



HAL
open science

**Vers une vision intégrale des systèmes multi-agents :
Contribution à l'intégration des concepts d'agent,
d'environnement, d'organisation et d'institution.**

John Tranier

► **To cite this version:**

John Tranier. Vers une vision intégrale des systèmes multi-agents : Contribution à l'intégration des concepts d'agent, d'environnement, d'organisation et d'institution.. Informatique [cs]. Université Montpellier II - Sciences et Techniques du Languedoc, 2007. Français. NNT: . tel-00203489

HAL Id: tel-00203489

<https://theses.hal.science/tel-00203489>

Submitted on 10 Jan 2008

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.

Numéro d'identification :

ACADÉMIE DE MONTPELLIER

UNIVERSITÉ MONTPELLIER II
— SCIENCES ET TECHNIQUES DU LANGUEDOC —

THÈSE

présentée à l'Université des Sciences et Techniques du Languedoc
pour obtenir le diplôme de DOCTORAT

SPÉCIALITÉ : **Informatique**
Formation Doctorale : **Informatique**
École Doctorale : **Information, Structures, Systèmes**

Vers une vision intégrale des systèmes multi-agents

**Contribution à l'intégration des concepts d'agent,
d'environnement, d'organisation et d'institution**

par

John Tranier

Soutenue le 18 décembre 2007 devant le Jury composé de :

Olivier BOISSIER, Professeur, École Nationale Supérieure des Mines de Saint-Étienne Rapporteur
Amal EL FALLAH-SEGHROUCHNI, Professeur, Université Paris VI Rapporteur
Zohra BELLAHSÈNE, Professeur, Université Montpellier 2 Examineur
Fabien MICHEL, Maître de Conférences, Université de Reims Champagne-Ardenne Examineur
Tiberiu STRATULAT, Maître de Conférences, Université Montpellier 2 Examineur
Jacques FERBER, Professeur, Université Montpellier 2 Directeur

Numéro d'identification :

ACADÉMIE DE MONTPELLIER

UNIVERSITÉ MONTPELLIER II
— SCIENCES ET TECHNIQUES DU LANGUEDOC —

THÈSE

présentée à l'Université des Sciences et Techniques du Languedoc
pour obtenir le diplôme de DOCTORAT

SPÉCIALITÉ : **Informatique**
Formation Doctorale : **Informatique**
École Doctorale : **Information, Structures, Systèmes**

Vers une vision intégrale des systèmes multi-agents

**Contribution à l'intégration des concepts d'agent,
d'environnement, d'organisation et d'institution**

par

John Tranier

Soutenue le 18 décembre 2007 devant le Jury composé de :

Olivier BOISSIER, Professeur, École Nationale Supérieure des Mines de Saint-Étienne Rapporteur
Amal EL FALLAH-SEGHROUCHNI, Professeur, Université Paris VI Rapporteur
Zohra BELLAHSÈNE, Professeur, Université Montpellier 2 Examineur
Fabien MICHEL, Maître de Conférences, Université de Reims Champagne-Ardenne Examineur
Tiberiu STRATULAT, Maître de Conférences, Université Montpellier 2 Examineur
Jacques FERBER, Professeur, Université Montpellier 2 Directeur

Sommaire

Chapitre 1 Introduction	9
1.1 Contexte	9
1.2 Problématique	14
1.3 Proposition et contributions	14
1.4 Plan de la thèse	16
Chapitre 2 Les dimensions d'un système multi-agent : état de l'art	17
2.1 Agent	18
2.1.1 Agent réactif	18
2.1.2 Agent cognitif	19
2.1.3 Agent hybride	20
2.2 Environnement	21
2.2.1 La nécessité d'un environnement de première classe	22
2.2.2 Différents types d'environnement	24
2.2.3 Enjeux pour la conception d'un environnement	26
2.2.4 Le principe Influence/Réaction	28
2.2.5 La séparation entre le corps et l'esprit d'un agent	29
2.2.6 Modélisation explicite du dynamisme	30
2.2.7 Conclusion	31
2.3 Interaction	31
2.3.1 Stigmergie	32
2.3.2 Langage de communication agent	33
2.3.3 Protocole d'interaction	35
2.3.4 Conclusion	36
2.4 Dimension sociale	36
2.4.1 Organisation	40
2.4.2 Norme	45
2.4.3 Agent normatif	48

