans la Figure 4.2) :

$$\mathcal{X}_{bdi} = \langle \mathcal{B}, \mathcal{M}, \mathcal{E}, \mathcal{P}_{lib}, \mathcal{I}, plan, ctx, \mathcal{S}_{\mathcal{M}}, \mathcal{S}_{\mathcal{E}}, \mathcal{S}_{\mathcal{O}}, \mathcal{S}_{\mathcal{I}} \rangle$$

Avec :

- $\mathcal{B} = hist \cup \mathcal{O}$  est l’ensemble des croyances de l’agent, composé de :

- $hist$  est l'ensemble contenant l'historique des actions exécutées et des messages échangés avec les autres agents jusqu'au moment actuel. Comme précisé dans la Section 3.2, l'identité personnelle est constituée par la continuité au fil du temps des propriétés physiques et mentales qui forment un individu. Ce composant est implémenté dans plusieurs architectures d'agent, comme par exemple MicroPSI [Bach, 2003] via les composants *Short Term Memory*, *Long Term Memory*, et *Memory Maintenance* qui capturent l'historique des interactions d'un agent avec son environnement;
- $\mathcal{O} = \mathcal{O}_{env} \cup \mathcal{O}_{ag}$  est l'ensemble des observations qui reflète les perceptions de l'agent sur les propriétés physiques de son environnement et sur les actions physiques et sociales des autres agents.
- $\mathcal{M}$  est une boîte de réception qui contient l'ensemble des messages reçus par l'agent;
- $\mathcal{E} = \mathcal{E}_{int} \cup \mathcal{E}_{ext}$  est l'ensemble des événements internes et externes à traiter par l'agent. Il faut noter que les désirs, qui capturent les motivations de l'agent, sont implicitement inclus dans l'ensemble des événements;
- $\mathcal{P}_{lib}$  est la librairie de plans qui capture le savoir-faire de notre agent IDEAS (plus d'informations sur les techniques de planification dans le chapitre suivant);
- $\mathcal{I}$  est l'ensemble des intentions, qui capturent les décisions de l'agent;
- $plan$  est la fonction de planification qui s'occupe d'unifier un événement sélectionné avec des plans de  $\mathcal{P}_{lib}$ ;
- $ctx$  est un gestionnaire de contexte qui retourne un score qui vérifie l'applicabilité et le degré de réussite d'un plan (unifié avec un événement) étant donné les contraintes de l'environnement (d'autres éléments à prendre en compte comme les capacités techniques de l'agent sont détaillés dans le chapitre suivant)

$$ctx : \mathcal{P}_{lib} \times 2^{\mathcal{O}_{env}} \rightarrow \mathbb{R}$$

Par exemple, un drone ne devrait pas pouvoir exécuter l'action `fly` quand il y a trop de vent;

- $\mathcal{S}_{\mathcal{M}}, \mathcal{S}_{\mathcal{E}}, \mathcal{S}_{\mathcal{O}}, \mathcal{S}_{\mathcal{I}}$  sont des fonctions de sélection de messages, d'évènements, de plans, et d'intentions.

**Fonctions de sélection** Nous avons vu que l'architecture BDI de base contient quatre fonctions de sélection, ayant chacune un comportement par défaut.  $\mathcal{S}_{\mathcal{M}}$  sélectionne les



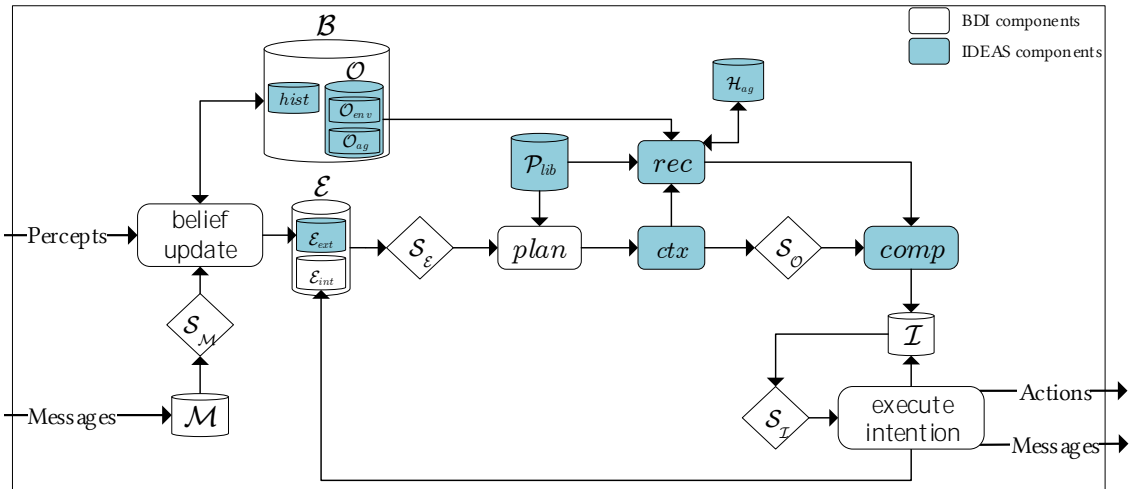


FIGURE 4.2 : Un modèle abstrait d'IDEAS

messages depuis l'ensemble  $\mathcal{M}$  dans l'ordre de leur réception. De même,  $\mathcal{S}_{\mathcal{E}}$  sélectionne le premier évènement ajouté à l'ensemble  $\mathcal{E}$ .  $\mathcal{S}_{\mathcal{O}}$  sélectionne les plans selon leur ordre d'apparition dans  $\mathcal{P}_{lib}$ .  $\mathcal{I}$  est un ensemble d'intentions, représentées comme un ensemble d'actions à exécuter. Par défaut, l'agent accorde la même attention à toutes les intentions.  $\mathcal{S}_{\mathcal{I}}$  sélectionne et retire la première action de la première intention dans  $\mathcal{I}$ , et déplace cette intention en fin de l'ensemble.

Dans notre modèle IDEAS, nous personnalisons les fonctions de sélection, afin que les messages, les évènements, et les intentions associés à des agents partageant ses identités sociales soient plus prioritaires que ceux des agents appartenant à un *out*-groupe. En ce qui concerne la sélection des plans, il faut rappeler d'abord qu'un agent :

- n'est pas capable d'exécuter seul toutes les actions constituant un plan,
- sera éventuellement amené à déléguer l'exécution de quelques actions ou sous-plans à d'autres agents.

Nous proposons qu'un agent sélectionne, en fonction de son identité, les plans pouvant être satisfaits par des agents partageant ses identités sociales. Par exemple, des agents avec lesquels il a déjà coopéré avec réussite, ou des agents qui lui ont déjà fourni les mêmes actions dans le plan à sélectionner. Il existe d'autres critères de sélection pris en compte dans le modèle IDEAS telles que la qualité de service, l'utilité, et le coût.

Dans ce qui suit, nous allons présenter les composants permettant d'implémenter le processus de construction et de maintenance de l'identité de l'agent, et le processus d'engagement.

### 4.2.2 Construction et maintenance d'identité

Nous considérons que l'identité d'un agent est composée de :

- l'ensemble de ses croyances sur l'état de l'environnement, et sur les autres agents ( $\mathcal{B}$ ),
- les désirs qui le motive à agir sur l'environnement et interagir avec les autres agents ( $\mathcal{E}_{ext}$ ),
- les hypothèses sur les tâches poursuivies et les appartenances et activités sociales des autres agents ( $\mathcal{H}_{ag}$ ),
- le processus de comparaison de l'identité ( $comp$ ).

Ainsi, l'identité de l'agent est représentée par :

$$id = \langle \mathcal{B}, \mathcal{E}_{ext}, \mathcal{H}_{ag}, comp \rangle$$

- $comp$  est le *comparateur d'identité*, défini comme suit :

$$comp : id_{new} \times id_{standard} \rightarrow \mathbb{R}$$

Après la sélection d'un évènement externe, ce processus estime le coût/utilité de l'identité potentiellement adoptée  $id_{new}$  par rapport à la norme d'identité de l'agent  $id_{standard}$ .  $comp$  fonctionne comme un système de prise de décision, filtrant les évènements externes à traiter en fonction de leurs conséquences sur l'identité de l'agent. De cette façon, les interférences négatives peuvent être évitées pour préserver son identité ou négociées pour l'adapter.

- $\mathcal{H}_{ag}$  est l'ensemble des hypothèses sur les adhésions des autres agents, les coalitions formées et les tâches poursuivies. Ces hypothèses sont mises à jour dans chaque cycle de raisonnement.

La question qui se pose est de savoir comment un agent peut faire des hypothèses sur les autres agents pour pouvoir décider ceux qui font partie de ses *in*-groupes, et *out*-groupes, et par conséquence raisonner sur les évènements générés à partir de leurs sollicitations. Comme nous l'avons déjà mentionné (Section 3.6), l'auto-conception de l'identité de chaque agent dépend en partie de ce qu'il *pense* des autres. Plus des agents partagent les mêmes significations intersubjectives, plus est efficace et coordonné leur fonctionnement dans des tâches collectives. L'intersubjectivité devient ainsi un processus au carrefour des composants de l'architecture de notre agent IDEAS. Dans cette thèse, nous traduisons ce processus par un mécanisme de reconnaissance d'intention.

**La reconnaissance d'intention** De nombreux sous-domaines contribuent au domaine de recherche dynamique de la reconnaissance du comportement d'autres agents [Sukthankar et al., 2014] : la reconnaissance d'intention, la reconnaissance d'objectifs, la reconnaissance de plan, la reconnaissance d'activité, l'abduction mentale, l'interprétation d'intention, etc. Mais ils partagent des défis et des applications similaires. Cette division n'est pas destinée à séparer la reconnaissance en sous-domaines de recherche isolés, ils sont étroitement liés et ils peuvent se transformer les uns en les autres. Mieux que cela, la résolution de l'un peut résoudre les autres. Par exemple, de nombreuses travaux utilisent le but et l'intention de manière interchangeable [Sadri, 2010]. De même, la reconnaissance d'intention et la reconnaissance de plan peuvent avoir la même signification [Sindlar et al., 2008, Pereira, 2013], alors que d'autres travaux considèrent la reconnaissance de but/intention comme un cas particulier de reconnaissance de plan [Blaylock and Allen, 2006]. La reconnaissance d'intention peut être utile pour de nombreux domaines de recherche tels que la robotique [Parker, 1998], l'intelligence ambiante [Rashidi et al., 2011], la vision par ordinateur [Tavakkoli et al., 2007], l'interaction homme-machine, la surveillance du trafic, les applications militaires, les jeux vidéo et bien d'autres [Sukthankar et al., 2014, Pereira, 2013].

L'architecture de notre agent IDEAS inclut un système de reconnaissance d'intention multi-agent. Ce système a besoin de la librairie de plans  $\mathcal{P}_{lib}$  pour "reconnaître" les plans suivis par les agents, du gestionnaire de contexte  $ctx$  pour favoriser ou pénaliser des hypothèses selon l'état de l'environnement, des observations ( $\mathcal{O}$ ) qui servent comme entrée du système à partir desquelles il va chercher des plans qui les expliquent, et une fonction de reconnaissance de plan ( $rec$ ). Ce système est représenté par un tuple :

$$\mathcal{R} = \langle \mathcal{P}_{lib}, ctx, \mathcal{O}, \mathcal{H}_{ag}, rec \rangle$$

tel que

$$rec : 2^{\mathcal{H}_{ag}} \times 2^{\mathcal{O}} \rightarrow \mathcal{H}_{ag}$$

est la fonction de reconnaissance de plan, qui met à jour l'ensemble des hypothèses  $\mathcal{H}_{ag}$  en fonction des nouvelles perceptions (l'implémentation de cette fonction est détaillée et illustrée dans le chapitre suivant). En utilisant ce système de reconnaissance d'intention, un agent est capable d'identifier :

- les tâches qu'il poursuit,
- les coalitions auxquelles il appartient,

ce qui lui permet d'avoir une représentation de son identité, c'est-à-dire ce qu'il doit

préservé avant s'engager à coopérer avec d'autres agents.

### 4.2.3 Processus d'engagement centré identité

La coopération fonctionne à travers l'allocation de certaines tâches par un agent à un autre agent sollicité via une requête répondant à un engagement [Castelfranchi and Falcone, 2002]. Dans le modèle BDI de base, l'agent se base sur l'applicabilité d'un plan (unifié avec un évènement) pour décider de son engagement à le suivre. Cet engagement se traduit simplement par l'ajout de la séquence des actions du plan comme une intention dans l'ensemble  $\mathcal{I}$ . Une fois que l'agent est passé à l'étape de l'exécution des intentions, le modèle de base propose plusieurs stratégies d'engagement (*blind*, *open-minded*, *single-minded*, présentées dans la Section 2.2.3 en page 15). Ces dernières sont mises en place lorsqu'un des plans de l'agent n'aboutit pas : est-ce qu'il choisit un autre plan pour traiter l'évènement ? est-ce qu'il abandonne le traitement de l'évènement ? est-ce qu'il garde le même plan tant que l'objectif n'est pas satisfait ? Les stratégies proposées dans le modèle BDI de base répondent à ces questions.

Dans le contexte où un agent est sollicité par d'autres agents du même SMA, on peut comprendre qu'ils sont pré-programmés pour qu'ils s'entre-aident, et qu'ils partagent probablement les mêmes objectifs. La gestion de conflits entre ces agents sera en grande partie effectuée pour coordonner les actions des agents, et éviter les redondances et la re-planification des tâches. Cependant, lorsqu'un agent est sollicité par d'autres qui ne sont pas forcément du même SMA que lui, ou avec lesquels il n'a jamais coopéré pour participer dans la réalisation de certaines actions ou dans l'allocation de certaines tâches, évaluer l'applicabilité d'un plan n'est pas suffisante pour l'adopter et s'engager à le suivre. Un agent ne doit mettre dans ses intentions que des plans n'interférant pas négativement avec ses autres intentions, et des plans favorisant la cohésion de SMA. Plus généralement, tout engagement de l'agent doit être pris de manière à ce qu'il conserve son identité, et adapter cette dernière pour qu'elle profite des changements du contexte physique et sociale de l'agent.

Ainsi, nous avons besoin dans notre modèle IDEAS d'un processus d'engagement permettant aux agents de choisir les meilleures options qui réduisent les interférences négatives entre l'exécution des actions des agents appartenant à des SMA coexistants. Nous proposons de modéliser ce processus d'engagement en tant que DCOP (*Distributed Constraint Optimization Problem*) [Modi et al., 2005]. DCOP est un formalisme pour la résolution de problèmes distribués en maximisant le bien-être social d'un groupe d'agents. Ce formalisme est appliqué dans une large gamme d'applications multi-agents, qui par-

tagent des caractéristiques communes avec notre mécanisme d'engagement basé sur l'identité : les agents choisissent un petit nombre de possibilités ; l'efficacité du groupe dans son ensemble dépend des choix des agents en interaction ; chaque agent ne peut communiquer qu'avec quelques autres agents de son voisinage local ; et le problème doit être résolu de manière entièrement distribuée.

**Distributed Constraint Satisfaction Problem** Un problème d'optimisation de contraintes distribuées (DCOP) modélise un certain nombre d'agents devant prendre certaines décisions. Ces agents ont un nombre fini de choix possibles pour chaque décision. Ensuite, les synergies entre les décisions sont capturées à l'aide de contraintes : des fonctions qui spécifient une valeur différente pour chaque combinaison de choix d'un sous-ensemble de décisions des agents. Par conséquent, l'objectif d'un algorithme DCOP est de trouver le choix pour chaque décision qui maximise ou minimise la somme globale des fonctions. Un DCOP est représenté par un tuple  $\langle \mathcal{A}, \mathcal{X}, \mathcal{D}, \mathcal{C}, \alpha \rangle$ , tel que :

- $\mathcal{A} = \{A_1, \dots, A_{|\mathcal{A}|}\}$  est un ensemble d'agents impliqués dans le problème ;
- $\mathcal{X} = \{X_1, \dots, X_n\}$  est un ensemble de variables de décision, où chaque variable est contrôlée par un agent de l'ensemble  $\mathcal{A}$  ;
- $\mathcal{D} = \{D_X | X \in \mathcal{X}\}$  est un ensemble de domaines finis de variables, tel que chaque variable  $X_i$  prend des valeurs dans  $D_{X_i}$  ;
- $\mathcal{C} = \{f_1, \dots, f_m\}$  est un ensemble de contraintes, où chaque contrainte est une fonction définie sur un sous-ensemble  $\chi_i = \{X_i, \dots, X_j\} \subseteq \mathcal{X}$ . Chaque fonction  $f_i$  affecte une valeur d'utilité pour chaque combinaison de valeurs de  $\chi_i$ . Plus formellement,  $f : D_{X_i} \times \dots \times D_{X_j} \rightarrow \mathbb{R} \cup \{-\infty\}$ .
- $\alpha : \mathcal{X} \rightarrow \mathcal{A}$  est une fonction qui fait correspondre à chaque variable de décision son agent de contrôle. Un agent peut contrôler plusieurs variables, tandis qu'une variable est contrôlée par un seul agent.

Une solution d'un problème DCOP est d'avoir une affectation de toutes les variables de décision qui maximise la fonction *objectif*, qui est la fonction qui capture l'utilité agrégée d'une instantiation complète  $\mathcal{X}^*$  pour toutes les variables  $X_i$  :

$$\mathcal{X}^* = \operatorname{argmax} \sum_{i=1}^m f_i(\chi_i)$$

Une fois que le processus d'engagement est modélisé en tant que DCOP, il peut être résolu par des algorithmes dédiés qui garantissent que les agents s'engagent à répondre aux sollicitations des autres tout en préservant leurs identités.

### 4.2.3.1 Modèle d'engagement

Le processus d'engagement est mis en place lorsque plusieurs agents sont sollicités à réaliser des parties différentes d'un plan. Nous l'utilisons d'une manière distribuée dans chacun de ces agents, ayant comme finalité de leur permettre de décider quelles parties du plan ils devront choisir pour que chacun puisse minimiser les interférences négatives entre ses choix et l'ensemble de ses évènements et intentions, autrement dit, préserver son identité. Ce processus peut être modélisé en DCOP comme suit :

- $\mathcal{A} = \{a_1, \dots, a_n\}$  est un ensemble d'agents auxquels on demande de participer à la réalisation des parties d'un plan ;
- $\mathcal{X} = \{x_1, \dots, x_m\}$  est l'ensemble des variables correspondant aux parties du plan en question ;
- $\mathcal{D} = \{D_x | x \in \mathcal{X}\}$  est un ensemble d'agents qui ont accepté de fournir une partie du plan  $\mathcal{X}$  ;
- $\mathcal{C} = \mathcal{I}_S \cup \Upsilon$ , telle que  $\mathcal{I}_S$  est un ensemble de **fonctions de score d'identité** (détaillée dans la section suivante), et  $\Upsilon = \{v(x_i), \dots, v(x_m)\}$  est un ensemble des utilités associées à chaque partie du plan dans  $\mathcal{X}$ .

L'objectif est de trouver l'ensemble des affectations  $\mathcal{X}^*$  pour les actions  $\mathcal{X}$ , en maximisant les utilités des actions et les scores d'identité.

### 4.2.3.2 Fonction de score d'identité

Dans notre modèle d'engagement, les contraintes sont binaires, c'est-à-dire qu'elles se produisent entre des paires d'actions  $\{x_i, x_j\}$ . Une fonction  $I_{i,j} \in \mathcal{I}_S : 2^{D_{x_i}} \times 2^{D_{x_j}} \rightarrow \mathbb{R}^+$  représente le score d'identité correspondant à chaque combinaison d'agents ayant accepté de fournir les actions concernées dans  $I_{i,j}$ . Par exemple pour une action  $x_1$  qui peut être exécutée par des agents dans le domaine  $\{a_1, a_2\}$ , et une action  $x_2$  qui a comme domaine  $\{a_1, a_5, a_4\}$ ,  $I_{1,2}$  correspond à :  $I_{1,2} = \{\{x_1 = a_1, x_2 = a_1\}, \{x_1 = a_1, x_2 = a_5\}, \{x_1 = a_1, x_2 = a_4\}, \{x_1 = a_2, x_2 = a_1\}, \{x_1 = a_2, x_2 = a_5\}, \{x_1 = a_2, x_2 = a_4\}\}$ .

Chaque fonction  $I_{i,j}$  calcule le score global d'identité d'un ensemble d'agents et leurs actions fournies en additionnant les scores d'identité locaux  $\delta$  calculés par chaque action

impliquée dans  $I_{i,j}$ , qui est définie comme suit :  $I_{i,j} = \delta_{i \rightarrow j} + \delta_{j \rightarrow i}$  telle que :

$$\delta_{i \rightarrow j} = \begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \\ \omega_4 \end{bmatrix} \left[ |eff^+(\mathcal{T}(x_j))| \quad -|eff^-(\mathcal{T}(x_j))| \quad \frac{v(x_j)}{v(x_i)} \quad trust(x_j) \right] \quad (4.1)$$

- $x_i$  est la variable d'action qui calcule le score local d'identité,
- $x_j$  est la variable d'action pour laquelle  $x_i$  calcule le score local d'identité,
- $trust(x_j)$  renvoie un score correspondant au degré d'acointance calculé pour l'agent fournissant  $x_j$ .