2.4.4	Engagement social	49
2.4.5	Contrôle et sanction	56
2.4.6	Institution	58
2.4.7	Searle : construction de la réalité sociale	60
2.5	Conclusion	67
Chapitre 3 Vers une vision intégrale des systèmes multi-agents		69
3.1	La théorie de la vision intégrale	70
3.2	Pourquoi une approche intégrale pour les systèmes multi-agents?	74
3.3	Une carte intégrale pour les systèmes multi-agents	75
3.4	Analyser au travers des 4 quadrants	81
3.4.1	La communication	81
3.4.2	Le contrôle social	82
3.5	Approche de conception pour les systèmes multi-agents	83
3.5.1	Séparer les quatre dimensions	83
3.5.2	Les relations entre les quatre dimensions	87
3.5.3	Quadrants – Comportements – Capacités	88
3.5.4	Intérêts de cette approche de conception	89
Chapitre 4 MASQ : un méta-modèle fondé sur les 4 quadrants		91
4.1	Introduction	92
4.1.1	Concepts primitifs	92
4.1.2	Principes	93
4.2	Esprit	95
4.3	Objet	97
4.4	Corps : incarnation d’un esprit dans un objet	100
4.5	Espace brut	105
4.5.1	Dimension interne d’un espace brut	106
4.5.2	Relations entre espaces bruts	114
4.5.2.1	Mécanisme d’interaction entre espaces bruts	114
4.5.2.2	Patrons de conception	116
4.5.3	Structuration de l’environnement	121
4.6	Culture	123
4.6.1	Interprétation collective	124
4.6.2	Perception d’une réalité institutionnelle	128
4.6.3	Internalisation des règles institutionnelles	131
4.6.4	Distinction entre culture et espace brut	133

4.6.5	Discussion	134
4.7	Synthèse	136
Chapitre 5 Modélisation et conception avec MASQ		139
5.1	Modélisation d'un wiki	139
5.1.1	Description du wiki	140
5.1.2	Modélisation avec MASQ	143
5.1.3	Synthèse	150
5.2	Le projet <i>Warbot</i>	151
5.2.1	Description	151
5.2.2	Modélisation MASQ de <i>Warbot</i>	154
5.2.3	Aspects implémentatoires	158
5.2.4	Conclusion	160
Chapitre 6 Conclusion et perspectives		161
6.1	Bilan	161
6.2	Perspectives	164
Table des figures		167
Liste des tableaux		169
Bibliographie		171

Chapitre 1

Introduction

Les caractéristiques et les attentes des applications informatiques ont considérablement changé ces dernières années, soulevant du même coup un nombre important de défis à relever. Les concepteurs d'applications doivent maintenant faire face à la *décentralisation*, à la *distribution* et au besoin de *fédérer* des systèmes hétérogènes. De plus, ils doivent être en mesure de fournir des solutions *robustes* et capables de *s'adapter* dans des environnements qui peuvent être aussi imprévisibles et versatiles que l'Internet.

Face à ces nouveaux enjeux, les techniques classiques ne parviennent qu'à proposer des réponses limitées.

1.1 Contexte

C'est dans ce contexte que les systèmes multi-agents (SMA) ont été proposés comme nouveau paradigme de conception. Ils se situent à l'intersection entre l'intelligence artificielle distribuée (IAD) et la vie artificielle. Ils amènent une décentralisation des données (connaissance) et du contrôle (comportement) : la résolution de problèmes ou la simulation de systèmes complexes ne résulte pas d'un calcul effectué par un programme monolithique, mais des interactions entre plusieurs entités autonomes.

D'un point de vue génie logiciel l'approche multi-agent constitue un changement radical de conception : on passe de la programmation *impérative* à la programmation *intentionnelle*. La programmation impérative consiste à décrire le « comment » en exposant pas à pas comment il faut effectuer la tâche à réaliser. Au contraire, la programmation intentionnelle vise à décrire le « quoi ». Ce qui se traduit en termes de programmation agent par la description de *buts* à atteindre pour un agent et de capacités de raisonnement pour déterminer comment atteindre ces buts.

Les SMA apportent de nouvelles solutions pour traiter des problèmes complexes, mais le prix à payer est une certaine perte de contrôle au niveau l'ensemble du système. Il n'est pas possible de donner pour ces systèmes de validation formelle. Souvent le résultat d'un SMA n'est pas prévisible, on ne peut que l'observer à son exécution. Mais justement les SMA sont destinés à opérer dans des contextes dynamiques, distribués, imprévisibles pour lesquels les systèmes prouvés échouent. Le rôle des SMA est de donner des moyens pour traiter des problèmes pour lesquels nous ne sommes pas en mesure de déterminer à l'avance ce que serait une bonne solution ou quel serait le moyen d'y parvenir.

Systeme multi-agent

En toute première approche nous pouvons décrire un SMA comme étant un ensemble d'entités actives et autonomes, les **agents**, qui évoluent dans un **environnement**.

Dans [Russell et Norvig, 2003], les auteurs décrivent un agent de la manière suivante : "An agent is anything that can be viewed as perceiving its environment through sensors and acting upon the environment through effectors". Cette relation générale entre un agent et son environnement est illustrée sur la figure 1.1.