Ce degré varie selon que les agents fournissant les deux actions appartiennent au même système ou groupe social ou ont déjà collaboré avec succès,

- $\mathcal{T}(x_j)$  est la tâche reconnue par le système de reconnaissance de plan étant donné  $x_i$  et  $x_j$  comme observations,
- $|eff^+|$  et  $|eff^-|$  sont respectivement le nombre d'actions avec des effets positifs et négatifs sur l'environnement par rapport aux désirs et aux intentions de l'agent qui les évalue.

Par exemple, si l'objectif est de maintenir la température d'une chambre à une certaine valeur, l'ouverture de la fenêtre est considérée comme une action indésirable ayant des effets négatifs.

- $v(x_i)$  et  $v(x_j)$  sont respectivement les utilités des actions  $x_i$  et  $x_j$ .

Les intervalles de ces termes sont normalisés et priorisés en utilisant des poids ( $\omega_1, \omega_2, \omega_3, \omega_4$ ). Chacun des poids joue sur le comportement global des agents (plus de détails seront fournis dans la section suivante) tels qu'on peut favoriser l'évitement des interférences négatives ou l'affectation des actions au maximum des agents d'un même système.

#### 4.2.4 Cycle de vie

Le cycle de vie illustré dans la Figure 4.2 est un adaptation d'un cycle de vie BDI, avec l'ajout des composants et des processus décrits ci-dessus. Un agent BDI commence par mettre à jour ses croyances une fois qu'il a obtenu de nouvelles informations (perceptions, messages).

**Reconnaissance d'intention** Une fois que la fonction de  $\mathcal{S}_M$  sélectionne un message de l'ensemble  $\mathcal{M}$ , l'agent doit comprendre la signification intentionnelle derrière

ce message. Un agent peut, par exemple, recevoir des messages d'information lui permettant de mettre à jour ses croyances sur l'environnement et sur les autres agents. Il peut aussi recevoir des sollicitations pour exécuter une tâche, un plan, ou une action dans un cadre coopératif. Notre proposition d'ajouter cette étape de reconnaissance d'intention est motivé par le fait que cette dernière n'est, généralement, pas fournie comme information dans les requêtes des agents. Par exemple, un agent recevra une requête d'ouvrir une porte d'une chambre, sans être informé si l'intention derrière l'exécution de cette action est dans un contexte de sécurité ou de confort des occupants de la chambre. Ainsi, l'agent doit reconnaître l'intention derrière chaque sollicitation pour la prendre en compte lors de la décision de s'engager à répondre favorablement ou pas.

**Mise à jour de l'identité** Après avoir sélectionné un évènement, l'agent a ajouté (dans cette étape) une nouvelle croyance à l'ensemble  $\mathcal{B}$  et, éventuellement, un nouveau désir à l'ensemble des évènements  $\mathcal{E}$ , et une nouvelle trace dans son historique d'interactions *hist*. De plus, il doit mettre à jour les hypothèses qu'il a sur les coalitions formées, les appartenances des agents de son voisinage et leurs tâches poursuivies. Cette mise à jour se base sur le même mécanisme de reconnaissance d'intention utilisé dans la troisième étape. Avec ces changements et mises à jour, l'agent a une nouvelle identité.

Après la mise à jour de son identité, l'agent peut maintenant agir sur un évènement qu'il a sélectionné dans une étape précédente, à travers le choix d'un plan applicable parmi tous les plans qui peuvent être unifiés avec cet évènement.

**Engagement** Dans cette étape, l'agent a un plan applicable qui a une grande chance réussite étant donné l'état actuel de l'environnement. Il a déjà reconnu l'intention derrière l'évènement que le plan sélectionné traite. S'engager à réaliser ce plan se traduit par l'ajouter dans son ensemble d'intentions, et par conséquent mettre à jour l'identité de l'agent. Avant de faire cette décision, nous utilisons le processus de comparaison d'identité (*comp*) pour examiner si la nouvelle identité va remplacer l'identité standard ou bien s'il faut préserver l'identité standard et rejeter automatiquement le plan déjà sélectionné, et le remplacer par un autre qui peut satisfaire l'évènement à traiter. C'est-à-dire, l'agent doit vérifier si l'exécution des actions du plan choisi n'interférera pas négativement dans l'exécution des intentions dans  $\mathcal{I}$ . Une fois que cette vérification faite, l'agent peut ajouter le plan auquel il peut s'engager dans l'ensemble de ses intentions pour pouvoir les exécuter.



## 4.3 Synthèse

Ce chapitre a présenté le modèle IDEAS et a parcouru les différentes étapes de son cycle de vie. Ce dernier étend un cycle de vie d'un agent BDI typique en intégrant des processus et des éléments de la théorie de l'identité, pour permettre à l'agent de raisonner sur les sollicitations des autres. Le modèle IDEAS est doté de la capacité de :

- anticiper les interférences négatives qui peuvent surgir entre les intentions de l'agent et la réalisation d'une action ou l'adoption d'une tâche d'un autre agent sollicitateur,
- favoriser les sollicitations qui interféreraient positivement dans la satisfaction des objectifs de l'agent,
- privilégier la participation dans la réalisation des tâches impliquant des agents du même SMA ou des anciens coéquipiers avec lesquels l'agent IDEAS a déjà coopéré avec succès.

À travers ces capacités, un agent IDEAS peut avoir le contrôle sur son identité en choisissant de la conserver (en restant prudent et ne pas prendre des risques vis-à-vis des agents préalablement inconnus), ou de l'adapter (en négociant le choix des actions qu'il peut fournir, et en étendant son réseau d'accointances).

Le choix d'un modèle BDI pour illustrer le processus de construction et maintenance d'identité, et le processus d'engagement ne signifie pas que notre contribution est dépendante de ce modèle. Nous sommes convaincus que les deux processus proposés peuvent s'implémenter dans toute architecture d'agent, ou même dans des architectures logicielles qui ne sont pas orientées agent, telle celle utilisée dans le projet de recherche qui finance cette thèse. Comme déjà mentionné, la quasi-totalité des architectures d'agents étudiées dans ce chapitre ont déjà quelques éléments de base (e.g. connaissances sur l'environnement, connaissances sur les autres agents, objectifs, ...), et elles sont toutes extensibles pour prendre en compte la description de l'identité d'un agent.

Dans l'état actuel de notre proposition, l'identité d'un agent est principalement basée sur les coalitions auxquelles il appartient et les objectifs qu'il veut atteindre. Or, l'identité d'un individu est beaucoup plus riche que ça, surtout en terme de mécanismes sociaux (mobilité individuelle [Ellemers et al., 1997], caractère distinctif personnel [Sedikides and Strube, 1997], créativité sociale [Jackson et al., 1996], compétition sociale, gestion des conflits, ...). De plus, nous n'avons pas exploité le concept d'identité collective, qui est dans un autre niveau conceptuel car il s'agit d'un concept d'un groupe d'agents, tandis que l'identité personnelle et sociale définissent un individu.



# Implantation

---

<b>5.1</b>	<b>Contexte applicatif</b> . . . . .	<b>66</b>
5.1.1	Défis . . . . .	66
5.1.2	Architecture de l' <i>avatar</i> . . . . .	68
<b>5.2</b>	<b>Étude de cas</b> . . . . .	<b>71</b>
5.2.1	Scénario . . . . .	71
5.2.2	La planification par réseaux hiérarchisés de tâches . . . . .	72
<b>5.3</b>	<b>La construction de l'identité d'un agent IDEAS</b> . . . . .	<b>77</b>
5.3.1	Le mécanisme de reconnaissance d'intention dans IDEAS . . . . .	79
5.3.2	Le processus de construction d'identité . . . . .	81
<b>5.4</b>	<b>Le processus d'engagement des agents IDEAS</b> . . . . .	<b>83</b>
<b>5.5</b>	<b>Synthèse</b> . . . . .	<b>86</b>

---

Après avoir détaillé l'implémentation de notre modèle IDEAS et des deux processus caractérisant notre contribution, nous présentons dans ce chapitre nos choix techniques pour implémenter ce modèle. Nous commençons par définir les défis relevés par un *avatar* et nous présentons son architecture. Un avatar qui est une entité logicielle représentant un objet physique sur le Web, développée dans le cadre du projet ANR ASAWoO. Ensuite, nous décrivons une étude de cas pour illustrer le déroulement des processus de construction d'identité d'un agent IDEAS et d'engagement centré identité.

## 5.1 Contexte applicatif

Le Web des Objets (WoT) est une extension de l'Internet des Objets qui a comme principale motivation l'utilisation des technologies, des standards, et des protocoles du Web pour permettre les interactions entre des objets physiques. Le premier avantage de cette extension est de viser à accéder à toutes sortes d'objets connectés par le biais de technologies et d'outils standard, réalisant ainsi l'interopérabilité entre ces objets et évitant les divergences entre les technologies propriétaires de leurs constructeurs. Il permet de réutiliser des plateformes éprouvées, des technologies et des techniques d'interaction avec les utilisateurs, de faciliter le développement des applications, et de réduire les délais de mise sur le marché. Deuxièmement, en donnant aux objets une existence en tant que ressources Web, le WoT vise à enrichir les possibilités des applications Web en comblant le fossé entre le monde virtuel et le monde physique. En effet, être capable de contrôler la température de sa maison intelligente depuis son téléphone portable est déjà une réalité, tout comme permettre à un pot de fleurs de faire un *tweet* quand son sol s'assèche. Inversement, représenter un objet physique en tant que ressource Web lui donne une existence dans le monde de l'information. Même les objets inanimés peuvent maintenant être associés un code QR, de sorte que des personnes peuvent poster des commentaires sur un restaurant qu'ils ont visité ensemble et récupérer tous ces commentaires dans le blog du restaurant. Des applications plus puissantes et utiles sont maintenant construites pour utiliser des objets connectés via des technologies Web. Mais la construction de telles applications reste toujours une activité manuelle.

### 5.1.1 Défis

Le travail réalisé dans thèse s'inscrit dans le cadre du projet ASAWoO<sup>1</sup> (*Adaptive Supervision of Avatar/Object Links for the Web of Objects*) financé par l'Agence Nationale pour la Recherche sous la référence <ANR-13-INFR-0012>. Ce projet a pour objectif l'extension des objets physiques en leur proposant des représentations virtuelles, tels que chaque objet-représentation forment une entité cyber-physique sur le Web, qui définit et incarne les comportements de haut niveau de l'objet. Pour ce faire, nous associons un *avatar*<sup>2</sup> [Jammont et al., 2014] à chaque objet physique qui le représente dans le monde virtuel. Cette représentation consiste à fournir aux constructeurs de matériels, aux développeurs de logiciels et aux utilisateurs finaux une abstraction qui rend les objets physiques acces-

1. <https://projet.liris.cnrs.fr/asawoo/>

2. Dans ce chapitre, agent et avatar seront utilisés d'une manière interchangeable.

sibles sur le Web, et étend leur statut et leurs capacités en des fonctionnalités homogènes et compréhensibles par l'utilisateur. L'avatar repose sur plusieurs fonctionnalités qui répondent à différents défis que nous détaillons dans la suite :

**Description sémantique des capacités et fonctionnalités** Les fonctionnalités (la couche applicative) et les capacités (la couche objet) décrites sémantiquement sont une caractéristique clé d'un avatar. Elles sont décrites sous forme d'ontologies en utilisant le langage OWL (Web Ontology Language), qui permet de traiter et de raisonner sur ces descriptions en utilisant des outils standards du Web sémantique (par exemple, des bases de données *triplestores*, et des moteurs de requêtes SPARQL).

**Adaptation au contexte** Les objets peuvent avoir des capacités hétérogènes à utiliser dans divers environnements. Par conséquent, un avatar doit pouvoir fonctionner sur différents objets, quel que soient leurs types et leurs fabricants. Ce défi est abordé dans le projet en concevant un modèle de contexte sémantique à plusieurs niveaux, ainsi que son moteur de raisonnement. Ce moteur peut répondre aux demandes concernant quatre objectifs d'adaptation :

- Où est-ce que le code de l'avatar doit être déployé ? par exemple, le code de l'avatar d'une porte ou une fenêtre ne sera pas déployé et exécuté comme le code de l'avatar d'une Raspberry Pi (le premier sur le cloud, alors que le deuxième est sur l'objet connecté lui même).
- Quelle interface de communication choisir ? par exemple, le protocole de communication change en fonction de la distance entre deux drones.
- Quelles fonctionnalités exécuter et exposer pour qu'elles soient exploitables par d'autres objets ? Par exemple, un drone n'exposera pas des capacités de prise de vidéo lorsqu'il n'a pas un espace de stockage suffisant.
- Quel objet physique choisir pour une tâche collective ? Par exemple, un robot agricole va choisir, selon la distance, le drone le plus proche, ou selon le niveau de batterie, le robot qui a le plus d'autonomie pour lui demander de prendre des photos aériennes d'une partie d'un champ.

**Services tolérants aux perturbations** Les mouvements libres des appareils mobiles (par exemple, les robots ou les drones) et la courte portée des interfaces de communication sans fil, accompagnées d'interférences radio, provoquent des perturbations fréquentes et imprévisibles. Ces perturbations peuvent également résulter de la désactivation ou la mise en veille des appareils en raison de leur puissance limitée. Des protocoles de routage et des mécanismes de transmission de données doivent

donc être conçus pour les réseaux qui souffrent de perturbations fréquentes et imprévisibles.

**Composition de services sémantiques** Les avatars exposent les fonctionnalités et les capacités que leurs objets peuvent fournir en tant que services Web RESTful. Pour collaborer avec d'autres avatars, ils doivent manipuler l'exécution de ces services sémantiques dans le cadre d'applications WoT (par exemple, la régulation de température). Cette manipulation consiste à découvrir et appeler des services Web sémantiques d'une manière automatique.

**Fonctionnalités collaboratives** Les avatars doivent être capables de répondre à des tâches complexes, qui dépassent la capacité individuelle de chacun, d'une manière collective. Ceci implique la nécessité de protocoles d'allocation de tâches, de négociation, et de coopération pour répondre aux besoins des utilisateurs de la manière la plus efficace, optimisée, et autonomie possible.

### 5.1.2 Architecture de l'*avatar*

Plusieurs composants dans l'architecture de l'avatar permettent de relever ces défis. Ces composants sont, entre autres un environnement cloud pour déployer le code de l'avatar des objets qui n'ont pas une grande capacité de calcul et de communication, des répertoires de pilotes pour permettre l'avatar à échanger des messages avec son objet physique correspondant, des ontologies pour permettre aux avatars de pourvoir gérer les capacités et les fonctionnalités qu'il peuvent fournir individuellement ou collectivement selon le contexte, des *marketplaces* pour permettre aux développeurs de fournir des applications de type Web des Objets. Le reste des composants est intégré dans l'architecture logicielle de l'avatar.

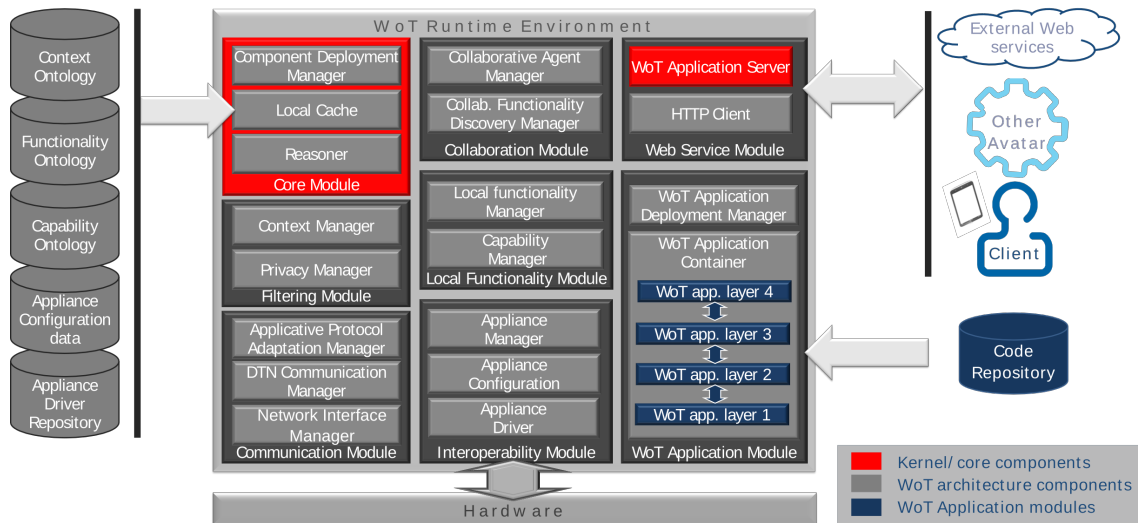


FIGURE 5.1 : L'architecture logicielle de l'avatar

Cette architecture est modélisée comme une plateforme de type OSGi (Open Service Gateway initiative) à base de composants (déployés sous forme de *bundles*) qui peuvent être dynamiquement installés, démarrés, arrêtés, mis à jours et désinstallés sans avoir besoin de redémarrer toute la plateforme. L'avatar consacre des composants spécifiques au contrôle de l'objet physique, tandis que d'autres implémentent le comportement autonome, adaptatif et collaboratif des avatars, assurant l'interopérabilité avec les objets et les logiciels externes qui utilisent les standards du Web. Il faut noter que l'environnement d'exécution est entièrement découplé de l'architecture logique de l'avatar. Ce dernier peut adapter dynamiquement la distribution de ses composants à différents emplacements (par exemple, objet, passerelle réseau local et cloud) pour améliorer leur exécution. L'architecture logicielle de l'avatar (que nous avons publiée dans [Mrissa et al., 2015]) est constituée de six modules fonctionnels, décrits ci-dessous.

**Module d'interopérabilité avec l'objet (*Interoperability Module*)** Ce module offre une interface uniforme pour interagir avec l'objet physique auquel l'avatar est associé. Il s'appuie sur un répertoire de fichiers et d'outils pour configurer l'objet (c'est-à-dire pour définir les méthodes à utiliser pour communiquer avec lui). À l'aide de ce module, l'avatar charge et utilise les pilotes pour échanger des messages avec l'objet physique. Il effectue aussi une introspection de l'objet pour découvrir ses capacités physiques.

**Noyau (*Core Module*)** Ce module contient d'autres composants que l'avatar réutilise dans plusieurs étapes de son cycle de vie. Le *reasoner* permet gérer tout ce qui concerne le raisonnement sémantique sur les ontologies (ontologies et requêtes

SPARQL). Le *functionality manager* se charge de décider quelles fonctionnalités déployer en fonction des capacités de l'objet dans le contexte actuel. Le *deployment manager* décide où et quand déployer le code des autres composants de l'avatar.

**Module de communication (*Communication Module*)** Ce module est chargé de sélectionner la bonne interface et le bon réseau (Wi-Fi, Bluetooth ou Zigbee) et les protocoles applicatifs (CoAP ou HTTP) en fonction des interfaces de communication disponibles et des besoins de performance (débit et consommation énergétique).

**Module de services Web (*Web Service Module*)** Ce module est conçu pour exposer les fonctionnalités d'un avatar en tant que services RESTful [Fielding, 2000]. Ainsi, d'autres avatars et applications Web peuvent découvrir et appeler ces services pour interagir avec l'avatar (et l'objet physique). Ce module de services Web permet également à un avatar de découvrir et d'invoquer des services distants proposés par des avatars ou des serveurs sur le Web. Il est aussi chargé de mettre en œuvre les processus de négociation inter-avatar, en utilisant un schéma de communication basé sur les bonnes pratiques utilisées dans le style architectural REST.

**Module d'applications WoT (*Wot Application Module*)** Les applications de type Web des Objets exposent le comportement de haut niveau d'un objet connecté. Ils fournissent aux utilisateurs des interfaces Web pour contrôler et interagir avec des objets physiques. Ces applications sont exécutées dans un conteneur d'application WoT, en partie distribué sur l'objet connecté et en partie sur l'infrastructure Web (comme un cloud) grâce au *Deployment Manager*.

Nous avons déployé dans l'architecture d'avatar du projet ASAWoO plusieurs composants du modèle IDEAS. Par exemple, le *module d'interopérabilité avec l'objet* implémente l'exécution des actions. Le *module de communication* implémente l'envoi/réception des messages. Les bibliothèques de plans sont implémentées en tant qu'ontologies basées sur le répertoire de fonctionnalités de l'avatar, alors que les croyances (observations et historique des actions) et les événements sont répertoriés dans d'autres ontologies dédiées. La fonction de *sélection de l'évènement* est implémentée dans le *context manager*, ce dernier contient aussi la *sélection de plans*.

**Module de collaboration** Une partie des travaux de cette thèse est implémentée dans ce module. Il s'agit du module qui s'occupe de définir quelle situation d'interaction établir avec chaque avatar du voisinage de son objet physique. Ainsi, c'est à travers ce modèle que l'agent est capable de former des coalitions, négocier pour le propriétaire de son objet physique, s'engager à accepter des sollicitations, etc. Nous intégrons la *fonction de reconnaissance d'intention*, et le *comparateur de l'identité*



dans le *module de collaboration*.

## 5.2 Étude de cas

Dans cette section, nous présentons une étude de cas pour illustrer notre contribution, et un des choix techniques que nous avons faits pour implémenter cette contribution.