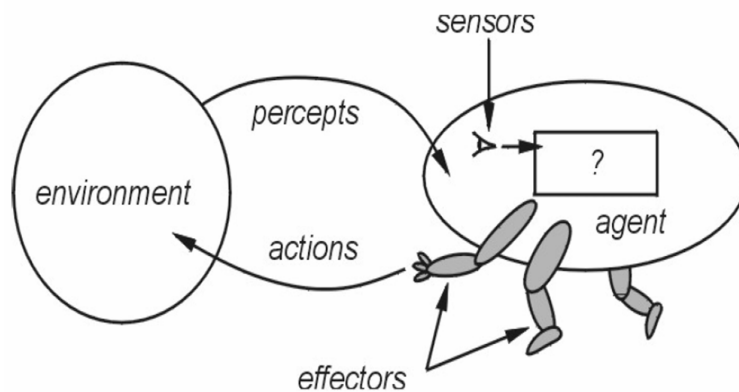


FIG. 1.1 – Interaction entre un agent et son environnement

D'après cette description, la notion d'agent regroupe aussi bien des êtres humains, des robots ou encore des agents virtuels (incarné par un programme informatique). L'environnement est le contexte pour les activités et les interactions des agents. Il peut être le monde réel (dans le cas de robots par exemple) ou bien un univers virtuel modélisé de manière informatique. Sans environnement il n'y a pas d'action ou de perception possible, on ne peut donc pas parler d'agent sans parler d'environnement [Weyns *et al.*, 2004a]. D'autre part il est important de préciser qu'un agent n'est capable de percevoir

que partiellement¹ son environnement.

Il n’y a pas de consensus absolu pour donner une définition exacte du terme « agent », il existe plutôt différentes définitions acceptées, chacune étant plus adaptée à un contexte particulier. Nous nous contenterons donc ici d’énumérer un ensemble de propriétés communément admises pour les agents.

Autonomie Un des fondements du paradigme multi-agent est de considérer un agent comme *autonome*. A partir du moment où un agent est exécuté sur un système, celui-ci à un contrôle absolu sur son comportement. Aucun élément extérieur, tel quel l’environnement ou un autre agent ne peut contrôler un agent. Cette autonomie constitue la différence principale entre un agent et un objet. Un objet est doté de méthodes qu’il suffit d’invoquer pour en obtenir l’exécution. Pour un agent, l’invocation de méthodes des objets est remplacée par la *négociation*.

Réactivité Un agent est *réactif*, c’est à dire qu’il maintient une interaction permanente avec son environnement et réagit aux changements qu’il perçoit dans celui-ci. Le comportement d’un agent peut s’adapter aux évolutions de son environnement.

Pro-activité Un agent est *pro-actif*, c’est-à-dire qu’il ne se contente pas de réagir à des évènements extérieurs, il est aussi capable de prendre l’initiative de son comportement. On attribue aux agents des buts qui les motivent (que ceux-ci soient explicites ou implicites), et le comportement qu’ils mettent en œuvre tend à satisfaire leurs buts.

Sociabilité Un agent est une entité sociale dans le sens où il est capable d’interagir avec les autres agents. Par essence, un agent vit en société, il doit donc prendre en compte les autres qui peuvent l’aider à atteindre ses objectifs ou au contraire le perturber dans cette tâche.

Ce paradigme qui consiste à modéliser un problème par un ensemble d’agents autonomes qui évoluent dans un environnement a constitué dans les années 80 le fondement du domaine des SMA. En l’état ce paradigme a l’avantage d’être simple sur le principe et donc facile à comprendre. Cependant il est si simple qu’il en est pauvre. Toutes les constructions nécessaires pour résoudre un problème donné doivent être conçues de manière *ad-hoc* par le modélisateur.

¹La perception limitée d’un agent sur son environnement est fondamentale, elle garantit que la distribution du système est bien réelle. Si le système pouvait être observé globalement par un unique agent, il ne serait plus distribué [Jonquet, 2006].

Évolution des systèmes multi-agents

Depuis, ce paradigme a été considérablement enrichi par l'introduction de nouveaux concepts et de nouvelles méthodologies dans le but de rendre ce paradigme réellement exploitable. Aujourd'hui, il est clair qu'un SMA ne se résume plus à la simple mise en commun d'agents autonomes au sein d'un environnement.