### 5.2.1 Scénario

Considérons un scénario d'agriculture de précision (illustré en Figure 5.2) décrit plus en détail dans le chapitre d'ouvrage que nous avons co-rédigé avec les autres membres du projet ASAWoO [Médini et al., 2017]. Des réseaux de capteurs comprenant des anémomètres, des thermomètres et des pluviomètres sont déployés pour surveiller divers paramètres environnementaux liés aux applications agricoles. Des asperseurs automatiques fournissent de l'eau aux vignes. Des tags RFID sont utilisés pour identifier chaque rangée de vignes. Des *agribots* (robots agricoles) sont utilisés pour effectuer plusieurs tâches agricoles (fertilisation, taille, pulvérisation, désherbage, traitement d'image, ...). Des drones survolent les champs de vigne pour observer leurs états et effectuer des tâches aériennes.

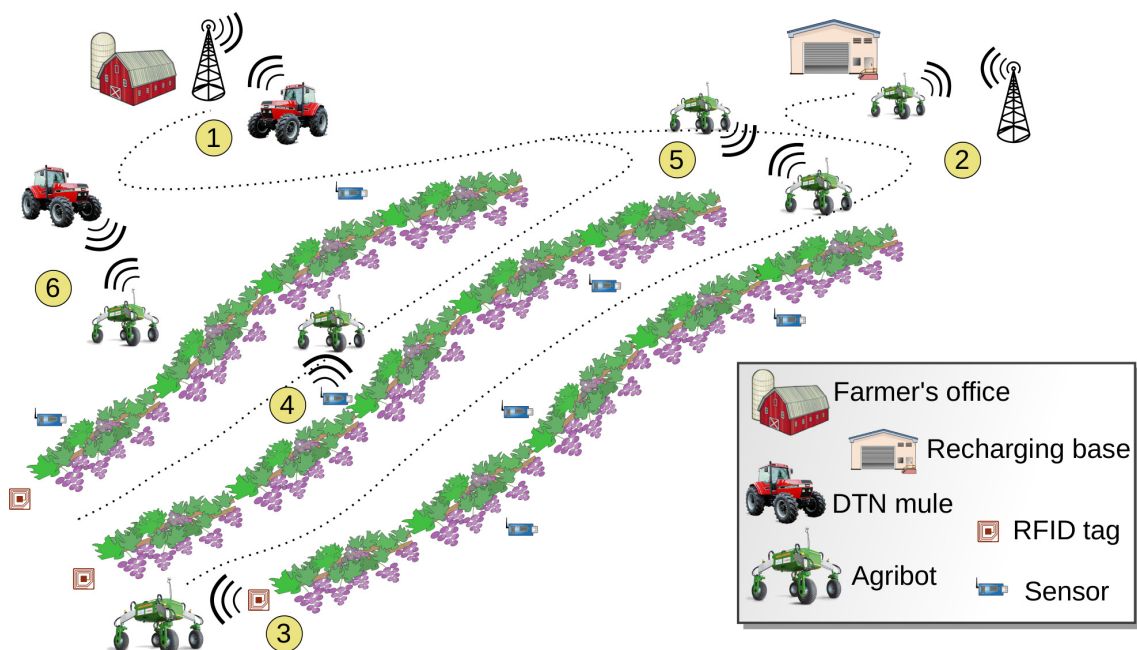


FIGURE 5.2 : Illustration d'un scénario de viticulture

### 5.2.2 La planification par réseaux hiérarchisés de tâches

Chaque objet de cette étude de cas est représenté par un avatar, qui a des tâches haut niveau à réaliser. Les approches de planification sont une solution pour définir les actions qui mènent à des objectifs donnés. La technique de planification que nous utilisons dans l'implémentation de la contribution de cette thèse est appelée HTN (*Hierarchical Task Network*) [Georgievski and Aiello, 2015]. Son principal avantage est la capacité de représenter l'état de l'environnement à un niveau d'abstraction supérieur à celui de la planification classique, en établissant une séquence d'actions sur la base des tâches à accomplir plutôt que sur les objectifs à atteindre. De plus, les approches HTN peuvent résoudre les problèmes de planification plus rapidement que de nombreux planificateurs classiques, comme GRAPHPLAN [Blum and Furst, 1997].

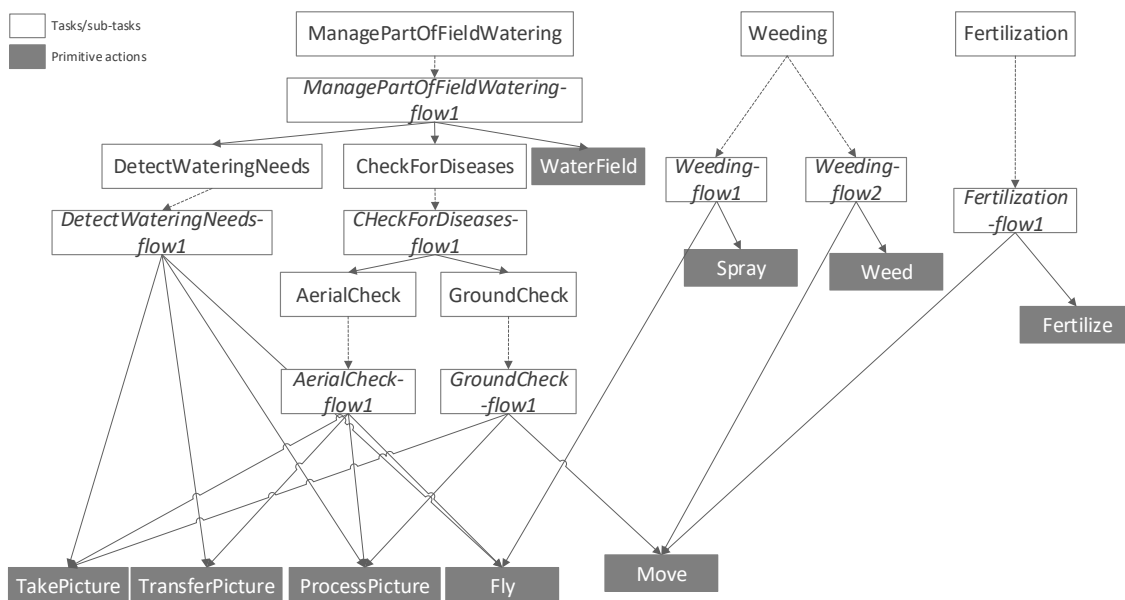


FIGURE 5.3 : La librairie de plan utilisée dans l'étude de cas.

La hiérarchie de fonctionnalités illustrée dans la Figure 5.3 correspond à un scénario d'arrosage automatique. La fonctionnalité `ManagePartOfFieldWatering` peut être exposée par l'avatar d'un asperseur ou un tag RFID d'un rang de vignes, par exemple. Elle est décomposée en `DetectWateringNeeds` pour détecter si une partie d'un champ de vignes a besoin d'être arrosée, `CheckForDiseases` pour vérifier si une partie est atteinte par des maladies, et `WaterField` pour arroser. `DetectWateringNeeds` consiste à prendre une photo aérienne (`TakePicture`), la transférer (`TransferPicture`) à un avatar ou au cloud pour appliquer des algorithmes de traitement d'image (`ProcessPicture`) afin d'extraire des informations. Pour vérifier si une partie du champ de vignes est atteinte

d'une maladie, une analyse du sol (GroundCheck) ou aérienne (AerialCheck) est nécessaire.

HTN est l'une des techniques les plus utilisées dans les applications du monde réel [Ghaleb et al., 2004] pour des raisons telles que l'expressivité, la représentation de domaine et la façon de résoudre les problèmes. Cette technique représente les actions et l'état de l'environnement de la même manière que la planification classique : l'état de l'environnement étant un ensemble de propositions logiques, et le modèle d'action étant un triplet  $\langle \text{action}, \text{préconditions}, \text{effets} \rangle$  :

- *action* est le nom de l'action (avec ou sans paramètres),
- *préconditions* est la liste des conditions préalables qui doivent être satisfaites pour effectuer *action*,
- *effets* est la liste des propositions logiques qui sont ajoutées ou supprimées des croyances après l'exécution de l'action, pour décrire le nouvel état de l'environnement.

Contrairement à la planification classique, HTN consiste à accomplir des tâches plutôt qu'à atteindre des objectifs. Un problème de planification HTN est représenté par le quadruple  $\langle s_i, w_i, A, M \rangle$ , tel que :

- $s_i$  est l'état initial de l'environnement,
- $w_i$  est un réseau de tâches hiérarchiques (comme c'est illustré dans la Figure 5.3),
- $A$  est un ensemble d'actions de planification classique (nom, préconditions, effets),
- $M$  est un ensemble de méthodes HTN (nom, préconditions, contraintes). Cet ensemble est utilisé pour décomposer récursivement une tâche en sous-tâches tout en respectant si des préconditions sont vérifiées. La décomposition se termine lorsqu'on arrive à des actions primitives.

**La planification HTN pour les agents BDI** Dans le cycle de raisonnement d'un agent BDI, l'agent répond à des événements en sélectionnant les plans les plus appropriés pour traiter un événement donné. Ce traitement se fait à base de la librairie de plans qui contient son savoir-faire. Une librairie de plans BDI est généralement un ensemble de règles sous la forme  $e : \psi \leftarrow P$  telle que  $P$  est le corps d'un plan (la séquence des actions à exécuter), et  $e$  est l'évènement à traiter lorsque le contexte  $\psi$  est favorable. Si l'exécution du plan échoue à un moment donné, l'agent cherche un plan alternatif pour l'ajouter à ses intentions. Ce processus se répète jusqu'à ce que le plan choisi pour traiter l'évènement sélectionné réussisse, ou qu'il n'y ait plus de plans alternatifs. Il y a plusieurs similarités entre les planificateurs HTN et les langages de programmation BDI

dans la représentation et la manipulation de la librairie de plans.

- Premièrement, les planificateurs HTN et les langages BDI supposent une représentation explicite des croyances de l'agent et un ensemble de tâches ou d'actions primitives que l'agent peut exécuter sur l'environnement.
- Deuxièmement, les connaissances procédurales sur le domaine sont disponibles dans les deux sous forme de méthodes de réduction (pour HTN) et de règles de plan (pour BDI).
- Troisièmement, les deux techniques trouvent des solutions en réduisant les tâches de haut-niveau en tâches de niveau inférieur en faisant appel à un ensemble donné d'instructions de réduction.

D'autres détails sur les avantages de l'utilisation de la technique de planification HTN dans des langages de programmation BDI sont fournis dans les spécifications du langage CANPLAN [Sardina and Padgham, 2011] qui propose une extension HTN d'un langage BDI appelé CAN [Winikoff et al., 2002] (Conceptual Agent Notation, qui est à son tour une extension du langage AgentSpeak [Rao, 1996]).

**La planification HTN dans le Web sémantique** L'utilisation des ontologies dans le Web des Objets est nécessaire pour permettre la représentation des connaissances sous forme interopérable par des objets physiques. Cependant, l'intégration des ontologies dans des systèmes d'agents et des formalismes de planification est encore un chemin de recherche à ses étapes initiales. Plusieurs travaux proposent une représentation sémantique des domaines de planification en ontologies. Dans ce cadre, Freitas et al. [2014] ont proposé un modèle sémantique pour représenter des librairies de plans HTN à l'aide d'ontologies implémentées en langage OWL. Ce modèle est accompagné d'algorithmes de conversion de ces librairies de plans HTN (plus spécifiquement, celles utilisées dans le planificateur SHOP [Nau et al., 1999]) et des langages de représentation de plans BDI (à savoir, AgentSpeak) en ontologies, et vice-versa. Cette conversion profite des similarités entre les plans HTN et ceux utilisés dans la programmation des agents BDI.

Nous nous inspirons du travail de Freitas et al. [2014] pour implémenter les répertoires de fonctionnalités des avatars, pour son originalité dans l'intégration des ontologies avec le formalisme HTN et avec les langages de programmation orientée agent. Ainsi, les classes et les propriétés que nous utilisons dans la formalisation de notre ontologie de plans HTN sont illustrées dans la Figure 5.4, et décrites telles que les classes sont les nœuds, et les propriétés sont des arcs :

**Action** Une tâche non décomposable qui peut être appliquée directement sur l'environ-



Par exemple, voici comment nous représentons l'instance *domain-viticulture* de la classe *DomainDefinition* :

```
### http://liris.cnrs.fr/asawoo/vocab#domain-viticulture
asawoo:domain-viticulture
  rdf:type owl:NamedIndividual, asawoo:DomainDefinition ;
  asawoo:has-action asawoo:action-processpicture ,
    asawoo:action-takepicture ,
    asawoo:action-transferpicture ,
  asawoo:action-waterfield ;
  asawoo:has-method asawoo:method-aerialcheck ,
  asawoo:method-checkfordiseases ,
  asawoo:method-detectwateringneeds ,
  asawoo:method-fertilization ,
  asawoo:method-groundcheck ,
  asawoo:method-managepartoffieldwatering ,
  asawoo:method-weeding .
```

La méthode *DetectWateringNeeds* est représentée comme suit :

```
### http://liris.cnrs.fr/asawoo/vocab#method-detectwateringneeds
asawoo:method-detectwateringneeds
  rdf:type owl:NamedIndividual, asawoo:Method ;
  asawoo:has-flow asawoo:detectwateringneeds-flow1 ;
  asawoo:utility "85.47974820969246"^^xsd:double .
```

Ici, nous illustrons les propriétés qui lient les classes. Par exemple, la propriété *has-flow* lie des flux (*detectwateringneeds-flow1*) à des méthodes (*detectwateringneeds*) pour indiquer quels flux décomposent une méthode. Cette décomposition se fait à travers la propriété *has-flow*, dans un premier temps, pour savoir quel flux suivre dans la décomposition, en fonction de l'état de l'environnement. Dans un deuxième temps, la propriété *has-postcondition* (comme illustré dans la Figure 5.4) permet de savoir quelles sont les actions et/ou sous-tâches résultantes de la décomposition suivant ce flux.

Pour savoir, à l'aide d'une seule propriété, quelles tâches/actions composent une tâche donnée, il existe dans le langage OWL un type de propriété qui ne lie pas explicitement deux classes dans une ontologie (comme c'est le cas dans toutes les propriétés de la Figure 5.4), mais qui fait cette liaison d'une manière implicite en composant une propriété avec d'autres. Ainsi, nous avons créé une propriété *planStep* définie comme suit :

```
### http://liris.cnrs.fr/asawoo/vocab#planStep
asawoo:planStep rdf:type owl:ObjectProperty, owl:TransitiveProperty ;
                owl:propertyChainAxiom ( asawoo:has-flow
                                           asawoo:has-postcondition ) .
```

La requête SPARQL suivante permet de constituer l'ensemble des actions qui composent une tâche donnée en paramètre (*DetectWateringNeeds*) :

```
SELECT ?action WHERE { {
  asawoo:method-detectwateringneeds asawoo:planStep+ ?action.
  ?action rdf:type asawoo:Action.
} }
```

### 5.3 La construction de l'identité d'un agent IDEAS

L'intersubjectivité contribue à la continuité d'un groupe en favorisant sa formation et sa cohérence (discuté dans la Section 3.6). Par conséquent, pour définir son identité, un agent pourra utiliser un tel concept pour identifier ce qu'il partage avec les autres. Plus un groupe partage des significations intersubjectives, plus grande est la cohésion et la communication du groupe. Ainsi, pour savoir si l'autre partage ces significations, un agent doit prendre en compte l'intention derrière les actions et les sollicitations des autres agents dans son interaction avec eux. À cette fin, nous traduisons le concept d'intersubjectivité par un processus de reconnaissance d'intention.

**La reconnaissance d'intention dans les systèmes multi-agents** Il y a plusieurs approches de modélisation du problème de reconnaissance d'intention. On distingue entre trois formalismes qui classifient ces approches :

- la logique [Allen et al., 1991], qui définissent les relations entre les actions et les buts à travers des spécifications formelles, comme la grammaire algébrique ;
- les probabilités [Charniak and Goldman, 1993], qui se basent sur la formation d'hypothèses probabilistes sur les intentions des agents observés, en fonction la probabilité d'évènement de leurs actions. Ces approches utilisent par exemple le raisonnement bayésien ou les automates de Markov à états cachés ;
- les statistiques/cas [Blaylock and Allen, 2003, Kerkez and Cox, 2003], qui font appel à l'expérience et les situations déjà rencontrées par l'agent observateur pour reconnaître l'intention de l'agent observé.

Une grande proportion des techniques s'inscrit dans les formalismes basés sur la logique [Sadri, 2010] et les probabilités [Blaylock and Allen, 2006]. La reconnaissance d'intention multi-agent s'intéresse à l'explication de l'activité d'un ensemble d'agents en identifiant des structures d'équipes et leurs objectifs [Banerjee et al., 2010]. Quelle que soit la forme de raisonnement employée, qu'elle soit abductive, déductive, probabiliste ou mixte, la reconnaissance de l'intention exige une certaine connaissance de la façon dont les actions atteignent des objectifs. Il est montré que ce problème est NP-complet [Banerjee et al., 2015], et rencontre plusieurs difficultés qui commencent par celles héritées dans la reconnaissance d'intention *mono-agent*, à savoir :

- le coût de la maintenance manuelle des bibliothèques de plans [Zhuo and Li, 2011],
- l'absence d'un moyen de savoir si le plan de l'agent ou l'équipe observée a échoué ou a été abandonné,
- l'observabilité partielle des actions des agents.

En l'absence d'une bibliothèque de plans, les agents peuvent utiliser des modèles d'action [Zhuo et al., 2012] pour reconnaître les plans pouvant être exécutés par une équipe observée. Ces techniques nécessitent une trace des activités d'une équipe d'agents, un état initial et un but. L'objectif est de générer tous les plans possibles pour passer de l'état initial à l'état du but, et conserver les plans pouvant être expliqués par la trace de l'équipe observée. Il existe d'autres classes d'approches qui ne reposent pas sur des bibliothèques de plans ou des modèles d'action. Les approches qui se basent sur les rôles peuvent reconnaître les intentions en identifiant les rôles des agents et les structures des équipes. Un rôle est une représentation de l'identification des services et fonctions d'un agent au sein d'une équipe, il définit les relations, en terme d'actions, tâches, intentions, et plans, entre des agents appartenant à une équipe [Weyns et al., 2005]. Dans ce cadre, Soon et al. [2003] proposent une approximation d'une intention d'équipe en considérant les activités qui ont lieu entre les rôles des agents dans une structure hiérarchique (appelée *RoleGraph*). Cette dernière contient des informations importantes sur l'équipe, en faisant abstraction des détails de bas niveau. Les agents peuvent ensuite faire correspondre ces activités entre des rôles avec des *RoleGraphs* qu'ils connaissent déjà. D'autres approches se basent sur la reconnaissance d'intention à partir de l'identification des coalitions (ou collusions). Kerr and Cohen [2011] proposent une technique permettant de détecter la présence des coalitions (ou collusions) et d'identifier leurs membres. Cette technique n'exige aucune connaissance sur les plans exécutés par les coalitions, elle se base sur les propriétés fondamentales des actions observables elles-mêmes. Généralement, un agent fait partie d'une coalition parce qu'il s'attend à un avantage, qui semble être fondamental à l'existence de coalitions, et il est utilisé comme base de cette technique. Plus précisément,



les *flux des intérêts* entre les agents sont observés, c'est-à-dire les actions d'un agent qui sont bénéfiques pour un autre.

### 5.3.1 Le mécanisme de reconnaissance d'intention dans IDEAS

La reconnaissance de plan a été modélisé de différentes manières par de nombreuses approches prenant de nombreux types d'entrée [Sukthankar et al., 2014]. En ce qui concerne le contexte multi-SMA, en particulier avec l'utilisation du Web sémantique dans le Web des objets [Guinard et al., 2011], les agents intègrent des ontologies de domaine comme une librairie de plan dans leur processus de prise de décision. Pour implémenter notre mécanisme de reconnaissance d'intention multi-agent, nous trouvons que l'algorithme proposé par Zhuo and Li [2011], appelé MARS (*MultiAgent Recognition System*) est le plus approprié pour le contexte de cette thèse. Ce choix est motivé par la prise en compte de l'observabilité partielle de l'environnement, et par l'incomplétude des librairies de plans.

L'objectif de cet algorithme est de reconnaître des plans multi-agents en se basant sur des traces partiellement observables d'un ensemble d'agents, et des librairies de plans (qui peuvent être partielles). Nous utilisons cet algorithme comme une base de notre fonction de reconnaissance de plan ( $\mathcal{R}$ ). MARS prend en entrée la librairie de plans  $\mathcal{P}_{lib}$  et une trace d'équipe  $\mathcal{O}_{env}$ , et produit en sortie un ensemble d'occurrences qui couvre  $\mathcal{O}_{env}$ . Cet algorithme se décompose en trois étapes :

1. la création d'un ensemble d'occurrences candidates qui correspondent chaque sous-ensemble d'observations  $\mathcal{O}_{env}$  à un plan dans  $\mathcal{P}_{lib}$  ;
2. la génération des contraintes "légères" en associant des poids à chaque candidat, qui dépend du nombre d'actions dans un plan, de l'utilité du plan et de l'observabilité des actions ;
3. la génération des contraintes "dures" pour forcer le choix du meilleur plan de  $\mathcal{P}_{lib}$  pour chaque candidat de l'étape 1.

Techniquement, le problème de reconnaissance de plan est modélisé comme une instance d'un problème *Partial Weighted Max-SAT*, dont la solution optimale est la meilleure mise en correspondance entre les actions observées avec les plans de la librairie de plans. Ensuite, le problème peut être résolu en utilisant un solveur Max-SAT.

**Partial Weighted MaxSat Problem** Le problème *Maximum SATisfiability (Max-SAT)* est l'un des homologues d'optimisation du problème de satisfiabilité booléenne (*SAT*).

Dans *Max-SAT*, étant donné une formule booléenne dans la forme normale conjonctive (*CNF*, *Conjunctive Normal Form*), on veut déterminer le nombre maximum de clauses qui peuvent être satisfaites par une affectation complète. Une clause est une disjonction de littéraux, et un littéral est simplement une variable ou sa négation.

Il y a deux variantes importantes du problème *Max-SAT* :

- *weighted Max-SAT* : c'est un problème *Max-SAT* dans lequel un poids positif est associé à chaque clause. La solution consiste à maximiser la somme des poids des clauses satisfaites par n'importe quelle affectation.
- *partial Max-SAT* : c'est un problème *Max-SAT* dans lequel certaines clauses ne peuvent être falsifiées par aucune solution.

La combinaison des deux variantes est appelée le problème *weighted partial Max-SAT*. Beaucoup de solveurs existent dans la communauté SAT (Les solveurs les plus à jour sont en compétition sur <http://www.maxsat.udl.cat>), comme Sat4J [Le Berre and Parrain, 2010] et MaxSatz [Li et al., 2007].

L'attribution de poids appropriés pour les contraintes légères, dans l'étape 2, joue un rôle important dans l'efficacité et la précision de l'algorithme de reconnaissance d'intention. Nous adaptons l'équation qui calcule le poids dans l'algorithme MARS comme suit :

$$\omega_i = ctx(\mathcal{T}_i, \mathcal{O}_{env}) \times v(\mathcal{T}_i) \times \lambda(c_i) \quad (5.1)$$

- $ctx(\mathcal{T}_i, \mathcal{O}_{env})$  correspond au gestionnaire de contexte décrit dans le modèle BDI de base. Nous utilisons l'implémentation déjà existante de Terdjimi et al. [2017] dans l'architecture de l'avatar, qui entre autres renvoie un score indiquant la précision d'une tâche  $\mathcal{T}_i$  étant donné le contexte actuel, en supposant que la confiance sur l'évènement d'une action dépend de son applicabilité et de sa précision dans l'état actuel de l'environnement. Par exemple, nous sommes plus confiants qu'une tâche consistant à piloter un drone ne se produit pas lorsque la vitesse du vent est élevée.
- $v(\mathcal{T}_i)$  est l'utilité d'une tâche, étant donné qu'un agent est plus motivé de participer à des tâches avec des utilités élevées.
- $\lambda(c_i)$  est le taux d'observation, c'est-à-dire le degré de confiance que le candidat  $c_i$  pour lequel on calcule le poids a eu lieu.

Zhuo and Li [2011] supposent que le nombre d'actions observées affecte la confiance en l'occurrence d'un candidat  $c_i$ . Ainsi, le taux d'observation  $\lambda(c_i)$  est défini par :

$$\lambda(c_i) = \frac{2|\mathcal{T}_i| - |null(\mathcal{T}_i)| - |null(\mathcal{O}_{env})|}{2|\mathcal{T}_i|} \times N(c_i)$$

- $|\mathcal{T}_i|$  correspond au nombre total d'actions dans une tâche.
- $|\text{null}(\mathcal{T}_i)|$  est le nombre de tâches ou d'actions indéfinies dans la librairie de plans (par exemple, un agent peut avoir une librairie qui manque certaines étapes dans la décomposition d'une tâche).
- $|\text{null}(\mathcal{O}_{env})|$  est le nombre d'observations manquantes d'une activité d'équipe (par exemple, un agent  $a_1$  peut savoir que  $a_2$  a effectué une action à un moment donné, mais ne connaît pas nécessairement le nom de l'action).
- $N(c_i)$  est le nombre d'agents (i) appartenant au même système de l'agent observateur ou (ii) ceux avec lesquels il a déjà coopéré (en supposant qu'un agent a plus de confiance dans le résultat de reconnaissance des activités de ses accointances).

Pour simplifier, nous présentons dans l'étude de cas un agribot (*agribot1*) pour illustrer la construction de l'identité, et un drone (*drone1*) pour illustrer le processus d'engagement.

### 5.3.2 Le processus de construction d'identité

Nous souhaitons doter chaque agent de la capacité de construire son identité personnelle et sociale afin d'avoir une représentation de ce qu'il doit préserver lorsqu'il interagit avec les autres. Concernant l'identité personnelle, elle peut être directement dérivée des :

- attributs de l'objet connecté correspondant (identifiant, fabricant, date de sortie, ...) à l'avatar,
- ses capacités (capteurs, actionneurs, ...),
- ses objectifs (individuels ou spécifiés par le développeur),
- d'autres composants comme l'historique des interactions (voir Figure 3.8 pour plus de détails).

En ce qui concerne l'identité sociale, chaque agent observe constamment ses voisins (qu'ils soient membres du même SMA ou pas). Nous utilisons un système de reconnaissance d'intention pour que chaque agent puisse définir quelles tâches collectives il poursuit et dans quelles coalitions il est impliqué. Par exemple, *agribot1* construit constamment une base d'observations  $\mathcal{O}_{env}$  qui inclut l'historique des actions exécutées par chaque agent dans son voisinage. Lorsque l'occasion se présente, l'agribot actualise son identité sociale en reconnaissant les coalitions formées et leurs tâches respectives. Le système de reconnaissance de plan fonctionne comme suit : pour chaque ensemble  $o$  de tous les sous-ensembles d'observations  $\mathcal{O}_{env}$ , pour chaque action dans  $o$ , la requête suivante récupère toutes les tâches de haut niveau qui peuvent expliquer le fait que certaines actions données (par exemple, *move*) sont exécutées :

```
SELECT ?method WHERE {
  asawoo:move asawoo:partOfMethod ?method.
}
```

avec *partOfMethod* une propriété définie comme l'inverse de la propriété *planStep*, sachant que "la reconnaissance de plan est la planification en sens inverse" [Ramirez and Geffner, 2009].

```
### http://liris.cnrs.fr/asawoo/vocab#partOfMethod
asawoo:partOfMethod rdf:type owl:ObjectProperty ;
  owl:propertyChainAxiom (
    [ owl:inverseOf asawoo:has-postcondition ]
    [ owl:inverseOf asawoo:has-flow ]
  ) .
```

Ensuite, nous associons des poids à chaque tâche retournée par la requête. Par exemple, la requête SPARQL<sup>3</sup> ci-dessus permet de reconnaître la tâche Weeding pour l'actions observées {agribot2:Move}. Nous associons à cette tâche reconnue le poids suivant :

$$\omega_i = \left( \frac{2|\text{Weeding}| - |\{\text{null}_{\text{Weeding}}\}| - |\{\text{null}_o\}|}{2|\text{Weeding}|} \times N(\{\text{drone2}, \text{agribo3}\}) \right) \times v(\text{Weeding}) \times \text{Ctx}(\text{Weeding}, \mathcal{O}_{env})$$

La tâche Weeding contient 4 actions, sans aucune étape manquante dans la librairie de plan (voir Figure 5.3). Supposons qu'il n'y a pas d'observations manquantes dans le sous-ensemble courant  $o$ , et que *agribot1* (l'observateur) ait collaboré avec *agribot2*, donc  $N = 1$ .

$$\omega_i = \left( \frac{8 - 0 - 0}{8} + 1 \right) \times v(\text{Weeding}) \times \text{Ctx}(\text{Weeding}, \mathcal{O}_{env})$$

$\text{ctx}(\text{Weeding}, \mathcal{O}_{env})$  est calculé par le gestionnaire de contexte de l'avatar, et l'utilité  $v(\text{Weeding})$  est donnée dans la librairie de plan comme une des propriétés de la tâche.

Dans cet exemple, nous avons considéré une seule tâche reconnue pour illustrer le calcul de poids, les étapes générales restantes sont bien expliquées dans l'algorithme MARS

3. Les ontologies définissant la librairie de plans sont accessibles sur : <https://gitlab.com/asawoo/semantic-repos/tree/master/ontologies>

[Zhuo and Li, 2011]. En utilisant cette technique de reconnaissance de plan, *agribot1* identifie l'une des tâches qu'il poursuit (Weeding) et l'une des coalitions auxquelles il est impliqué (*agribot2*).

Nous avons doté l'agent de la capacité de construire son identité personnelle et sociale. À partir de maintenant, l'agent sait ce qu'il doit préserver avant de s'engager à poursuivre d'autres tâches. En d'autres termes, il doit prendre en considération qu'il est engagé auprès de *agribot2* de réaliser la tâche Weeding avant d'accepter de coopérer avec un autre agent.

## 5.4 Le processus d'engagement des agents IDEAS

Le drone (*drone1*) a deux missions : une mission de pulvérisation et une mission de supervision aérienne. Dans ce contexte, il souhaite accomplir la tâche `ManageFieldWatering` qui est décomposable comme illustrée dans la Figure 5.3. Supposons que *drone1* n'est pas capable d'accomplir seul cette tâche de la manière la plus optimale, il formera donc un coalition pour traiter `ManageFieldWatering` de manière collective. C'est là où *Contract Net Protocol* nous servira comme protocole de formation de coalition.

**Contract Net Protocol** *Contract Net Protocol* (CNP) est un protocole de partage de tâches défini par Smith [1980] pour l'allocation décentralisée de tâches sur la base d'appels d'offres par un *manager* et des soumissions de propositions par des *contractors*. De nombreuses extensions et variantes de ce protocole ont été proposées. Dans cette thèse, nous utilisons celle proposée par Aknine et al. [2004]. Cette extension avait pour but de permettre à un agent de gérer plusieurs processus de négociation en parallèle et de réduire les situations de désengagement des *contractors*. Les primitives de communication utilisées par les *managers* et les *contractors* dans cette extension sont :

- *Announce* : un agent propose l'exécution d'une tâche.
- *PreBid* : un agent fait une offre temporaire pour l'exécution d'une tâche.
- *PreAccept* : un agent est temporairement accepté pour l'exécution d'une tâche.
- *PreReject* : un agent est temporairement rejeté pour l'exécution d'une tâche.
- *DefinitiveBid* : un agent fait une offre définitive pour l'exécution d'une tâche.
- *DefinitiveAccept* : un agent est définitivement accepté pour l'exécution d'une tâche.
- *DefinitiveReject* : un agent est définitivement rejeté pour l'exécution d'une tâche.

Nous trouvons l'extension de Aknine et al. [2004] très appropriée à notre problème, car un agent doit raisonner sur les coalitions potentielles auxquelles il participerait et les

actions qu'il réaliserait avant de s'engager à coopérer.

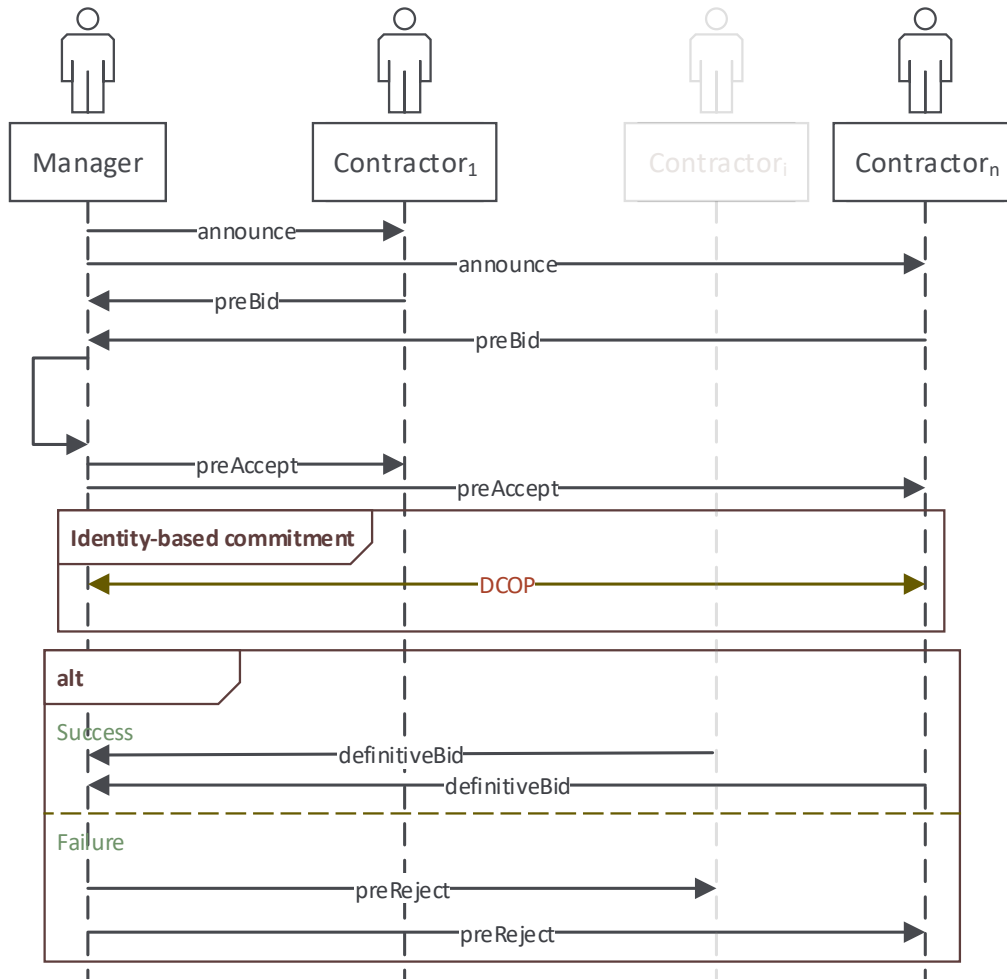


FIGURE 5.5 : Adaptation du protocole *Extended CNP* proposé par Aknine et al. [2004]

Une fois que `ManageFieldWatering` est décomposée, *drone1* (le *manager*) envoie (comme illustré dans la figure 5.5) des messages *Announce* à tous les autres agents de son voisinage demandant l'exécution de toutes les sous-tâches et les actions dérivées de la décomposition. Lorsque *agribot1* reçoit le message *Announce*, il répond à cette requête en envoyant un message *PreBid* pour chacune des sous-tâches ou actions qu'il peut effectuer. Dans ce scénario, *agribot1* ne peut fournir que `GroundCheck`.

Après un *timeout* donné, le *manager* évalue les propositions soumises par les *contractors* et sélectionne les meilleures  $k + 1$  *PreBids* pour chaque sous-tâche ou action demandée. Une coalition est considérée comme  $k$ -robuste [Okimoto et al., 2015] (pour un entier positif  $k$ ) si la suppression de  $k$  agents conduit à une équipe restante qui peut toujours accomplir les tâches donnée. Ensuite, *drone1* enverra des messages *PreAccept* aux meilleurs  $k + 1$  *contractors*.

La coalition n'est pas encore formée puisque les agents ne connaissent ni leurs coéquipiers ni leurs sous-tâches (ou actions) fournies. Ainsi, même si le meilleur enchérisseur pour chaque action rejoint la coalition, il peut être impliqué dans l'accomplissement de quelques tâches interférant négativement dans la réalisation de ses propres tâches. Après l'envoi de *PreAccepts*, le *manager* établira un réseau de contraintes tel que chaque contrainte implique des actions dont les contractors déploiera un *factor graph*<sup>4</sup> afin qu'ils puissent résoudre les contraintes liées à la préservation de leurs identités et optimiser la qualité globale de la solution (équation 4.1). La résolution du problème DCOP<sup>5</sup> permettra à *agribot1* de s'engager à rejoindre la coalition menée par *drone1* tout en préservant l'objectif Weeding (identifié à partir de l'étape de construction de l'identité).

Les algorithmes résolvant les problèmes DCOP peuvent être catégorisés en des algorithmes *exacts* (comme ADOPT [Modi et al., 2005], DPOP [Petcu and Faltings, 2005] et leurs extensions) qui garantissent la solution optimale mais exigent un surcoût de coordination exponentiel, et des algorithmes *non-exacts* (comme Max-Sum [Farinelli et al., 2008] et ses extensions) qui sont basés sur l'échange local d'informations et peuvent fournir une qualité de solution acceptable tout en étant efficaces en termes de coût de calcul et de communication.

**Max-Sum algorithm** Max-Sum fonctionne sur un *factor graph*, qui est un graphe biparti non orienté. Ses sommets représentent des variables et des contraintes (appelés *factors*), et ses arêtes relient les facteurs à leurs variables. Les messages vont passer des variables aux facteurs, et des facteurs aux variables. Chaque variable  $X_i$  calcule le message  $Q_{i \rightarrow j}(X_i)$  à envoyer au *factor*  $f_j$  en ajoutant les valeurs reçues de ses *factors* (sauf le *factor*  $f_j$ ).

$$Q_{i \rightarrow j}(X_i) = \sum_{k \in M_i - j} R_{k \rightarrow i}(X_i)$$

À contrario, un *factor*  $f_j$  calcule le message  $R_{j \rightarrow i}(X_i)$  à envoyer à la variable  $X_i$  en ajoutant sa propre valeur aux valeurs reçues de ses variables (sauf la variable  $X_i$ ).

$$R_{j \rightarrow i}(X_i) = \max_{X_j - i} [f_j(X_j) + \sum_{k \in N_j - i} Q_{k \rightarrow j}(X_k)]$$

L'algorithme se termine lorsqu'un les nouveaux messages reçus par les nœuds sont identique aux derniers messages reçus dans l'itération précédente, ou lorsqu'une limite de

4. Le déploiement du graphe factoriel sort du cadre de cet article.

5. Le code source de l'ensemble du mécanisme de coopération est disponible sur : <https://gitlab.com/asawoo/applications>

temps ou un nombre donné d'itérations est atteint. Max-Sum est complet pour les *factor graphs* en forme d'arbre, sous-optimal pour les graphes en boucle voire ne pas converger du tout. Mais il a été démontré que l'algorithme calcule des solutions de meilleure qualité avec un calcul acceptable par rapport à plusieurs algorithmes complets évalués [Farinelli et al., 2008].

## 5.5 Synthèse

Dans ce chapitre, nous avons présenté le contexte applicatif de cette thèse, qui est le projet ANR ASAWoO. Ce contexte impose des défis à relever par les différentes contributions du projet, qui sont implémentées sous forme de modules faisant partie de l'architecture logicielle de l'avatar. Ensuite, nous avons décrit un scénario de viticulture pour illustrer la contribution de cette thèse. Pour mettre en œuvre notre contribution, nous avons dû procéder à des choix techniques. Ainsi, le choix d'adopter HTN comme représentation de librairie de plans est principalement motivé par la présence de travaux visant l'intégration de cette technique avec le modèle BDI. Le choix de représenter sémantiquement les données que l'agent utilise est avant tout pour s'aligner aux spécifications techniques fixées par les membres du projet ASAWoO. En effet, dans le contexte du WoT, l'utilisation des ontologies permet une interopérabilité dans la représentation et un meilleur raisonnement sur les connaissances des avatars. Notons que notre modèle IDEAS pourrait fonctionner avec des architectures autres que BDI (comme MWAC [Jamont et al., 2010]), d'autres choix techniques auraient pu être adoptés en fonction du contexte applicatif.

Nous avons ensuite illustré les deux processus qui permettront à nos agents de pouvoir prendre en compte leurs identités lors de l'engagement à faire partie d'une coalition. Pour construire son identité, un agent continue à observer son environnement pour tirer des conclusions sur les actions et les interactions des agents autour de lui. À partir de chaque observation (ou ensemble d'observations) l'agent se sert d'un algorithme de reconnaissance d'intention pour identifier les objectifs haut-niveau de l'agent (ou de l'ensemble d'agents) observé. Ainsi, il pourrait savoir à travers son exécution d'actions, quels agents font partie de la même coalition que lui, et quels agents pourraient être des potentiels futurs collaborateurs. Dans le choix de l'algorithme de reconnaissance d'intention, nous avons pris en compte les algorithmes capables de reconnaître l'intention d'une équipe d'agents observés, de gérer l'incomplétude de la librairie de plans et l'observabilité partielle des actions des agents. Le processus d'engagement est basé sur une



---

extension du protocole CNP, à laquelle nous avons ajouté une phase de négociation entre les enchérisseurs choisis par le *manager*. Cette phase est modélisée comme un problème d'optimisation de contraintes que nous résolvons en utilisant l'algorithme Max-Sum. Ce processus permet aux agents (aux enchérisseurs) de trouver un compromis entre assurer une bonne qualité de la solution produite collectivement, en réduisant le nombre d'interférences négatives entre leurs actions, ou en maximisant le nombre d'interférences positives de ces actions, avec la possibilité de favoriser la réalisation d'une tâche par des agents partageant une même identité sociale. Ces comportements se manifestent, si dans la modélisation du processus d'engagement en DCOP, nous arrivons à les capturer dans la fonction à optimiser. Cette manifestation dépend aussi du choix des techniques utilisées pour résoudre le problème. Enfin, le chapitre suivant va nous permettre de valider la pertinence de notre modèle.