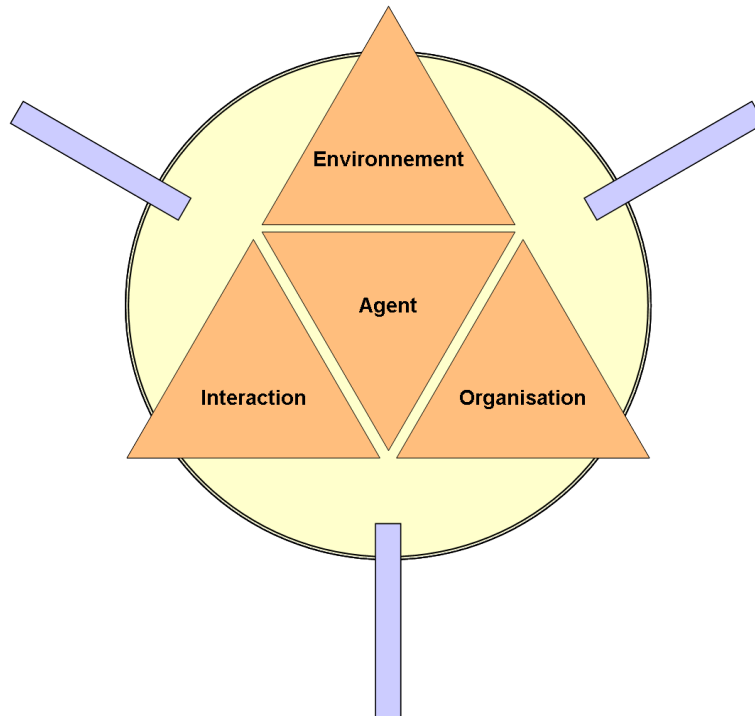


FIG. 1.2 – L'approche VOYELLES

Dans son approche VOYELLES [Demazeau, 2001] (figure 1.2), Demazeau distingue quatre dimensions orthogonales qui interviennent dans l'analyse, la conception, l'implémentation et l'exécution d'un système multi-agent : l'*Agent*, l'*Environnement*, l'*Interaction* et l'*Organisation* (A+E+I+O).

En ce qui concerne la dimension de l'organisation, nous préférons le terme de dimension sociale, car dans l'approche VOYELLES cette dimension regroupe aussi les normes, les lois ou les institutions.

Chacune de ces dimensions a constitué un axe de recherche spécifique pour les SMA. Nous donnons ici un rapide aperçu de chacun d'eux avant de les aborder plus en détail au chapitre 2.

Agent : c'est sur cette dimension que l'effort a été le plus important. Elle comprend les éléments qui sont nécessaires à la définition un agent autonome, c'est-à-dire son architecture interne, les mécanismes qu'il met en œuvre pour raisonner et produire son comportement, ses représentations de la connaissances, etc. Le résultat le plus significatif dans ce domaine est l'architecture BDI (*Belief, Desire, Intention*) [Rao et Georgeff, 1992].

Environnement : l'environnement, nous l'avons vu, est le milieu dans lequel les agents évoluent. C'est plus récemment que certains travaux se sont concentrés spécifiquement sur cet élément du système. Les problématiques principales qui sont soulevées par sa conception sont le traitement de l'action, de la perception, et de sa dynamique propre. La conférence *E4MAS (Environment for Multi-Agent Systems)* illustre l'intérêt qui porté à l'environnement par une partie de la communauté multi-agents.

Interaction : cette dimension comprend les moyens par lesquels les agents peuvent interagir. L'interaction a été étudié principalement pour faciliter la *coordination* et la *coopération* entre les agents. Les travaux dans cette voie ont notamment donné naissance aux *langages de communication* tels que FIPA-ACL [FIPA, 2001] et KQML [Finin et al., 1994], ainsi qu'aux *protocoles d'interaction*.

Dimension sociale : certains travaux ont mis en évidence l'intérêt de considérer la dimension sociale d'un SMA, d'une part pour structurer explicitement le système, et d'autre part pour réguler l'activité des agents tout en préservant leur autonomie. Ces travaux se sont largement inspirés de la sociologie, on parle de la « métaphore sociale ». Les principales notions qui ont été proposées dans cette direction sont les *organisation*, les *normes*, les *engagements sociaux* et les *institution*.

D'autre part, un aspect qui retient une attention toute particulière est la conception de SMA ouverts. Un SMA ouvert est un SMA dans lequel les agents sont *hétérogènes* et *égocentriques*, et sont développés par différents concepteurs, en utilisant des langages et des architectures différentes. La conception de tels systèmes pose une difficulté majeure : il n'est pas possible de faire de présuppositions sur l'architecture interne des agents.

1.2 Problématique

La problématique que nous proposons d'aborder dans cette thèse découle du constat suivant : les concepts fondamentaux d'un SMA que sont l'agent, l'environnement, l'interaction, et les concepts sociaux (organisation, norme, institution) ont été principalement étudiés séparément. Les travaux portant sur les modèles d'agent ont analysé la relation entre la structure interne d'un agent et son comportement, ainsi que l'interaction entre les agents ; mais ils n'ont généralement pas pris en compte l'environnement et les organisations. Les travaux qui se sont focalisés sur l'environnement ont peu ou pas considéré les problématiques sociales. Et réciproquement, les travaux sur les organisations et les concepts normatifs ont généralement négligé la notion d'environnement.

Cet état de fait a eu pour conséquence de donner naissance à différentes écoles : les approches centrées agent, les approches centrées organisation, et le domaine spécifique de la simulation qui met en avant l'environnement.