# Évaluation

---

<b>6.1 Objectifs</b> . . . . .	<b>89</b>
<b>6.2 Description du cadre expérimental</b> . . . . .	<b>90</b>
<b>6.3 Intérêt et pertinence du modèle</b> . . . . .	<b>92</b>
6.3.1 Interférences négatives et positives . . . . .	92
6.3.2 Nombre d'agents partageant une même identité sociale . . . . .	94
<b>6.4 Coût de l'exploitation du modèle</b> . . . . .	<b>96</b>
6.4.1 Comparaison de la qualité de service . . . . .	96
6.4.2 Temps d'exécution . . . . .	97
6.4.3 Évaluation de la précision de l'algorithme de reconnaissance d'intention . . . . .	97
<b>6.5 Synthèse</b> . . . . .	<b>98</b>

---

## 6.1 Objectifs

Ce dernier chapitre présente des résultats expérimentaux visant à montrer la pertinence de notre approche, qui est basée sur l'identité dans la coopération des agents dans des environnements multi-SMAs, et le coût de sa mise en œuvre. Les agents dans de telles conditions sont amenés à établir des coopérations tout en préservant leurs intérêts. Il est donc judicieux de considérer comme premier critère d'évaluation le taux d'interférences négatives que les agents peuvent éviter. Ce critère traduit la capacité des agents à préserver les objectifs qu'ils poursuivaient avant de s'engager dans une nouvelle coopération, lors de la négociation de leurs choix d'actions à fournir.

Le deuxième critère est l'utilité globale de la tâche à réaliser collectivement par des agents. Ce critère montre comment chaque agent est amené à adapter son identité pour privilégier l'utilité globale des actions fournies (par l'ensemble des agents de la coalition) à la préservation de ses objectifs.

Le troisième critère s'intéresse à la tendance des agents qui partagent la même identité sociale à se mettre en coalitions pour agir ensemble. Ce critère devient important lorsqu'un agent ne peut pas raisonner sur les actions des autres pour évaluer leurs interférences sur ses objectifs et évaluer leurs coûts/utilités. Ainsi, il peut se baser sur son historique d'interactions comme critère de décision s'il s'engage à coopérer ou non avec un autre agent.

Enfin, le quatrième critère considère le passage à l'échelle de l'implémentation en terme de temps d'exécution en fonction du nombre de tâches et des actions qui les composent, et en terme d'impact de la taille de ces derniers sur la précision de l'algorithme de reconnaissance d'intention.

## 6.2 Description du cadre expérimental

Pour pouvoir observer ces critères, nous avons mené des expérimentations sur une population de 100 agents appartenant à 10 systèmes distincts, avec 1000 prédicats possibles qui décrivent l'état de l'environnement. Étant donné l'échelle du scénario de viticulture présenté dans le chapitre précédent, nous avons choisi de générer des bibliothèques plus complexes pour mettre en évidence l'apport de notre contribution. Les agents générés correspondent à des objets physiques installés dans les champs de vignes. Les systèmes représentent un ensemble d'objets installés par le même fournisseur, appartenant au même propriétaire, ou déployés par le même concepteur d'un SMA. Les prédicats peuvent repérer les différents états de l'environnement (ligne de vigne arrosée, sol humide, batterie pleine, ...).

Les agents de chaque système partagent une bibliothèque de plans qui décrit comment un certain nombre de tâches peuvent être décomposées en actions. Notre objectif derrière ces expérimentations est d'observer comment les comportements des agents des différents SMA vont être affectés en fonction de leurs identités. Pour ce faire, nous avons choisi quatre critères à observer, que nous avons déjà présentés dans l'équation 4.1 (page 61) qui retourne un score d'identité pour une attribution d'agents à des actions :

$$\delta_{i \rightarrow j} = \begin{bmatrix} \omega_1 \\ \omega_2 \\ \omega_3 \\ \omega_4 \end{bmatrix} \left[ |eff^+(\mathcal{T}(x_j))| \quad -|eff^-(\mathcal{T}(x_j))| \quad \frac{v(x_j)}{v(x_i)} \quad trust(x_j) \right]$$

Dans cette équation les coefficients  $\omega_1, \omega_2, \omega_3, \omega_4$  pondèrent respectivement le nombre d'actions avec des effets positifs (par rapport à l'agent  $i$  qui les évalue), le nombre d'actions avec des effets négatifs, le rapports des utilités des actions fournies par les deux agents  $i$  et  $j$  impliqués dans l'équation, et le degré d'*accointance* qui existe entre les deux agents.

Nous avons généré 4 librairies de plans HTN pour chaque système, qui varient entre 3 et 30 tâches pouvant être décomposées en 10 à 90 actions. Ces choix représentent différentes échelles du problème dans la limite de la complexité que notre implémentation actuelle peut résoudre. Au delà de ces chiffres présentés, l'algorithme ne converge pas vers la solution optimale ou ne donne pas des résultats meilleurs que ceux présentés pour ces chiffres. Les librairies de plans générées sont similaires à celles illustrées dans la Figure 5.3 (page 72), qui contient 7 tâches décomposables en 9 actions. Pour chaque taille de librairie de plans, nous faisons tourner le processus d'engagement sur un échantillon de 4 tâches décomposables en différents nombres d'actions.

Pour chaque tâche à réaliser collectivement par les agents, nous observons :

- le nombre d'interférences négatives entre les actions des agents,
- le nombre d'interférences positives entre les actions des agents,
- le nombre d'agents appartenant à un même système ou ayant déjà coopéré,
- la qualité des choix des actions par des agents.

Cette comparaison se fait entre les actions attribuées aux agents des :

- coalitions formées sans processus d'engagement, c'est-à-dire en se basant sur les meilleures enchères pour chaque action dans la tâche à décomposer (que nous nommons *Original bidders*),
- nouvelles coalitions résultantes du processus d'engagement (nommées *Definitive bidders*).

Nous ajoutons à cette comparaison la solution la plus optimale calculée par un algorithme de pointe étant donné le modèle du problème.

## 6.3 Intérêt et pertinence du modèle

Nous évaluons dans cette section la pertinence de notre contribution à travers la présentation de taux d'interférences négatives qu'un agent peut éviter, et l'influence de l'identité des agents sur le choix de leurs coéquipiers.

### 6.3.1 Interférences négatives et positives

La variation du nombre d'interférences négatives entre les actions des agents de deux coalitions, avant et après le processus d'engagement est un facteur important pour décider de la réussite de notre modèle à doter les agents de la capacité de choisir les actions auxquelles ils s'engagent à réaliser en fonction de leurs identités. La Figure 6.1 montre qu'en moyenne 20% d'interférences négatives sont évitées après le changement des attributions des actions à de nouveaux agents. Ce changement reflète l'affectation d'une valeur non nulle (en occurrence, 1) au coefficient  $\omega_2$  et la mise à 0 du reste des coefficients de l'équation 4.1. Comme le montre le Tableau 6.1, nous notons qu'en moyenne 42% des agents retirent leur *PreBids* pour éviter de réaliser des actions qui interfèrent négativement dans la réalisation de leurs tâches. Nous remarquons aussi qu'en augmentant l'échelle du problème, (30 tâches et 90 actions) on ne gagne pas autant au niveau d'évitement d'interférences négatives. Ceci est dû à la précision de l'algorithme de reconnaissance d'intention qui diminue avec l'augmentation de la taille de l'espace de recherche. Nous discutons la précision de l'algorithme de reconnaissance d'intention dans la Section 6.4.3. En ce qui concerne les interférences positives, elles montrent une plus grande variabilité du fait que plusieurs agents partagent les mêmes tâches. Pour tester la variation d'interférences positives, nous affectons une valeur (=1) non nulle au coefficient  $\omega_1$  et 0 aux autres coefficients.

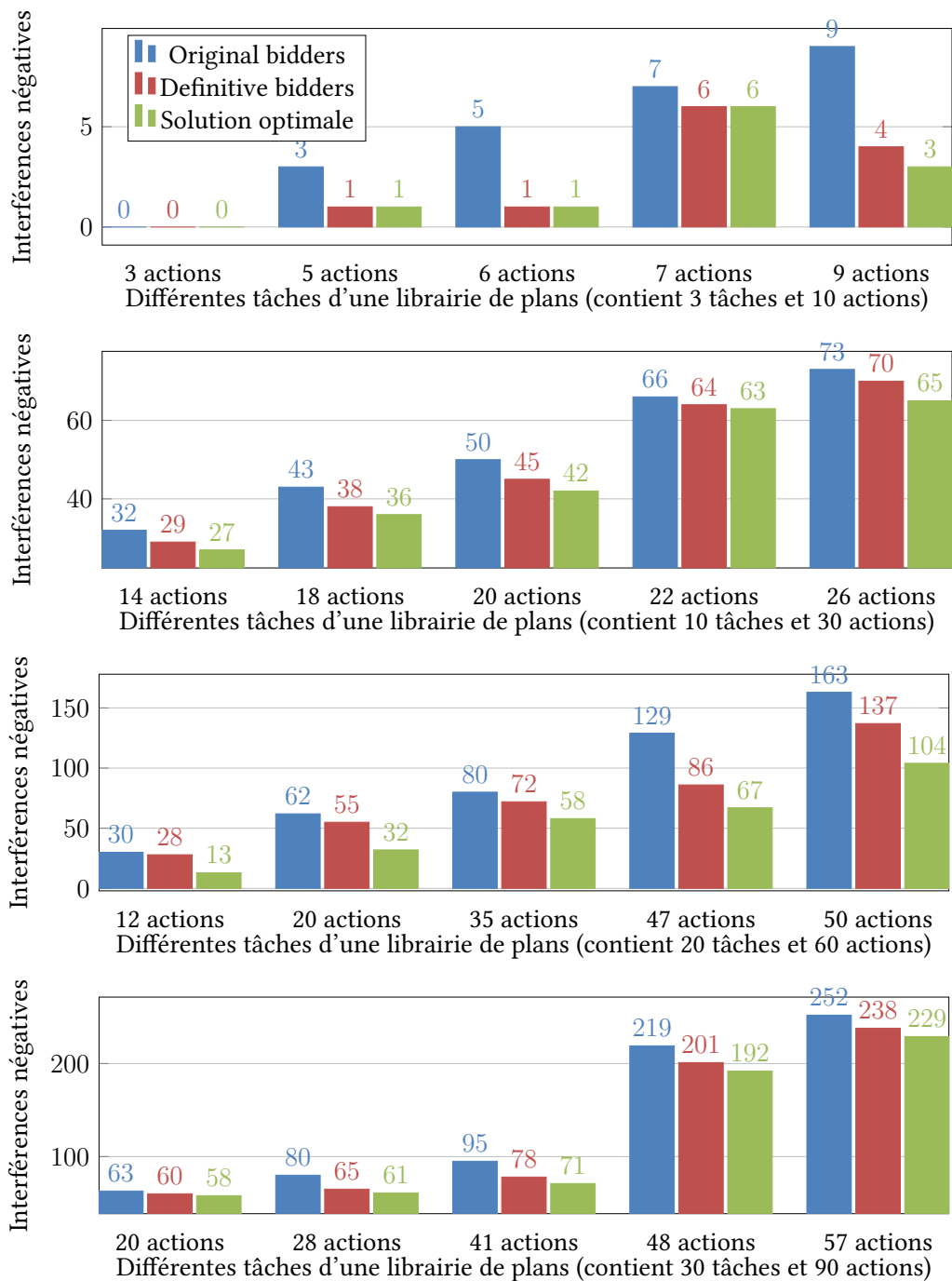


FIGURE 6.1 : Comparaison d'interférences négatives entre un nombre croissant d'actions dans des équipes 3-robustes effectuant des tâches aléatoires.

**Comparaison avec la solution d'un *solver* optimal** Une fois que le processus d'engagement est modélisé en tant que DCOP, il peut être résolu par des algorithmes dédiés. Dans l'implémentation de la contribution de cette thèse, nous nous sommes particulière-

ment intéressés à la classe d’algorithmes *non-exacts* pour résoudre le modèle DCOP. Ces algorithmes ne garantissent pas la solution la plus optimale, mais ils sont moins coûteux en terme de calcul et de communication, ce qui a motivé notre choix. Dans la Figure 6.1, nous avons comparé les résultats obtenus en utilisant notre implémentation de l’algorithme Max-Sum avec la solution la plus optimale. Nous remarquons que les résultats obtenus ne sont pas au même niveau d’optimisation que la meilleure solution possible. Cette différence dépend principalement de la nature suboptimale de l’algorithme. De plus, dans la création du *factor graph* (qui capture les dépendances entre les allocations d’actions à des agents), nous avons fait le choix de créer des *factors* pour chaque paire d’agents participants, ce qui influence la topologie du *factor graph* créé, qui est en effet très dense et surtout cyclique. Dans ces conditions, Max-Sum (et la classe d’algorithmes de propagation de croyances) ne converge pas vers la solution la plus optimale calculée par un solveur. Cependant, notre objectif derrière les expérimentations présentées dans ce chapitre est de fournir une preuve de concept qui montre la pertinence de la prise en compte de l’identité des agents dans le processus d’engagement. Cet objectif nous paraît plus prioritaire, dans un premier temps, qu’assurer un niveau élevé de d’optimisation et de performance.

### 6.3.2 Nombre d’agents partageant une même identité sociale

La variation de la dynamique des coalitions avant et après le processus d’engagement reflète la tendance des agents à changer de choix pour l’adapter à leurs identités. Dans le tableau 6.1, le nombre de *bidders* indique les agents qui ont accepté de participer dans la réalisation d’une tâche en fournissant au moins une action.

Nous remarquons qu’il y a toujours des agents qui refusent de participer à la coalition suite à une décision prise lors du processus d’engagement. Pour des coalitions plus nombreuses, nous remarquons que le nombre de *Definitive Bidders* est réduit car il y a plusieurs actions qui sont attribuées à un même agent. Dans les résultats présentés dans le tableau, nous avons testé l’influence du degré d’*accointance* des agents sur leurs choix des actions. Nous remarquons que le nombre d’agents membres d’une même coalition qui sont de systèmes différents diminue, car les actions sont attribuées à un même agent ou à un maximum d’agents du même système ou qui ”se connaissent”. Nous testons ce comportement en affectant une valeur non nulle aux coefficients  $\omega_4$  dans l’équation 4.1 (et aussi à  $\omega_2$ , pour orienter le comportement des agents vers l’évitement d’interférences négatives).



Librairie de plans		Actions composantes	Équipes 3-robustes			Nb d'ag. avec une id. soc. partagée	
Tâches	Actions		<i>Original bidders</i>	<i>Definitive bidders</i>	Agents retirés (%)	Sans IDEAS	Avec IDEAS
3	10	3	3	3	2 (66%)	0	0
		5	5	5	2 (40%)	0	0
		6	3	3	2 (66%)	1	1
		7	5	5	3 (60%)	1	0
		9	5	5	2 (66%)	0	0
10	30	14	9	9	4 (44%)	1	1
		18	7	7	4 (57%)	1	1
		20	7	7	2 (60%)	1	1
		22	8	9	4 (50%)	1	2
		26	7	7	2 (28%)	1	2
20	60	12	9	10	2 (22%)	2	1
		20	9	9	3 (33%)	3	3
		35	10	12	3 (30%)	3	2
		47	9	7	2 (33%)	2	1
		50	10	9	2 (22%)	3	1
30	90	20	9	9	4 (44%)	4	3
		28	10	9	3 (30%)	4	4
		41	13	10	5 (38%)	5	3
		48	12	12	3 (25%)	6	5
		57	13	11	4 (30%)	7	5

TABLE 6.1 : Tâches aléatoires décomposées en actions composantes.

## 6.4 Coût de l'exploitation du modèle

Dans cette section, nous évaluons les effets de l'exploitation de notre modèle sur la qualité du service fourni par les nouvelles coalitions formées après le processus d'engagement, et le coût en terme de temps d'exécution du processus d'engagement.

### 6.4.1 Comparaison de la qualité de service

Pour pouvoir évaluer la qualité de la solution fournie après le processus d'engagement, nous étudions la variation de la somme des utilités des actions fournies par les agents des coalitions avant et après le processus d'engagement. Comme le montre la Figure 6.2, nous remarquons qu'à cause du retrait de certains agents qui étaient meilleurs enchérisseurs dans la première coalition, la somme des utilités des actions fournies par leurs remplaçants va forcément baisser. Cependant, cette perte reste acceptable même si l'écart avec la meilleure qualité possible devient de plus en plus important dans l'échelle de 60 actions à fournir. Dans tous les cas cet écart reste nettement au dessus de la qualité de la pire des solutions (*Worst Bidders*).

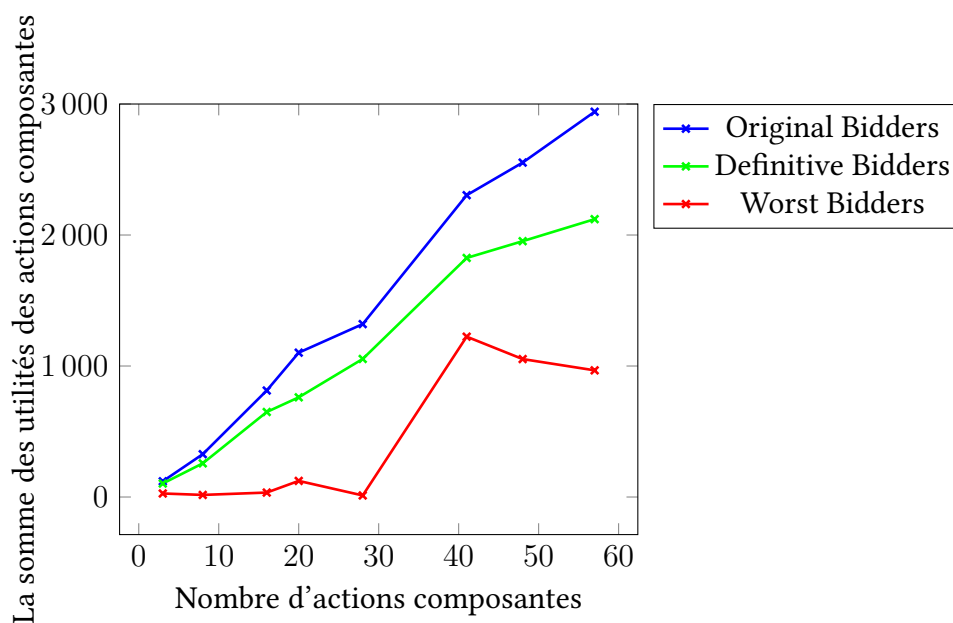


FIGURE 6.2 : La variation de la somme des utilités des actions composantes entre les coalitions formées avant et après le processus d'engagement.

### 6.4.2 Temps d'exécution

Pour pouvoir évaluer le coût de l'implémentation de notre proposition en terme de temps d'exécution, nous montrons, dans la Figure 6.3, l'évolution du temps (en millisecondes) de l'exécution du processus d'engagement centré identité par une équipe 5-robuste d'agents appartenant chacun à un SMA différent. Le temps d'exécution comprend le déroulement du protocole *Extended CNP*, tel qu'un algorithme de reconnaissance d'intention est utilisé pour vérifier les conséquences de chaque attribution d'actions à des agents sur l'identité de chacun. Nous remarquons que, par exemple, une équipe 5-robuste prendra environ 16 secondes pour négocier l'attributions des 50 actions d'une tâche à réaliser collectivement. En ce qui concerne le passage à l'échelle, l'état actuel de l'implémentation du processus de négociation montre ses limites en terme de temps d'exécution, alors que le processus de reconnaissance d'intention garde une évolution polynomiale dans des cas de plus en plus complexes, tandis que le temps pris par le processus basique de *Contract Net Protocol* reste négligeable.

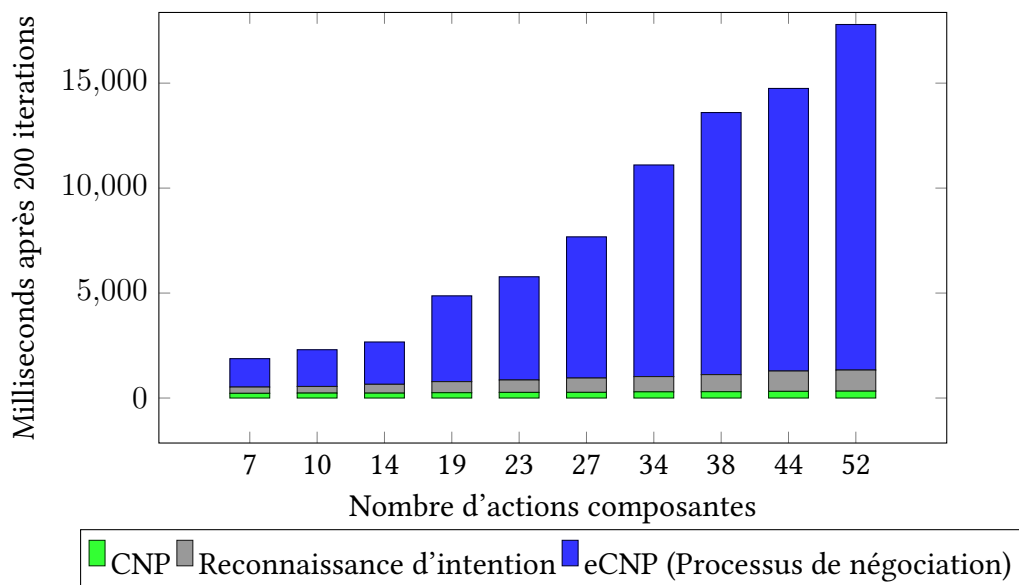


FIGURE 6.3 : Le temps de résolution pour le processus d'engagement par rapport au nombre d'actions composantes pour une équipe 5-robuste.

### 6.4.3 Évaluation de la précision de l'algorithme de reconnaissance d'intention

Les deux processus de notre approche (construction/maintenance d'identité, et engagement) reposent essentiellement sur la prise en compte des significations intentionnelles

des actions des agents. Pour ce faire, nous utilisons MARS [Zhuo and Li, 2011], l'algorithme de reconnaissance de plans multi-agents, en adaptant la fonction qui calcule les poids dans l'algorithme. La précision des résultats obtenus par l'algorithme influence la capacité d'un agent à construire une identité fiable, à base de laquelle il évaluera l'impact de chacune de ses décisions. D'un autre côté, devant à une sollicitation, ou lors d'un processus d'engagement, l'agent s'appuie sur le résultat de l'algorithme pour pouvoir décider s'il accepte ou non l'action en question, après avoir reconnu l'intention et les implications derrière son exécution. Dans la Figure 6.4, nous évaluons la précision de l'algorithme de reconnaissance d'intention en fonction du nombre d'agents et du nombre de tâches (composées d'actions). Nous remarquons que le degré de précision baisse en fonction de la complexité de l'espace de recherche de ces deux facteurs. L'agent observateur réussit moins à identifier les tâches haut-niveau lorsque le nombre d'agents observés augmente. Ceci est dû aux différences entre les bibliothèques de plans des agents. Dans l'ensemble des expérimentations menées dans ce travail, nous supposons que les agents d'un même système partagent une même bibliothèque de plans, et lorsque nous générons un ensemble d'agents d'un nouveau système, la bibliothèque de plans de ce dernier a des tâches/actions manquantes ou supplémentaires par rapport aux autres bibliothèques de plans. Cette différence fait qu'un agent se trouverait dans des situations où il observe des actions qu'il n'a pas dans sa bibliothèque de plans, par exemple, ou à partir de l'observation d'un autre agent, la meilleure explication qu'il donnera à ses actions sera une tâche qui n'existe même pas dans la bibliothèque de plans de l'agent observé. Nous remarquons aussi que la précision des tâches reconnues baisse avec l'augmentation des bibliothèques des plans des agents, car l'agent trouve de plus en plus de tâches qui expliquent la même observation.

## 6.5 Synthèse

Ce chapitre a présenté des résultats d'expérimentations sur des agents appartenant à différents SMA coexistants. Ces résultats permettent d'évaluer à quel point notre approche peut contribuer à permettre une meilleure coopération des agents, en prenant en considération leurs identités et les significations intentionnelles des actions et des interactions des autres. Les expérimentations sont principalement basées sur la décision de chaque agent de s'engager ou non à coopérer avec les autres. Nous avons évalué différents comportements que les agents peuvent montrer en fonction de comment chacun considère son identité. La première remarque qui s'impose est que le comportement des agents

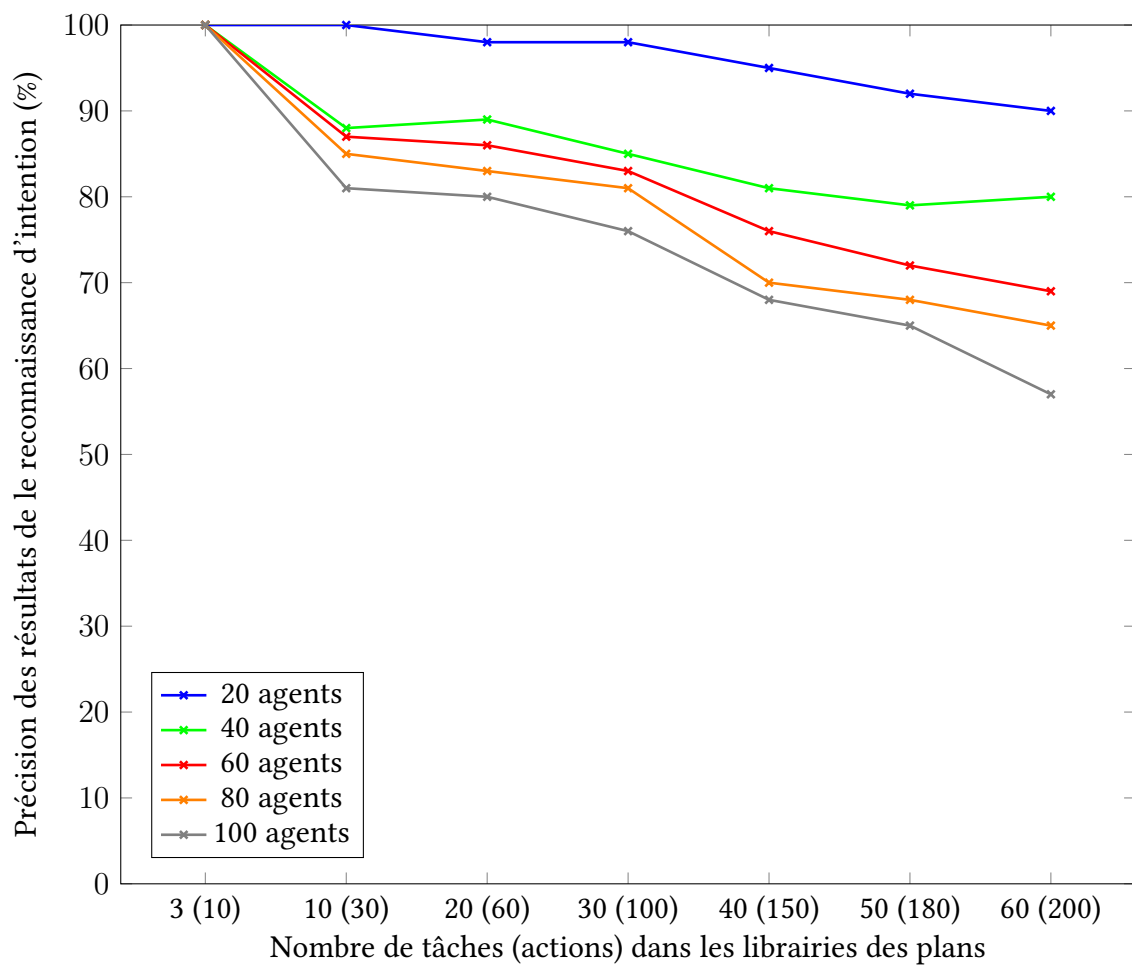


FIGURE 6.4 : L'évolution de la précision de l'algorithme de reconnaissance d'intention.

change d'une manière qui montre la réussite et la pertinence de notre inspiration de la théorie de l'identité comme cadre théorique pour modéliser nos agents. Cette réussite est soutenue par la paramétrabilité du modèle proposé. Nous avons illustré des cas où chaque agent cherche à préserver son identité en évitant des interférences négatives, à promouvoir son identité en cherchant à maximiser le nombre d'interférences positives, ou à montrer une sorte de cohésion en favorisant la coopération avec des agents qui partagent la même identité sociale.

Il y a plusieurs pistes d'amélioration de l'implémentation. Les résultats montrent un problème de passage à l'échelle de l'implémentation dans son état actuel. La première piste d'amélioration de la performance de notre solution concerne la modélisation du problème, en allégeant la topologie du *factor graph* par des heuristiques qui définissent quels agents doivent être mis en négociation, et quelles actions seront sujet de ces négociations, au lieu de faire négocier tous les agents de la coalition à former. Une autre amélioration concerne la précision de l'algorithme de reconnaissance d'intention. Ce dernier se base sur l'utilité et l'observabilité des actions observées pour reconnaître les tâches haut-niveau qui les impliquent. Nous l'avons amélioré en prenant en compte l'applicabilité d'une action étant donné l'état de l'environnement. Nous pouvons nous baser sur d'autres critères comme le degré de confiance en l'agent observé ou la fréquence de l'exécution de la tâche à reconnaître. Nous pouvons aussi améliorer les résultats obtenus en adaptant les choix techniques tels qu'un meilleur algorithme de résolution du problème de négociation, et les choix des technologies et protocoles de communication entre les agents.

Dans l'état actuel de l'implémentation, nous avons utilisé la version originale de l'algorithme Max-Sum, sans extension ou amélioration, sur un graphe de contraintes qui pourrait être plus efficace pour faciliter la recherche d'une solution optimale. De plus, pour converger vers cette solution, l'algorithme dépasse facilement les 200 itérations pour une équipe potentielle de 10 agents sollicités pour réaliser une tâche décomposable en 50 actions. Cependant, pour développer des applications WoT dans le cadre du projet ANR ASAWoO, il est nécessaire de respecter le style architectural REST, et utiliser HTTP comme protocole de communication entre les agents. Ces choix techniques nous contraignent à sacrifier le temps d'exécution en raison du nombre d'échanges entre les objets. D'un côté, REST comme style architectural convient très bien à l'exposition et la manipulation des données ou des services des agents sur le Web, mais il n'est pas adapté aux systèmes basés sur des processus d'exécution [Xu et al., 2008]. D'un autre côté, HTTP n'offre pas les mêmes performances comme d'autres protocoles tels que CoAP (*Constrained Application Protocol*) [Bormann et al., 2012] ou MQTT (*Message Queuing Telemetry*

*Transport*) [Banks and Gupta, 2014]. Nous avons montré ces limitations dans un de nos travaux [Khalfi et al., 2016], où nous avons modélisé le problème d'allocation de tâches dans un contexte Web des Objets comme un problème de satisfaction de contraintes (Distributed Constraint Satisfaction Problem - DisCSP [Yokoo et al., 1992]), que nous résolvons avec l'algorithme ABT (Asynchronous BackTracking) [Yokoo et al., 1992]. Enfin, l'objectif de cette thèse, en terme de construction pratique, est de fournir une preuve de concept de la pertinence de la théorie de l'identité comme cadre théorique permettant aux agents d'appréhender leurs interactions sociales.





# Conclusion

Les agents embarqués déployés dans des environnements physiques sont de plus en plus interopérables et amenés à coexister avec des agents d'autres systèmes dans un même espace physique. Un environnement n'est plus la propriété d'un unique SMA. La coopération ne se limite alors plus aux agents d'un même SMA : un agent doit donc être capable de répondre à des sollicitations exprimées de manière opportuniste par des agents extérieurs à son système.

## Bilan

Notre question globale de recherche porte sur la préservation des intérêts individuels et collectifs d'un agent lors de sa coopération dans un environnement *multi-SMA*. Pour ce faire, un agent est amené à considérer deux facteurs : le premier étant l'identification et la représentation explicite de ses intérêts, le deuxième est le raisonnement sur l'implication des actions et des interactions des autres agents sur ses intérêts identifiés. Ayant travaillé sur les notions et les concepts fondamentaux autour des interférences dans le cadre des interactions coopératives, nous avons divisé la question globale de recherche en deux questions.

La première sous-question est celle de la représentation des intérêts individuels et collectifs d'un agent. Le choix de la *théorie de l'identité* comme cadre théorique pour répondre à cette question est motivé par son intérêt, dans la psychologie sociale, à étudier la dynamique inter-personnelle, intra-personnelle et inter-groupes des individus. Elle permet la définition de leurs comportements dans des contextes sociaux variés. À travers la définition de l'*identité personnelle*, de l'*identité sociale* et de l'*identité collective*, nous avons

mis en évidence plusieurs similarités entre les éléments constitutifs de l'identité d'un individu et les composants de différentes architectures d'agent étudiées dans la littérature SMA. Cette étude nous a permis d'identifier l'élément de base pour répondre à la question soulevée : il s'agit d'un processus de construction et de maintenance de l'identité d'un agent. Ce processus est illustré dans un modèle d'agent que nous avons appelé IDEAS, dont la première implémentation est basée sur le modèle BDI et des composants inspirés d'autres architectures d'agents qui font référence à des éléments définissant l'identité. Ensuite, nous avons décrit le cycle de vie d'un agent IDEAS et mis l'accent sur la mise à jour de son identité.

La deuxième sous-question porte sur le raisonnement d'un agent sur les sollicitations des autres, c'est-à-dire comment doter un agent de la capacité d'accepter de répondre favorablement ou non à la sollicitation d'un autre agent. L'intérêt de cette question provient de la difficulté qu'a un agent à décider de coopérer ou non avec des agents d'autres SMA avec lesquels il n'a jamais interagi. La deuxième difficulté de cette sous-question se manifeste dans la nature des interactions coopératives entre des agents dans des contextes multi-SMA, qui se caractérise par la formation de coalitions ad hoc d'une manière opportuniste et temporaire. Pour répondre à ce défi, nous nous sommes inspirés d'un concept qui est à la base de l'interaction humaine, à savoir l'*intersubjectivité*. Ce concept consiste à prendre en considération la reconnaissance de la signification intentionnelle des activités des autres agents dans l'interaction avec eux. Nous avons proposé la reconnaissance d'intention multi-agent comme mécanisme implémentant le concept d'intersubjectivité. Ce mécanisme est utilisé dans un processus d'engagement pour permettre aux agents de choisir quelles sont les actions qu'ils vont s'engager à exécuter, tout en préservant leurs identités. La décision d'engagement est prise en fonction des tâches reconnues à partir des observations de l'activité des agents du voisinage. En lien avec la première question, la reconnaissance d'intention est aussi utilisée pour permettre à un agent d'identifier quelles tâches haut-niveau un agent partage avec les autres agents, et quelles coalitions sont formées pour réaliser ces tâches.

D'un point de vue pratique, nous avons implémenté notre contribution dans un contexte applicatif Web des objets, qui consiste à permettre à des objets physiques d'être accessibles sur le Web d'une manière interopérable en utilisant des standards et des protocoles du Web. Il y a plusieurs approches pour implémenter un système Web des objets avec les capacités des systèmes multi-agents. La première classe d'approches s'intéresse à *agentifier* les services Web (où les capacités des objets physiques sont exposées). Dans cette classe, on tend vers la création d'agents clients ou proxy de services Web qui contrôlent leurs appels. La deuxième classe d'approches consiste à intégrer des techno-

logies Web dans des plateformes multi-agents pour exposer les services internes d'un agent comme des services accessibles sur le Web. La première motivation de cette classe d'approches est de permettre à la fois l'interopérabilité entre plusieurs plateformes mais aussi à des services Web de communiquer avec des agents. Notre approche, basée sur la notion d'*avatar*, se distingue des deux classes dans le sens où nous cherchons à avoir des agents *vivants* dans le Web, embarqués dans chaque objet physique, tels que les protocoles, les standards et les technologies utilisées dans la communication, et les données des agents sont issues des recommandations du Web des objets. Dans notre implémentation, les agents sont des ressources REST, échangeant des données sémantiques en RDF (Resource Description Framework), sur lesquelles ils peuvent raisonner avec le langage SPARQL. Ces échanges comprennent des données comme des états de capteurs, des connaissances comme des bibliothèques de plans et des représentations de l'environnement, et aussi des processus comme la négociation et l'allocation de tâches.

Comme mentionné dans la section précédente, nous proposons le concept d'identité pour représenter les intérêts individuels et collectifs qu'un agent doit prendre en compte lors de sa coopération avec les autres. Le travail présenté dans ce manuscrit a focalisé sur une partie de la définition de l'identité. Nous avons réduit, dans un premier temps, la représentation de l'identité d'un agent aux objectifs qu'il est en train de poursuivre et aux coalitions auxquelles il appartient. Ces deux éléments sont au cœur des décisions et des choix sociaux d'un humain, mais l'identité (personnelle ou sociale) est loin d'être limitée à ces deux éléments. Ce choix de focalisation sur les objectifs et les coalitions est principalement lié au contexte technique de cette thèse qui est le Web des Objets. Dans ce contexte, chaque objet physique est installé pour détecter ou agir sur un environnement, et éventuellement interagir avec d'autres objets pour la réalisation de tâches plus complexes. Dans le cadre du projet ASAWoO, nous étendons chaque objet physique en proposant une représentation virtuelle telle que l'objet et sa représentation forment un objet cyber-physique accessible sur le Web. Pour ce faire, nous associons à chaque objet un *avatar* qui augmentera ses connaissances et ses capacités. Dans ces conditions, nous trouvons que pour avoir une représentation des intérêts d'un objet physique, les objectifs et les coalitions sont plus prioritaires que d'autres éléments, comme les rôles ou les émotions. Notre objectif était d'initier une réflexion sur les contraintes que peuvent rencontrer les agents dans des SMA coexistants, jeter un pont entre la théorie de l'identité et les systèmes multi-agents, et fournir une preuve de concept de la pertinence de ce cadre théorique. Les expérimentations présentées dans ce manuscrit montrent des résultats que nous pensons prometteurs au niveau de la qualité de la solution produite. Dans ces expérimentations, un agent intègre un processus de mise à jour de l'identité

dans son cycle de raisonnement, et un processus de décision lors de l'engagement à coopérer avec les autres. Ce dernier processus (modélisé comme un DCOP) nécessite plusieurs échanges de messages entre les agents pour qu'ils évaluent chacun l'impact de chacune des actions choisies en fonction des choix des autres agents. La première limitation concerne le choix de l'algorithme de résolution du problème DCOP. Dans la littérature, il est démontré que la classe d'algorithmes *non-exacts* est la plus adaptée à notre contexte, mais il nous reste tout de même le choix d'un algorithme dans cette classe, et aussi le choix de l'extension qui optimise au mieux le temps de convergence des agents vers la solution optimale.

## Perspectives

Nous nous sommes intéressés à permettre à un agent de construire et maintenir une identité pour le doter de la capacité de répondre à des sollicitations, tout en préservant ses intérêts individuels et collectifs. Un travail similaire à ce que nous avons mené pour le cas d'un agent sollicité est à effectuer pour les agents sollicités. Il nous paraît intéressant d'étudier les travaux de la *saillance d'identité* [Stryker and Serpe, 1982] pour permettre à un agent de choisir quelle identité « activer » lors de l'initiation d'une nouvelle coopération. En ce qui concerne l'inspiration de la théorie d'identité, et si nous prenons en compte la richesse des travaux effectués dans la littérature, nous devons approfondir notre étude de ce cadre théorique au delà de la présentation de quelques types d'identité, et l'implémentation d'un processus de construction d'identité, comme l'utilisation d'autres processus basés sur l'identité (comme l'influence sociale [Turner, 1991], par d'exemple).

Une autre perspective de cette thèse concerne le niveau conceptuel de l'étude des interactions sociales de nos agents. Dans la continuité des travaux qui se font dans l'équipe, nous pouvons d'améliorer les mécanismes d'exploitation de productions collectifs d'un SMA par un autre [Khenifar-Bessadi et al., 2017]. L'intérêt à considérer des SMA au niveau collectif nous motive à approfondir notre étude des identités collectives [Ashmore et al., 2004, Gardner and Garr-Schultz, 2017, Dovidio and Schellhaas, 2017] qui proposent des concepts et processus caractérisant un groupe d'agents. Ainsi, nous pouvons développer des processus sociaux permettant à un groupe d'agents de se créer une identité collective pour se représenter comme une seule entité, capable de solliciter d'autres entités collectives, d'exposer des actions collectives que le groupe peut réaliser, de recruter de nouveaux membres et promouvoir cette identité collective.