Pour chacun de ces courants de recherche, des besoins différents ainsi que des éléments de réponses ont été étudiés en ce qui concerne la conception de SMA ; mais les concepts utilisés sont souvent intriqués, ce qui rend difficile l'extraction de l'un d'entre eux dans une approche pour le mettre en relation avec des concepts utilisés dans des approches différentes. De plus, d'une approche à l'autre les mêmes termes n'ont pas toujours la même signification. Et de manière générale, les champs d'application des concepts sont mal définis : dans la conception d'un SMA, comment déterminer ce qui incombe aux agents ? à l'environnement ? aux organisations ? aux normes ou aux institutions ? Dans ce contexte, il est difficile de mettre en relation des modèles qui portent sur des aspects différents d'un SMA et par conséquent les méthodologies de conception existantes ne peuvent être que partielles. Il manque un cadre conceptuel commun qui permette de réconcilier les différentes approches, et ainsi de tirer profit de chacune d'elle dans la conception d'un SMA.

1.3 Proposition et contributions

Dans cette thèse nous mettons en avant l'importance de considérer un SMA suivant tous ses aspects : agent, environnement, interaction, organisation et institution. Ces différents aspects ayant principalement été étudiés séparément, nous proposons une approche fédératrice destinée à intégrer de manière cohérente agent, environnement, interaction, organisation et institution dans un même cadre conceptuel.

Dans cet objectif, nous proposons tout d'abord une restructuration conceptuelle du

domaine des SMA qui est directement inspirée des *quadrants* de la théorie de la « *vision intégrale* » qui est développée par [Wilber, 2001]. L'objectif de cette restructuration est de clarifier d'une part la fonction des différents aspects d'un SMA, et d'autre part les relations qu'ils entretiennent.

Ensuite, nous proposons le méta-modèle MASQ (*Multi-Agent System based on Quadrants*) qui permet de décrire un SMA en distinguant explicitement les aspects agent, environnement, organisation et institution.

Les contributions de ce travail de recherche se situent sur deux plans :

1. **sur le plan conceptuel**, le nouveau cadre conceptuel que nous proposons permet de :
 - clarifier les concepts fondamentaux en différenciant la fonction de chacun ; un résultat de cette clarification est la distinction entre structure organisationnelle et concepts normatifs (norme, engagement social, institution) ;
 - intégrer de manière cohérente les concepts d'agent, d'environnement, d'organisation et d'institution après avoir décrit explicitement leurs relations ;
 - unifier l'activité physique et l'activité sociale (ce cadre conceptuel permet de prendre en compte aussi bien des SMA basés sur la communication que des SMA utilisant des agents situés) ;
 - guider l'analyse pour la modélisation de problèmes avec une approche multi-agent ;
 - établir une « carte des SMA » sur laquelle il est possible de situer des modèles afin de mettre en évidence pour chacun d'eux les aspects qui sont traités et ceux qui ne le sont pas ; cette approche facilite l'intégration de modèles complémentaires ;
 - concevoir séparément chacune des dimensions proposée pour un SMA, ce qui est nécessaire pour la conception d'un SMA ouvert ;
2. **sur le plan formel**, nous proposons le méta-modèle MASQ (pour *Multi-Agent System based on Quadrants*) qui permet de décrire formellement un SMA suivant cette approche conceptuelle. MASQ repose sur quatre concepts primitifs : l'*esprit*, l'*objet*, l'*espace brut* et la *culture* ; un ensemble de relations entre ces concepts ; ainsi qu'un ensemble de lois pour décrire leur dynamique. Ce méta-modèle se veut suffisamment générique pour permettre l'intégration de modèles spécifiques selon le besoin du système considéré. L'objectif n'étant pas de déterminer quel modèle d'agent, d'environnement, d'organisation ou d'institution il faut utiliser, mais de fournir les moyens d'intégrer des modèles existants qui concernent ces différents concepts.

1.4 Plan de la thèse

La thèse sera organisée de la manière suivante :

Chapitre 2 Tout d’abord nous présentons un état de l’art orienté par les dimensions de l’approche VOYELLES : agent, environnement, interaction, et dimension sociale. Devant l’étendu du domaine et la disparité des travaux, nous nous efforcerons de présenter les principes essentiels qui se dégagent de chacune de ces thématiques.

Chapitre 3 Ce chapitre constitue le coeur de notre réflexion conceptuelle. Nous présenterons tout d’abord les *quadrants* de la théorie de la *vision intégrale* de Wilber et nous montrerons ensuite l’intérêt de son application aux SMA. Sur la base de cette théorie, nous proposerons un nouveau cadre conceptuel reposant sur quatre dimensions fondamentales pour analyser et concevoir un système multi-agent. Nous montrerons que ces dimensions fournissent un support qui permet de clarifier les concepts d’agent, d’environnement, d’interaction, d’organisation, de norme et d’institution, ainsi que les modèles qui sont proposés pour concevoir des systèmes multi-agents. Ensuite nous illustrerons l’intérêt d’utiliser cette théorie pour analyser des problèmes d’un point de vue multi-agent. Enfin nous proposerons une approche de conception adaptée aux systèmes ouverts qui découle de ce cadre conceptuel.