# Annexe A

```
<avatar/A1> a <Avatar>, sosa:Platform;
  rdfs:comment "un agent déployé dans un drone agricole."@fr ;
  sosa:hosts <drone/185> ;
  asawoo:beliefs <avatar/A1/beliefs> ;
  asawoo:intentions <avatar/A1/intentions> ;
  asawoo:events <avatar/A1/events> ;
  asawoo:plan-library <avatar/A1/plan-library> ;
  asawoo:social-identity <avatar/A1/social-identity>
.

<drone/185> a ssn:System ;
  ssn:hasSubSystem <drone/185/LightSensor/1>,
                  <drone/185/QuadMultispectralSensor/1>,
                  <drone/185/spray/1437> ;

# Le graphe contenant l'ensemble des évènements de l'avatar
GRAPH <avatar/A1/events> {
  _:b0 rdf:type asawoo:Achieve ;
    <delegated-goal> <Task/Weeding> ;
    <has-parameter> <parameter/row> ;
    <parameters-info> _:b1
  .
  _:b2 rdf:type asawoo:Achieve ;
    <delegated-goal> <Task/DetectWateringNeeds> ;
    <has-parameter> <parameter/field> ;
    <parameters-info> _:b3
  .
}

_:b1 {
  <parameter/row> <value> <field/4/row/207> .
}

_:b3 {
  <parameter/field> <value> <field/3> .
}

# Le graphe contenant l'ensemble des croyances de l'avatar
GRAPH <avatar/A1/beliefs> {
  <field/1> rdf:type asawoo:Field ;
    <has-row> <field/1/row/758>
  .
  <agribot/254> <isLocatedIn> <field/1/row/758> .
  <drone/185> <isLocatedIn> <field/5> .
  <AtmosphericPressureSensor/485> rdf:type sosa:Sensor ;
    sosa:observes <air_pressure> ;
  .
}
```

```

    sosa:madeObservation <Observation/578575>
.
<TemperatureSensor/971> rdf:type sosa:Sensor ;
  sosa:observes <air_temperature> ;
  sosa:madeObservation <Observation/149248>
.
<TemperatureSensor/138> rdf:type sosa:Sensor ;
  sosa:observes <field/5/soil_temperature> ;
  sosa:madeObservation <Observation/245474>
.
<PrecipitationSensor/319> rdf:type sosa:Sensor ;
  sosa:observes <field/5/rainfall_rate>, <field/5/thickness_of_rainfall_amount> ;
  sosa:madeObservation <Observation/439712>, <Observation/318447>
.
<WindSensor/138> rdf:type sosa:Sensor ;
  sosa:observes <wind_speed>, <wind_from_direction> ;
  sosa:madeObservation <Observation/139471>
.
<Observation/578575> rdf:type sosa:Observation ;
  sosa:hasSimpleResult "1027 mbar"^^cdt:atmosphericPressure
.
<agribot/254/move/4276> rdf:type asawoo:Action, sosa:Actuation ;
  sosa:madeByActuator <agribot/254> ;
  sosa:hasFeatureOfInterest <field/1/row/758> ;
  sosa:phenomenonTime [
    rdf:type time:Instant ;
    time:inXSDDateTimeStamp "2017-03-31T11:51:42+00:00"^^xsd:dateTimeStamp ]
.
<agribot/254/spray/2416> rdf:type asawoo:Action, sosa:Actuation ;
  sosa:madeByActuator <agribot/254> ;
  sosa:hasFeatureOfInterest <field/1/row/758> ;
  sosa:actsOnProperty <field/1/row/758/weeds> ;
  sosa:phenomenonTime [
    rdf:type time:Interval ;
    time:hasBeginning [
      rdf:type time:Instant ;
      time:inXSDDateTimeStamp "2017-03-31T11:55:13+00:00"^^xsd:dateTimeStamp ] ;
    time:hasEnd [
      rdf:type time:Instant ;
      time:inXSDDateTimeStamp "2017-03-31T12:15:37+00:00"^^xsd:dateTimeStamp ] ]
.
<drone/434/fly/1267> rdf:type asawoo:Action, sosa:Actuation ;
  sosa:madeByActuator <drone/434> ;
  sosa:hasFeatureOfInterest <field/5> ;
  sosa:phenomenonTime [
    rdf:type time:Instant ;
    time:inXSDDateTimeStamp "2017-03-28T13:14:04+00:00"^^xsd:dateTimeStamp ]
.
<drone/434/takepicture/2107> rdf:type asawoo:Action, sosa:Observation ;
  sosa:madeBySensor <drone/434> ;
  sosa:hasFeatureOfInterest <field/5> ;
  sosa:observedProperty <field/5/aerialimage> ;
  sosa:resultTime "2017-03-28T13:16:49Z"^^xsd:date
.
<drone/185/spray/1437> rdf:type asawoo:Action, sosa:Actuation ;
  sosa:madeByActuator <drone/185> ;
  sosa:hasFeatureOfInterest <field/5> ;
  sosa:actsOnProperty <field/5/rainfall_rate> ;
  sosa:phenomenonTime [
    rdf:type time:Interval ;
    time:hasBeginning [
      rdf:type time:Instant ;
      time:inXSDDateTimeStamp "2017-03-28T15:04:00+00:00"^^xsd:dateTimeStamp ] ;
    time:hasEnd [
      rdf:type time:Instant ;

```

```

        time:inXSDDateTimeStamp "2017-03-28T15:27:00+00:00"^^xsd:dateTimeStamp ] ]
}
.
# Le graphe contenant l'ensemble des intentions de l'avatar
GRAPH <avatar/A1/intentions> {
  [] a asawoo:intention ;
  asawoo:actionToPerform <Action/TransferPicture> ;
  asawoo:actionToPerform <Action/ProcessPicture> ;
  asawoo:has-source _:b0
}
.
# Le graphe représentant l'identité sociale de l'avatar
GRAPH <avatar/A1/social-identity> {
  [] a asawoo:recognition ;
  asawoo:recognizedavatar <agribot/254> ;
  asawoo:recognizedtask <Task/Weeding> ;
  asawoo:recognition-score "89.45842271050679"^^xsd:double
.
  [] a asawoo:recognition ;
  asawoo:recognizedavatar <drone/434> ;
  asawoo:recognizedavatar <drone/185> ;
  asawoo:recognizedtask <Task/ManagePartofFiendWatering> ;
  asawoo:recognition-score "23.01248579745232"^^xsd:double
.
  [] a asawoo:coalition ;
  asawoo:hasTeamMember [
    [] a asawoo:teamMembership ;
    asawoo:teamMember <drone/434> ;
    asawoo:performedAction <drone/434/fly/1267> . ] ;
  asawoo:hasTeamMember [
    [] a asawoo:teamMembership ;
    asawoo:teamMember <drone/434> ;
    asawoo:performedAction <drone/434/takepicture/2107> . ] ;
  asawoo:hasTeamMember [
    [] a asawoo:teamMembership ;
    asawoo:teamMember <drone/185> ;
    asawoo:performedAction <drone/185/spray/1437> . ] ;
  asawoo:hasGoal <Task/ManagePartofFiendWatering>
}
.
# Le graphe contenant la librairie des plans de l'avatar
GRAPH <avatar/A1/plan-library> {
  <Task/managepartoffieldwatering> a asawoo:Task ;
  asawoo:has-flow <Flow/managepartoffieldwatering-flow1> .
  <Flow/managepartoffieldwatering-flow1> a asawoo:MethodFlow ;
  asawoo:has-postcondition [
    a asawoo:PlanStep ;
    asawoo:has-step <Method/detectwateringneeds> ;
    asawoo:order 3 ] ;
  asawoo:has-postcondition [
    a asawoo:PlanStep ;
    asawoo:has-step <Method/checkfordiseases> ;
    asawoo:order 2 ] ;
  asawoo:has-postcondition [
    a asawoo:PlanStep ;
    asawoo:has-step <Action/WaterField> ;
    asawoo:order 1 ]
.
  <Method/detectwateringneeds> a asawoo:Method ;
  asawoo:has-flow <Flow/detectwateringneeds-flow1>
.
  <Flow/detectwateringneeds-flow1> a asawoo:MethodFlow ;
  asawoo:has-postcondition [
    a asawoo:PlanStep ;
    asawoo:has-step <Action/ProcessPicture> ;
    asawoo:order 2 ] ;

```

```

asawoo:has-postcondition [
  a asawoo:PlanStep ;
  asawoo:has-step <Action/TransferPicture> ;
  asawoo:order 4 ] ;
asawoo:has-postcondition [
  a asawoo:PlanStep ;
  asawoo:has-step <Action/TakePicture> ;
  asawoo:order 3 ] ;
asawoo:has-postcondition [
  a asawoo:PlanStep ;
  asawoo:has-step <Action/Fly> ;
  asawoo:order 1 ]
.
<Action/ProcessPicture> a asawoo:Action ;
  <has-paramater> <Parameter/field>
.
<Action/TakePicture> a asawoo:Action ;
  <has-paramater> <Parameter/field>
.
<Action/TransferPicture> a asawoo:Action ;
  <has-paramater> <Parameter/field>
.
<Action/Fly> a asawoo:Action ;
  <has-paramater> <Parameter/field>
.
<Method/checkfordiseases> a asawoo:Method ;
  asawoo:has-flow <Flow/checkfordiseases-flow1>
.
<Flow/checkfordiseases-flow1> a asawoo:Flow ;
  asawoo:has-postcondition [
    a asawoo:PlanStep ;
    asawoo:has-step <Method/GroundCheck> ;
    asawoo:order 1 ] ;
  asawoo:has-postcondition [
    a asawoo:PlanStep ;
    asawoo:has-step <Method/AerialCheck> ;
    asawoo:order 2 ]
.
<Method/groundcheck> a asawoo:Method ;
  asawoo:has-flow <Flow/groundcheck-flow1> .
<Method/groundcheck-flow1> a asawoo:MethodFlow ;
  asawoo:has-postcondition [
    a asawoo:PlanStep ;
    asawoo:has-step <Action:Move> ;
    asawoo:order 1 ] ;
  asawoo:has-postcondition [
    a asawoo:PlanStep ;
    asawoo:has-step <Action/TakePicture> ;
    asawoo:order 2 ] ;
  asawoo:has-postcondition [
    a asawoo:PlanStep ;
    asawoo:has-step <Action/ProcessPicture> ;
    asawoo:order 3 ]
.
<Method/aerialcheck> a asawoo:Method ;
  asawoo:has-flow <Flow/aerialcheck-flow1> .
<Flow/aerialcheck-flow1> a asawoo:MethodFlow ;
  asawoo:has-postcondition [
    a asawoo:PlanStep ;
    asawoo:has-step <Action/processpicture> ;
    asawoo:order 3 ] ;
  asawoo:has-postcondition [
    a asawoo:PlanStep ;
    asawoo:has-step <Action/transferpicture> ;
    asawoo:order 4 ] ;

```



```
asawoo:has-postcondition [
  a asawoo:PlanStep ;
  asawoo:has-step <Action/takepicture> ;
  asawoo:order 2 ] ;
asawoo:has-postcondition [
  a asawoo:PlanStep ;
  asawoo:has-step <Action/fly> ;
  asawoo:order 1 ]
.
<Method/weeding> a asawoo:Method ;
  asawoo:has-flow <Flow/weeding-flow2>, <Flow/weeding-flow1> .
<Flow/weeding-flow2> a asawoo:MethodFlow ;
  asawoo:has-postcondition [
    a asawoo:PlanStep ;
    asawoo:has-step <Action/weed> ;
    asawoo:order 2 ] ;
  asawoo:has-postcondition [
    a asawoo:PlanStep ;
    asawoo:has-step <Action/move> ;
    asawoo:order 1 ]
.
<Flow/weeding-flow1> a asawoo:MethodFlow ;
  asawoo:has-postcondition [
    a asawoo:PlanStep ;
    asawoo:has-step <Action/spray> ;
    asawoo:order 2 ] ;
  asawoo:has-postcondition [
    a asawoo:PlanStep ;
    asawoo:has-step <Action/fly> ;
    asawoo:order 1 ]
.
<Method/fertilization> a asawoo:Method ;
  asawoo:has-flow <Flow/fertilization-flow1> .
<Flow/fertilization-flow1> a asawoo:MethodFlow ;
  asawoo:has-postcondition [
    a asawoo:PlanStep ;
    asawoo:has-step <Action/fertilize> ;
    asawoo:order 1 ] ;
  asawoo:has-postcondition [
    a asawoo:PlanStep ;
    asawoo:has-step <Action/move> ;
    asawoo:order 2 ]
.
}
```



# Bibliographie

- D. Abrams and M. A. Hogg. *Social identifications : A social psychology of intergroup relations and group processes*. Routledge, 2006.
- M. R. Adler, A. B. Davis, R. Weihmayer, and R. W. Worrest. Conflict-resolution strategies for nonhierarchical distributed agents. In *Distributed artificial intelligence*, pages 139–161. Elsevier, 1998.
- S. Aknine, S. Pinson, and M. F. Shakun. An extended multi-agent negotiation protocol. *Autonomous Agents and Multi-Agent Systems*, 8(1) :5–45, 2004.
- J. Allen, H. Kautz, R. Pelavin, and J. Tennenberg. A formal theory of plan recognition and its implementation. *Reasoning About Plans*, pages 69–126, 1991.
- W. Alshabi, S. Ramaswamy, M. Itmi, and H. Abdulrab. Coordination, cooperation and conflict resolution in multi-agent systems. In *Innovations and advanced techniques in computer and information sciences and engineering*, pages 495–500. Springer, 2007.
- R. D. Ashmore, K. Deaux, and T. McLaughlin-Volpe. An organizing framework for collective identity : articulation and significance of multidimensionality. *Psychological bulletin*, 130(1) :80, 2004.
- L. Atzori, A. Iera, G. Morabito, and M. Nitti. The social internet of things (siot)—when social networks meet the internet of things : Concept, architecture and network characterization. *Computer Networks*, 56(16) :3594–3608, 2012.
- J. Bach. The micropsi agent architecture. In *Proceedings of ICCM-5, international conference on cognitive modeling, Bamberg, Germany*, pages 15–20, 2003.

- R. P. Bagozzi. On the concept of intentional social action in consumer behavior. *Journal of Consumer Research*, 27(3) :388–396, 2000.
- M. Baldoni, C. Baroglio, F. Capuzzimati, and R. Micalizio. Commitment-based agent interaction in jacamo+. *Fundamenta Informaticae*, 159(1-2) :1–33, 2018.
- T. Balke and N. Gilbert. How do agents make decisions? a survey. *Journal of Artificial Societies and Social Simulation*, 17(4) :13, 2014.
- B. Banerjee, L. Kraemer, and J. Lyle. Multi-agent plan recognition : Formalization and algorithms. In *AAAI*, 2010.
- B. Banerjee, J. Lyle, and L. Kraemer. The complexity of multi-agent plan recognition. *Autonomous Agents and Multi-Agent Systems*, 29(1) :40–72, 2015.
- A. Banks and R. Gupta. Mqtt version 3.1. 1. *OASIS standard*, 29, 2014.
- K. Barber, T. Liu, and D. Han. Strategic decision-making for conflict resolution in dynamic organized multi-agent systems. *A Special Issue of CERA Journal*, 2000.
- G. S. Basheer, M. S. Ahmad, and A. Y. Tang. A framework for conflict resolution in multi-agent systems. In *International Conference on Computational Collective Intelligence*, pages 195–204. Springer, 2013.
- N. Blaylock and J. Allen. Corpus-based, statistical goal recognition. In *Proceedings of the 18th International Joint Conference on Artificial Intelligence, IJCAI’03*, pages 1303–1308, San Francisco, CA, USA, 2003. Morgan Kaufmann Publishers Inc. URL <http://dl.acm.org/citation.cfm?id=1630659.1630845>.
- N. Blaylock and J. Allen. Fast hierarchical goal schema recognition. In *Proceedings of the National Conference on Artificial Intelligence*, volume 21, page 796. Menlo Park, CA ; Cambridge, MA ; London ; AAAI Press ; MIT Press ; 1999, 2006.
- A. L. Blum and M. L. Furst. Fast planning through planning graph analysis. *Artificial intelligence*, 90(1-2) :281–300, 1997.
- R. H. Bordini, L. Braubach, M. Dastani, A. E. F. Seghrouchni, J. J. Gomez-Sanz, J. Leite, G. O’Hare, A. Pokahr, and A. Ricci. A survey of programming languages and platforms for multi-agent systems. *Informatica*, 30(1), 2006.
- R. H. Bordini, J. F. Hübner, and M. Wooldridge. *Programming multi-agent systems in AgentSpeak using Jason*, volume 8. John Wiley & Sons, 2007.

- R. H. Bordini, M. Dastani, J. Dix, and A. E. F. Seghrouchni. *Multi-Agent Programming*. Springer, 2009.
- C. Bormann, A. P. Castellani, and Z. Shelby. Coap : An application protocol for billions of tiny internet nodes. *IEEE Internet Computing*, 16(2) :62–67, 2012.
- L. Braubach, A. Pokahr, and W. Lamersdorf. A universal criteria catalog for evaluation of heterogeneous agent development artifacts. *From Agent Theory to Agent Implementation (AT2AI-6)*, pages 19–28, 2008.
- M. B. Brewer. The many faces of socialidentity : Implications for political psychology. *Political psychology*, 22(1) :115–125, 2001.
- M. B. Brewer and W. Gardner. Who is this” we” ? levels of collective identity and self representations. *Journal of personality and social psychology*, 71(1) :83, 1996.
- P. J. Burke. An identity model for network exchange. *American Sociological Review*, 62 (1) :134–150, 1997. ISSN 00031224. URL <http://www.jstor.org/stable/2657456>.
- P. J. Burke. *Contemporary social psychological theories*. Stanford University Press, 2006.
- P. J. Burke and D. C. Reitzes. The link between identity and role performance. *Social psychology quarterly*, pages 83–92, 1981.
- P. J. Burke and J. E. Stets. *Identity theory*. Oxford University Press, 2009.
- R. Camacho, P. Carreira, I. Lynce, and S. Resendes. An ontology-based approach to conflict resolution in home and building automation systems. *Expert Systems with Applications*, 41(14) :6161–6173, 2014.
- J. Campos, C. Martinho, and A. Paiva. Conflict inside out : A theoretical approach to conflict from an agent point of view. In *Proceedings of the 2013 international conference on Autonomous agents and multi-agent systems*, pages 761–768. International Foundation for Autonomous Agents and Multiagent Systems, 2013.
- D. Capera, J.-P. Georgé, M.-P. Gleizes, and P. Glize. The amas theory for complex problem solving based on self-organizing cooperative agents. In *Enabling Technologies : Infrastructure for Collaborative Enterprises, 2003. WET ICE 2003. Proceedings. Twelfth IEEE International Workshops on*, pages 383–388. IEEE, 2003.

- C. Castelfranchi. Commitments : From individual intentions to groups and organizations. In *Proceedings of the First International Conference on Multiagent Systems (ICMAS)*, volume 95, pages 41–48, 1995.
- C. Castelfranchi. Modelling social action for ai agents. *Artificial Intelligence*, 103(1) : 157–182, 1998.
- C. Castelfranchi. The cognition of conflict : ontology, dynamics, and ideology. In *Conflict and Multimodal Communication*, pages 3–32. Springer, 2015.
- C. Castelfranchi and R. Falcone. Conflicts within and for collaboration. In *Conflicting agents*, pages 33–61. Springer, 2002.
- A. Cesta, M. Miceli, and P. Rizzo. Effects of different interaction attitudes on a multi-agent system performance. In *European Workshop on Modelling Autonomous Agents in a Multi-Agent World*, pages 128–138. Springer, 1996.
- E. Charniak and R. P. Goldman. A bayesian model of plan recognition. *Artificial Intelligence*, 64(1) :53–79, 1993.
- A. Ciorcea, O. Boissier, A. Zimmermann, and A. M. Florea. Reconsidering the social web of things : position paper. In *Proceedings of the 2013 ACM conference on Pervasive and ubiquitous computing adjunct publication*, pages 1535–1544. ACM, 2013.
- P. R. Cohen and H. J. Levesque. Intention is choice with commitment. *Artificial intelligence*, 42(2-3) :213–261, 1990.
- P. R. Cohen and H. J. Levesque. Teamwork. *Nous*, 25(4) :487–512, 1991.
- R. Conte and C. Castelfranchi. Mind is not enough. precognitive bases of social interaction. In *Proceedings of 1992 Symposium on Simulating Societies*, pages 93–110, 1992.
- C. H. Cooley. Human nature and the social order (new york : Scribner’s, 1902). *Social Organization*, 1902.
- M. Dastani and L. van der Torre. Programming boid-plan agents deliberating about conflicts among defeasible mental attitudes and plans. In *Autonomous Agents and Multiagent Systems, 2004. AAMAS 2004. Proceedings of the Third International Joint Conference on*, pages 706–713. IEEE, 2004.
- H. De Jaegher, E. Di Paolo, and S. Gallagher. Can social interaction constitute social cognition? *Trends in cognitive sciences*, 14(10) :441–447, 2010.

- M. De Weerd, A. Ter Mors, and C. Witteveen. Multi-agent planning : An introduction to planning and coordination. In *In : Handouts of the European Agent Summer*. Citeseer, 2005.
- J. D. DeLamater and A. Ward. *Handbook of social psychology*. Springer, 2006.
- J. A. Dias, J. J. Rodrigues, and L. Zhou. Cooperation advances on vehicular communications : A survey. *Vehicular communications*, 1(1) :22–32, 2014.
- J. F. Dovidio and F. M. Schellhaas. Collective identity. *The Wiley-Blackwell Encyclopedia of Social Theory*, pages 1–3, 2017.
- A. Eguchi, H. Nguyen, and C. W. Thompson. Everything is alive : towards the future wisdom web of things. *World Wide Web*, 16(4) :357–378, 2013.
- N. Ellemers, R. Spears, and B. Doosje. Sticking together or falling apart : In-group identification as a psychological determinant of group commitment versus individual mobility. *Journal of personality and social psychology*, 72(3) :617, 1997.
- A. Farinelli, A. Rogers, A. Petcu, and N. R. Jennings. Decentralised coordination of low-power embedded devices using the max-sum algorithm. In *Proceedings of the 7th international joint conference on Autonomous agents and multiagent systems-Volume 2*, pages 639–646. International Foundation for Autonomous Agents and Multiagent Systems, 2008.
- J. Ferber. *Les systèmes multi-agents : vers une intelligence collective*. InterEditions, 1995.
- J. Ferber. *Multi-agent systems : an introduction to distributed artificial intelligence*, volume 1. Addison-Wesley Reading, 1999.
- R. T. Fielding. *Architectural styles and the design of network-based software architectures*. PhD thesis, University of California, 2000.
- R. J. Fisher. *The social psychology of intergroup and international conflict resolution*. Springer Science & Business Media, 2012.
- A. Freitas, D. Schmidt, A. Panisson, F. Meneguzzi, R. Vieira, and R. H. Bordini. Semantic representations of agent plans and planning problem domains. In *International Workshop on Engineering Multi-Agent Systems*, pages 351–366. Springer, 2014.