Chapitre 4 Ce chapitre est consacré à la présentation du méta-modèle MASQ (pour *Multi-Agent System based on Quadrants*) qui permet de décrire de manière formelle un SMA en suivant l’approche conceptuelle présentée au chapitre 3. MASQ repose sur quatre concepts primitifs : *l’esprit*, *l’objet*, *l’espace brut* et *la culture* ; un ensemble de relations entre ces concepts ; ainsi qu’un ensemble de lois pour décrire leur dynamique.

Chapitre 5 Ce chapitre illustre au travers de deux exemples la modélisation et la conception avec le méta-modèle MASQ. Le premier exemple traité est un système de type wiki qui supporte la collaboration entre agents humains et qui est de nature « sociale ». Le deuxième exemple, *Warbot* (la guerre des robots), fait intervenir des agents logiciels ce qui nous permet d’illustrer une opérationnalisation des aspects culturels.

Chapitre 6 Finalement nous dresserons les conclusions issues de ce travail de recherche et quelques-unes des perspectives qui lui sont associées.

Chapitre 2

Les dimensions d'un système multi-agent : état de l'art

Contents

2.1	Agent	18
2.2	Environnement	21
2.3	Interaction	31
2.4	Dimension sociale	36
2.5	Conclusion	67

Dans ce chapitre, nous présentons un état de l'art orienté suivant l'approche VOYELLES : agent, environnement, interaction, organisation.

2.1 Agent

La définition d'un agent est bien évidemment un point essentiel dans la conception d'un SMA, c'est donc tout naturellement l'aspect qui a retenu le plus d'attention dans les recherches portant sur ce domaine.

Dans cette section nous allons considérer l'architecture interne d'un agent, c'est-à-dire tous les mécanismes qui modélisent sa propre dynamique.

Un agent est un processus cyclique composé de trois phases : *perception* – *délibération* – *action* [Pnueli, 1986] (figure 2.1).

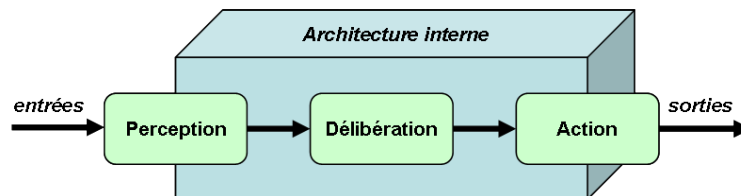


FIG. 2.1 – L'agent, un processus à 3 phases : *perception*, *délibération* puis *action*.

Le module de délibération est responsable du choix de l'action à exécuter, c'est le module le plus important dans la mesure où il définit les caractéristiques comportementales de l'agent. On classe [Wooldridge et Jennings, 1995] souvent les types d'agents en fonction de leur *comportement* individuel ; deux grandes catégories d'agents sont alors généralement distinguées : les agent *réactifs* et les agents *cognitifs*.

2.1.1 Agent réactif

Les agents réactifs sont caractérisés par des comportements simples qui se résument à des réactions aux stimuli de l'environnement. Ils ne possèdent pas de mémoire, ils n'ont pas de représentation de l'environnement, des autres agents qui sont présents dans le système, et d'eux-même ; leurs actions dépendent donc directement de leurs perceptions. Comme les agents réactifs ne sont pas capables de faire des traitements symboliques, leurs buts ne sont pas décrits explicitement mais sont « encodés » dans leur structure interne même et dans leur environnement.

Les approches à bases d'agents réactifs se fondent sur l'émergence de comportements collectifs sophistiqués à partir de comportements individuels simples. Elles défendent le principe qu'il n'est pas nécessaire que les agents soient individuellement intelligents pour

que le système dans son ensemble soit capable de résoudre des problèmes complexes [Ferber, 1999].

L'exemple le plus classique de système d'agents réactifs est celui des colonies de fourmis [Corbara *et al.*, 1993]. Alors que les fourmis ont un comportement très primitif et qu'il n'y a pas de relation d'autorité entre elles, leurs actions sont coordonnées de telle manière qu'il est possible pour la colonie de survivre et de résoudre des problèmes complexes tels que la recherche de nourriture ou la construction de nids.

La coordination entre agents réactifs repose sur des formes de communication indirecte au travers de l'environnement comme par exemple le dépôt de phéromone pour les fourmis.

Outre la résolution de problèmes complexes, les sociétés d'agents réactifs possèdent aussi la capacité de s'adapter dynamiquement à leur environnement. Par exemple, les fourmis sont capables de contourner un obstacle qui apparaît sur la route qu'elles empruntent entre la source de nourriture et la fourmilière. Il est possible de créer à partir d'agents réactifs des systèmes robustes et adaptatifs.

Les agents réactifs sont intéressants non pas au niveau individuel, mais au niveau des populations. La simplicité intrinsèque de leur comportement fait qu'ils ne présentent pas un grand intérêt individuellement, mais leur force vient du nombre et de leur capacité à faire émerger des organisations en formant des groupes.

2.1.2 Agent cognitif

Les agents cognitifs quant à eux possèdent une mémoire et sont capables de représentations symboliques. Ils possèdent une connaissance propre qui comprend une représentation (partielle) de leur environnement, des autres agents, ainsi que de leur savoir-faire, ce qui leur permet de gérer leurs interactions avec l'environnement et les autres agents.

On parle aussi d'*intentionnels*, *délibératifs* ou *rationnels* car ils possèdent des représentations explicites de leurs buts sur lesquelles ils sont capables de raisonner afin de produire des plans d'actions.

Ce qui distingue principalement les agents cognitifs des agents réactifs est leur faculté d'anticipation. Leur capacité à raisonner sur des représentations du monde leur permet de mémoriser des situations, de les analyser, de prévoir les changements induits par leurs actions et finalement de décider du comportement à adopter dans le futur. Un point crucial dans la conception d'agents cognitifs est donc la planification de leurs actions. Une revue détaillée des approches de planification d'action est proposée dans [El Fallah Seghrouchni, 2001; Balbo *et al.*, 2002].

```
Algorithme : BDI-interpreter
B := B_init;
I := I_init;
while true do
  get_perception(perc);
  B := belief_revision(B,perc);
  D := options(B,I);
  I := filter(B,D,I);
  plan = find_plan(B,I);
  execute(plan);
end
```

Algorithme 1: L'interpréteur BDI

L'architecture d'agents cognitifs la plus populaire est l'architecture BDI (Believes, Desires, Intentions) [Rao et Georgeff, 1992]. L'algorithme 1 donne le cycle de traitement d'un agent BDI qui détermine les actions à exécuter en prenant en compte ses perceptions. La première étape du cycle est la perception (fonction `get_perception`) qui est ensuite exploitée pour mettre à jour les croyances de l'agent (fonction `belief_revision`). La fonction `options` permet de générer un ensemble d'alternatives possibles (D) sur la base des croyances de l'agent (B) et de ses intentions actuelles (I). L'agent choisit ensuite une option parmi celles possibles (fonction `filter`). Le résultat est la nouvelle intention courante (I). Enfin un plan d'actions est généré (fonction `find_plan`) puis exécuté (`execute`).

Cette première version de l'algorithme donne une idée générale de l'approche BDI mais elle souffre d'un défaut : le plan est exécuté intégralement, ce qui ne permet pas de prendre en compte des modifications de l'environnement durant l'exécution du plan. Un agent pourrait donc s'engager dans une intention que l'évolution de l'environnement a rendu inconsistente.

Par la suite, [Rao et Georgeff, 1995] ont proposé un nouvel algorithme qui corrige ce problème (algorithme 2).

2.1.3 Agent hybride

Nous venons de présenter la distinction entre les agents de type réactif, qui réagissent uniquement aux stimuli de l'environnement, et les agents de type cognitif, qui possèdent une représentation symbolique de l'environnement qu'ils mettent à jour continuellement et qui leur permet de planifier leurs actions. Il faut cependant noter qu'il existe une multitude de possibilités entre ces deux extrêmes avec par exemple des agents qui exploitent des représentations numériques ou des représentations non symboliques.

```

Algorithme : BDI-interpreter
B := B_init;
I := I_init;
while true do
  get_perception(perc);
  B := belief_revision(B,perc);
  D := options(B,I);
  I := filter(B,D,I);
  plan = find_plan(B,I);
  while NOT (empty(plan) OR succeeded(I,B) OR impossible(I,B)) do
    action = next_action(plan);
    execute(action);
    get_perception(perc);
    B := belief_revision(B,perc);
    if reconsider(I,B) then
      D := options(B,I);
      I := filter(B,D,I);
    end
    if not(sound(plan,I,B)) then
      plan = find_plan(B,I);
    end
  end
end

```

Algorithme 2: L'interpréteur BDI modifié

Il est de plus possible de combiner les deux approches pour obtenir une architecture *hybride*. Dans une telle architecture un agent est composé de modules qui gèrent indépendamment la partie réflexe (réactive) et la partie réfléchie (cognitive) du comportement de l'agent. TouringMachines [Ferguson, 1992] et InteRRap [Müller, 1996] sont des exemples d'architectures hybrides.

2.2 Environnement

L'environnement d'un SMA est le contexte dans lequel les agents vont évoluer. Il fournit un support commun aux actions des agents, permettant ainsi l'interaction dans le système, et il constitue une source d'information à laquelle les agents peuvent accéder au travers de leur perception. Dans le cas de d'agents réels (comme des robots par exemple), l'environnement est le milieu dans lequel les agents sont plongés, ce milieu étant gouverné par les lois physiques naturelles. Dans le cas d'agents logiciels, un environnement artificiel possédant ses propres lois d'évolution doit être défini. Il est généralement considéré qu'un environnement est composé d'entités actives (les agents), d'entités passives (des objets) et de ressources qui peuvent être produites et consommées. Sans environnement il n'y a

pas d'action ou de perception possible, on ne peut donc pas parler d'agent sans parler d'environnement.