- W. L. Gardner and A. Garr-Schultz. Understanding our groups, understanding ourselves : the importance of collective identity clarity and collective coherence to the self. In *Self-Concept Clarity*, pages 125–143. Springer, 2017.
- B. Garrett. *Personal identity and self-consciousness*. Routledge, 2002.
- G. Gasser and M. Stefan. *Personal identity : complex or simple?* Cambridge University Press, 2012.
- J.-P. Georgé, M.-P. Gleizes, and V. Camps. Cooperation. In *Self-organising Software*, pages 193–226. Springer, 2011.
- I. Georgievski and M. Aiello. Htn planning : Overview, comparison, and beyond. *Artificial Intelligence*, 222 :124–156, 2015.
- M. Ghallab, D. Nau, and P. Traverso. *Automated Planning : theory and practice*. Elsevier, 2004.
- D. Guinard, V. Trifa, F. Mattern, and E. Wilde. From the iot to the wot : Resource-oriented architecture and best practices. In *Architecting the Internet of Things*, pages 97–129. Springer, 2011.
- R. Gupta and G. Kansal. A survey on comparative study of mobile agent platforms. *International Journal of Engineering Science and Technology*, 3(3), 2011.
- K. V. Hindriks, F. S. De Boer, W. Van Der Hoek, and J.-J. C. Meyer. Agent programming with declarative goals. In *International Workshop on Agent Theories, Architectures, and Languages*, pages 228–243. Springer, 2000.
- J.-M. Hoc. Towards a cognitive approach to human–machine cooperation in dynamic situations. *International journal of human-computer studies*, 54(4) :509–540, 2001.
- M. A. Hogg, D. J. Terry, and K. M. White. A tale of two theories : A critical comparison of identity theory with social identity theory. *Social psychology quarterly*, pages 255–269, 1995.
- L. E. Holmquist, F. Mattern, B. Schiele, P. Alahuhta, M. Beigl, and H.-W. Gellersen. Smart-its friends : A technique for users to easily establish connections between smart artefacts. In *Ubicomp 2001 : Ubiquitous Computing*, pages 116–122. Springer, 2001.
- D. Hussein, S. Park, S. N. Han, and N. Crespi. Dynamic social structure of things : A contextual approach in cps. *Internet Computing, IEEE*, 19(3) :12–20, 2015.



- L. A. Jackson, L. A. Sullivan, R. Harnish, and C. N. Hodge. Achieving positive social identity : Social mobility, social creativity, and permeability of group boundaries. *Journal of Personality and Social Psychology*, 70(2) :241, 1996.
- W. James. The principles of psychology new york. *Holt and company*, 1890.
- J.-P. Jamont. *Démarche, modèles et outils multi-agents pour l'ingénierie des collectifs cyber-physiques*. PhD thesis, Université Grenoble Alpes, 2016.
- J.-P. Jamont and M. Ocelllo. Meeting the challenges of decentralised embedded applications using multi-agent systems. *International Journal of Agent-Oriented Software Engineering*, 5(1) :22–68, 2015.
- J.-P. Jamont, M. Ocelllo, and A. Lagrèze. A multiagent approach to manage communication in wireless instrumentation systems. *Measurement*, 43(4) :489–503, 2010.
- J.-P. Jamont, L. Médini, and M. Mrissa. A web-based agent-oriented approach to address heterogeneity in cooperative embedded systems. In *Trends in Practical Applications of Heterogeneous Multi-Agent Systems. The PAAMS Collection*, pages 45–52. Springer, 2014.
- H. Jung and M. Tambe. Conflicts in agent teams. In *Conflicting agents*, pages 153–167. Springer, 2002.
- B. Kerkez and M. T. Cox. Incremental case-based plan recognition with local predictions. *International Journal on Artificial Intelligence Tools*, 12(04) :413–463, 2003.
- R. Kerr and R. Cohen. Detecting and identifying coalitions. In *Proceedings of the 16th AAAI Conference on Plan, Activity, and Intent Recognition, AAAIWS'11-16*, pages 41–46. AAAI Press, 2011. URL <http://dl.acm.org/citation.cfm?id=2908772.2908778>.
- E. M. Khalfi, J.-P. Jamont, F. Cervantes, and M. Barhamgi. Designing the web of things as a society of autonomous real/virtual hybrid entities. In *Proceedings of the 2014 International Workshop on Web Intelligence and Smart Sensing*, pages 1–5. ACM, 2014.
- E.-M. Khalfi, J.-P. Jamont, M. Mrissa, and L. Médini. A restful task allocation mechanism for the web of things. In *International Conference on Computing & Communication Technologies (RIVF)*, pages 73–78, 2016.

- E.-M. Khalfi, J.-P. Jamont, and M. Occello. Les identités au centre de la participation des agents à des actions collectives. In *Journées Francophones sur les Systèmes Multi-Agents (JFSMA 2017)*, 2017.
- A. Khenifar-Bessadi, J.-P. Jamont, M. Occello, C.-B. Ben-Yelles, and M. Koudil. De la coopération de collectifs multi-agents. un modèle d'utilisation de productions collectives. *Revue des Sciences et Technologies de l'Information-Série RIA : Revue d'Intelligence Artificielle*, 31(1-2) :97–132, 2017.
- G. Knoblich, S. Butterfill, and N. Sebanz. Psychological research on joint action : theory and data. In *Psychology of learning and motivation*, volume 54, pages 59–101. Elsevier, 2011.
- K. Kravari and N. Bassiliades. A survey of agent platforms. *Journal of Artificial Societies and Social Simulation*, 18(1) :11, 2015.
- J. K. Kuchar and L. C. Yang. A review of conflict detection and resolution modeling methods. *IEEE Transactions on intelligent transportation systems*, 1(4) :179–189, 2000.
- D. Layder. *Social and personal identity : Understanding yourself*. Sage, 2004.
- D. Le Berre and A. Parrain. The sat4j library, release 2.2, system description. *Journal on Satisfiability, Boolean Modeling and Computation*, 7 :59–64, 2010.
- C. M. Li, F. Manyà, and J. Planes. New inference rules for max-sat. *J. Artif. Intell. Res.(JAIR)*, 30(1) :321–359, 2007.
- J. Li, G. A. Wang, and H. Chen. Identity matching using personal and social identity features. *Information Systems Frontiers*, 13(1) :101–113, 2011.
- T.-H. Liu and K. S. Barber. Selection of conflict resolution strategies in dynamically organized sensible agent-based systems. In *AAAI/IAAI*, page 1193, 1998.
- J. Locke. *An essay concerning human understanding*. T. Tegg and Son, 1836.
- D. Martinho, J. Carneiro, G. Marreiros, and P. Novais. Dealing with agents' behaviour in the decision-making process. In *Intelligent Environments (Workshops)*, pages 4–14, 2015.
- C. Maternaghan and K. J. Turner. Policy conflicts in home automation. *Computer Networks*, 57(12) :2429–2441, 2013.

- G. J. McCall and J. L. Simmons. Identities and interactions. 1966.
- G. H. Mead. Mind, self and society from the standpoint of a social behaviourist. *Chicago : University of Chicago Press*, 1934.
- L. Médini, M. Mrissa, E.-M. Khalfi, M. Terdjimi, N. Le Sommer, P. Capdepuy, J.-P. Jamont, M. Occello, and L. Touseau. Building a web of things with avatars : A comprehensive approach for concern management in wot applications. In *Managing the Web of Things*, pages 151–180. Elsevier, 2017.
- P. J. Modi, W.-M. Shen, M. Tambe, and M. Yokoo. Adopt : Asynchronous distributed constraint optimization with quality guarantees. *Artificial Intelligence*, 161(1-2) :149–180, 2005.
- M. Mrissa, L. Médini, J.-P. Jamont, N. Le Sommer, and J. Laplace. An avatar architecture for the web of things. *IEEE Internet Computing*, 19(2) :30–38, 2015.
- B. Nansen, L. van Ryn, F. Vetere, T. Robertson, M. Brereton, and P. Douish. An internet of social things. In *Proceedings of the 26th Australian Computer-Human Interaction Conference on Designing Futures : the Future of Design*, pages 87–96. ACM, 2014.
- D. Nau, Y. Cao, A. Lotem, and H. Munoz-Avila. Shop : Simple hierarchical ordered planner. In *Proceedings of the 16th international joint conference on Artificial intelligence-Volume 2*, pages 968–973. Morgan Kaufmann Publishers Inc., 1999.
- H. W. Noonan. *Personal identity*. Psychology Press, 2003.
- T. Okimoto, N. Schwind, M. Clement, T. Ribeiro, K. Inoue, and P. Marquis. How to form a task-oriented robust team. In *Proceedings of the 2015 International Conference on Autonomous Agents and Multiagent Systems*, pages 395–403. International Foundation for Autonomous Agents and Multiagent Systems, 2015.
- R. Olfati-Saber, J. A. Fax, and R. M. Murray. Consensus and cooperation in networked multi-agent systems. *Proceedings of the IEEE*, 95(1) :215–233, 2007.
- T. J. Owens. Self and identity. In *Handbook of social psychology*, pages 205–232. Springer, 2006.
- L. E. Parker. Alliance : An architecture for fault tolerant multirobot cooperation. *IEEE transactions on robotics and automation*, 14(2) :220–240, 1998.

- D. Pereira, E. Oliveira, N. Moreira, and L. Sarmento. Towards an architecture for emotional bdi agents. In *Artificial intelligence, 2005. epia 2005. portuguese conference on*, pages 40–46. IEEE, 2005.
- L. M. Pereira. State-of-the-art of intention recognition and its use in decision making. *AI Communications*, 26(2) :237–246, 2013.
- A. Petcu and B. Faltings. A scalable method for multiagent constraint optimization. In *Proceedings of the 19th International Joint Conference on Artificial Intelligence, IJCAI’05*, pages 266–271, San Francisco, CA, USA, 2005. Morgan Kaufmann Publishers Inc. URL <http://dl.acm.org/citation.cfm?id=1642293.1642336>.
- A. Pokahr, L. Braubach, and W. Lamersdorf. A goal deliberation strategy for bdi agent systems. In *German Conference on Multiagent System Technologies*, pages 82–93. Springer, 2005.
- T. Postmes, S. A. Haslam, and R. I. Swaab. Social influence in small groups : An interactive model of social identity formation. *European review of social psychology*, 16(1) :1–42, 2005.
- M. Ramirez and H. Geffner. Plan recognition as planning. In *Proceedings of the 21st international joint conference on Artificial intelligence. Morgan Kaufmann Publishers Inc*, pages 1778–1783, 2009.
- A. S. Rao. Agentspeak (I) : Bdi agents speak out in a logical computable language. In *European Workshop on Modelling Autonomous Agents in a Multi-Agent World*, pages 42–55. Springer, 1996.
- A. S. Rao and M. P. Georgeff. Modeling rational agents within a bdi-architecture. *KR*, 91 :473–484, 1991.
- P. Rashidi, D. J. Cook, L. B. Holder, and M. Schmitter-Edgecombe. Discovering activities to recognize and track in a smart environment. *Knowledge and Data Engineering, IEEE Transactions on*, 23(4) :527–539, 2011.
- S. Resendes, P. Carreira, and A. C. Santos. Conflict detection and resolution in home and building automation systems : a literature review. *Journal of Ambient Intelligence and Humanized Computing*, 5(5) :699–715, 2014.
- R. Rönquist. The goal oriented teams (gorite) framework. In *International Workshop on Programming Multi-Agent Systems*, pages 27–41. Springer, 2007.

- F. Sadri. Logic-based approaches to intention recognition. *Handbook of Research on Ambient Intelligence : Trends and Perspectives*, pages 346–375, 2010.
- J. S. Santos, J. O. Zahn, E. A. Silvestre, V. T. Silva, and W. W. Vasconcelos. Detection and resolution of normative conflicts in multi-agent systems : a literature survey. *Autonomous Agents and Multi-Agent Systems*, pages 1–47, 2017.
- S. Sardina and L. Padgham. A bdi agent programming language with failure handling, declarative goals, and planning. *Autonomous Agents and Multi-Agent Systems*, 23(1) : 18–70, 2011.
- S. Sardina, L. de Silva, and L. Padgham. Hierarchical planning in bdi agent programming languages : A formal approach. In *Proceedings of the Fifth International Joint Conference on Autonomous Agents and Multiagent Systems*, AAMAS '06, pages 1001–1008, New York, NY, USA, 2006. ACM. ISBN 1-59593-303-4. doi : 10.1145/1160633.1160813. URL <http://doi.acm.org/10.1145/1160633.1160813>.
- G. Seddone. *Collective Intentionality, Norms and Institutions. A Philosophical Investigation about Human Cooperation*. Peter Lang, 2014.
- C. Sedikides and M. J. Strube. Self-evaluation : To thine own self be good, to thine own self be sure, to thine own self be true, and to thine own self be better. In *Advances in experimental social psychology*, volume 29, pages 209–269. Elsevier, 1997.
- S. Shapiro, S. Sardina, J. Thangarajah, L. Cavedon, and L. Padgham. Revising conflicting intention sets in bdi agents. In *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 2*, pages 1081–1088. International Foundation for Autonomous Agents and Multiagent Systems, 2012.
- O. Shehory and S. Kraus. Methods for task allocation via agent coalition formation. *Artificial intelligence*, 101(1) :165–200, 1998.
- C. Shin and W. Woo. Service conflict management framework for multi-user inhabited smart home. *Journal of Universal Computer Science*, 15(12) :2330–2352, 2009.
- M. P. Sindlar, M. M. Dastani, F. Dignum, and J.-J. C. Meyer. Mental state abduction of bdi-based agents. In *Declarative Agent Languages and Technologies VI*, pages 161–178. Springer, 2008.
- R. G. Smith. The contract net protocol : High-level communication and control in a distributed problem solver. *IEEE Transactions on computers*, (12) :1104–1113, 1980.

- D. Snow. Collective identity and expressive forms. *Center for the Study of Democracy*, 2001.
- S. Soon, A. Pearce, and M. Noble. A teamwork coordination strategy using hierarchical role relationship matching. In *Agents and Computational Autonomy*, pages 249–260. Springer, 2003.
- N. Spanoudakis and P. Moraitis. Engineering ambient intelligence systems using agent technology. *IEEE Intelligent Systems*, 30(3) :60–67, 2015.
- D. N. Stern. *The present moment in psychotherapy and everyday life (Norton series on interpersonal neurobiology)*. WW Norton & Company, 2004.
- J. E. Stets and P. J. Burke. Identity theory and social identity theory. *Social psychology quarterly*, pages 224–237, 2000.
- P. Stone, G. A. Kaminka, S. Kraus, J. S. Rosenschein, et al. Ad hoc autonomous agent teams : Collaboration without pre-coordination. In *Proceedings of the Twenty-Fourth Conference on Artificial Intelligence*, July 2010.
- G. Strawson. *Locke on personal identity : Consciousness and concernment*. Princeton University Press, 2014.
- S. Stryker and R. T. Serpe. Commitment, identity salience, and role behavior : Theory and research example. In *Personality, roles, and social behavior*, pages 199–218. Springer, 1982.
- G. Sukthankar, C. Geib, H. H. Bui, D. Pynadath, and R. P. Goldman. *Plan, activity, and intent recognition : Theory and practice*. Newnes, 2014.
- K. P. Sycara. Machine learning for intelligent support of conflict resolution. *Decision Support Systems*, 10(2) :121–136, 1993.
- E. Syukur, S. W. Loke, and P. Stanski. Methods for policy conflict detection and resolution in pervasive computing environments. In *Policy Management for Web workshop in conjunction with WWW2005 Conference*, 2005.
- H. Tajfel. Social and cultural factors in perception. *Handbook of social psychology*, 3 : 315–394, 1969.
- H. Tajfel. Experiments in intergroup discrimination. *Scientific American*, 223(5) :96–103, 1970.

- H. E. Tajfel. *Differentiation between social groups : Studies in the social psychology of intergroup relations*. Academic Press, 1978.
- A. Tavakkoli, R. Kelley, C. King, M. Nicolescu, M. Nicolescu, and G. Bebis. A vision-based architecture for intent recognition. In *Advances in Visual Computing*, pages 173–182. Springer, 2007.
- M. Terdjimi, L. Médini, M. Mrissa, and M. Maleshkova. Multi-purpose adaptation in the web of things. In *International and Interdisciplinary Conference on Modeling and Using Context*, pages 213–226. Springer, 2017.
- C. Tessier, L. Chaudron, and H.-J. Müller. *Conflicting agents : conflict management in multi-agent systems*, volume 1. Springer Science & Business Media, 2006.
- J. Thangarajah, M. Winikoff, L. Padgham, and K. Fischer. Avoiding resource conflicts in intelligent agents. In *ECAI*, volume 2, pages 18–22, 2002.
- A. Torreño, E. Onaindia, A. Komenda, and M. Štolba. Cooperative multi-agent planning : A survey. *ACM Computing Surveys (CSUR)*, 50(6) :84, 2017.
- J. C. Turner. Social categorization and the self-concept : A social cognitive theory of group behavior. *Advances in group processes*, 2 :77–122, 1985.
- J. C. Turner. *Social influence*. Thomson Brooks/Cole Publishing Co, 1991.
- V. Tuttlies, G. Schiele, and C. Becker. Comity-conflict avoidance in pervasive computing environments. In *OTM Confederated International Conferences” On the Move to Meaningful Internet Systems”*, pages 763–772. Springer, 2007.
- M. B. van Riemsdijk, M. Dastani, and J.-J. C. Meyer. Goals in conflict : semantic foundations of goals in agent programming. *Autonomous Agents and Multi-Agent Systems*, 18(3) :471, 2009.
- W. W. Vasconcelos, M. J. Kollingbaum, and T. J. Norman. Normative conflict resolution in multi-agent systems. *Autonomous agents and multi-agent systems*, 19(2) :124–152, 2009.
- D. Weyns, H. V. D. Parunak, F. Michel, T. Holvoet, and J. Ferber. Environments for multi-agent systems state-of-the-art and research challenges. In *Environments for multi-agent systems*, pages 1–47. Springer, 2005.

- M. Winikoff. Implementing commitment-based interactions. In *Proceedings of the 6th international joint conference on Autonomous agents and multiagent systems*, page 128. ACM, 2007.
- M. Winikoff, L. Padgham, J. Harland, and J. Thangarajah. Declarative and procedural goals in intelligent agent systems. In *International Conference on Principles of Knowledge Representation and Reasoning*. Morgan Kaufman, 2002.
- X. Xu, L. Zhu, Y. Liu, and M. Staples. Resource-oriented architecture for business processes. In *Software Engineering Conference, 2008. APSEC'08. 15th Asia-Pacific*, pages 395–402. IEEE, 2008.
- M. Yagita, F. Ishikawa, and S. Honiden. An application conflict detection and resolution system for smart homes. In *Proceedings of the First International Workshop on Software Engineering for Smart Cyber-Physical Systems*, pages 33–39. IEEE Press, 2015.
- M. Yokoo, T. Ishida, E. H. Durfee, and K. Kuwabara. Distributed constraint satisfaction for formalizing distributed problem solving. In *1992 12th International Conference on Distributed Computing System*, pages 614–621. IEEE, 1992.
- M. R. Zatelli, J. F. Hübner, A. Ricci, and R. H. Bordini. Conflicting goals in agent-oriented programming. In *Proceedings of the 6th International Workshop on Programming Based on Actors, Agents, and Decentralized Control*, pages 21–30. ACM, 2016.
- N. Zhong, J. H. Ma, R. H. Huang, J. M. Liu, Y. Y. Yao, Y. X. Zhang, and J. H. Chen. Research challenges and perspectives on wisdom web of things (w2t). *The Journal of Supercomputing*, 64(3) :862–882, 2013.
- H. H. Zhuo and L. Li. Multi-agent plan recognition with partial team traces and plan libraries. In *IJCAI*, volume 22, page 484, 2011.
- H. H. Zhuo, Q. Yang, and S. Kambhampati. Action-model based multi-agent plan recognition. In *Advances in Neural Information Processing Systems*, pages 368–376, 2012.



## Publications

- Médini, Lionel, Michael Mrissa, El-Mehdi Khalfi, Mehdi Terdjimi, Nicolas Le Sommer, Philippe Capdepuy, Jean-Paul Jamont, Michel Occello, and Lionel Touseau. "Building a Web of Things with Avatars : A comprehensive approach for concern management in WoT applications." In *Managing the Web of Things*, pp. 151-180. Elsevier, Morgan Kaufmann, 2017.
- Khalfi, El-Mehdi, Jean-Paul Jamont, and Michel Occello. "Les identités au centre de la participation des agents à des actions collectives." In *Actes des Journées Francophones sur les Systèmes Multi-Agents (JFSMA 2017)*. 2017.
- Khalfi, El-Mehdi, Jean-Paul Jamont, Michael Mrissa, and Lionel Médini. "A RESTful task allocation mechanism for the Web of Things." In *Proceedings of the International Conference on Computing & Communication Technologies (IEEE-RIVF)*, pp. 73-78. 2016.  
*Cet article a été récompensé par le prix du meilleur papier étudiant.*
- Khalfi, El Mehdi, Jean-Paul Jamont, Francisco Cervantes, and Mahmoud Barhamgi. "Designing the web of things as a society of autonomous real/virtual hybrid entities." In *Proceedings of the International Workshop on Web Intelligence and Smart Sensing*, pp. 1-5. ACM, 2014.