Pour un SMA informatique la notion d'environnement est ambiguë car on peut distinguer plusieurs niveaux d'environnements [Weyns *et al.*, 2004a]. Tout agent informatique nécessite un support matériel pour s'exécuter (processeur, réseau, etc.) ainsi qu'un support logiciel (système d'exploitation, machine virtuelle, etc.). L'infrastructure matérielle et logicielle constitue chacune un environnement qui détermine l'existence et les possibilités d'un agent. Enfin un SMA doit comporter un environnement conceptuel ou logique qui correspond à la modélisation du problème traité par le système. Les agents vont évoluer dans cet environnement pour résoudre le problème pour lequel ils sont conçus.

Pour distinguer le niveau infrastructure du niveau conceptuel, [Mertens *et al.*, 2004] parlent d'*environnement d'exécution* et d'*environnement d'application*.

Pour réduire cette ambiguïté inérente à l'environnement un modèle d'architecture 3-tier (figure 2.2) a été proposé dans [Weyns *et al.*, 2004a].

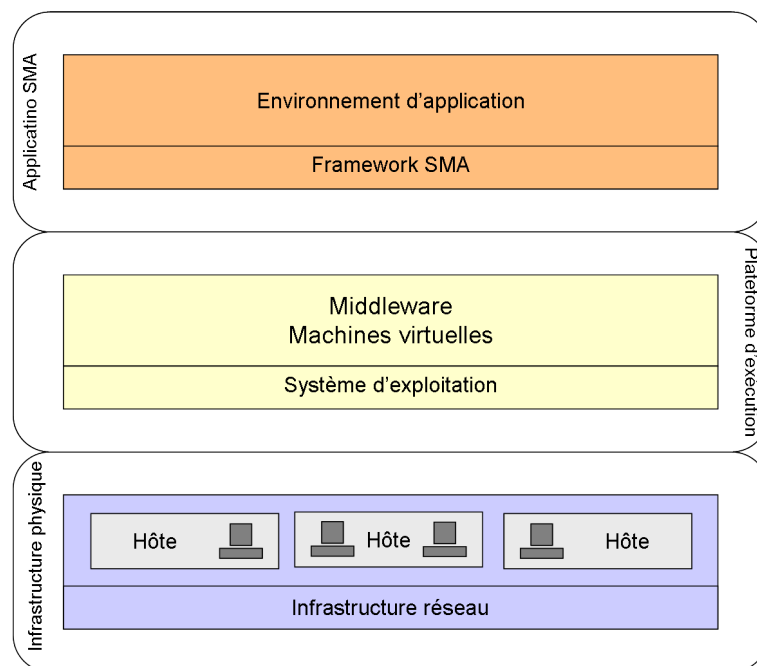


FIG. 2.2 – Modèle 3-tier pour les SMA

Dans cette thèse nous nous restreignons au niveau conceptuel d'un SMA, c'est donc l'environnement d'application que nous considérons.

2.2.1 La nécessité d'un environnement de première classe

Traditionnellement dans le domaine des SMA, la conception des agents a retenu beaucoup plus d'attention que celle de l'environnement, exception faite pour les systèmes

qui modélisent une composante spatiale dans laquelle les agents sont situés. Un des objectifs de cette thèse est de montrer qu'un environnement de première classe est nécessaire pour tout type de SMA.

La définition d'un environnement de première classe permet de représenter explicitement le problème à résoudre pour les agents, de définir la perception qu'ils peuvent en avoir, et les opérations, c'est-à-dire les actions, qu'ils peuvent effectuer sur cette représentation. Durant l'exécution du système, l'état de l'environnement décrit une réalité objective du point de vue des agents.

Pour un SMA qui modélise une composante spatiale, la nécessité d'un environnement explicite est évidente. Considérons par exemple une simulation de football multi-agents. Le système doit représenter l'aire de jeu, les joueurs (les agents) et le ballon. Les joueurs et le ballon doivent pouvoir être positionnés dans l'aire de jeu. Le système doit décrire les capacités de perception et d'action des joueurs ainsi que les lois physiques de cet univers virtuel. Tout cela nécessite donc la définition d'un environnement spécifique qui sera le contexte d'activités des agents.

Dans le d'un SMA qui ne comporte pas de composante spatiale (les agents n'étant alors pas situés), la notion d'environnement a souvent été négligée. En l'absence d'environnement d'application, l'environnement d'un agent se réduit à l'infrastructure (logicielle et matérielle) sous-jacente; c'est donc cette infrastructure qui détermine les capacités de perception et d'action d'un agent. Si l'on considère par exemple un agent s'exécutant sur une machine connectée à un réseau, cet agent aura alors comme capacité d'émettre et de recevoir des messages sur le réseau. Cependant, la plupart des applications nécessitent la définition de leurs propres contraintes qui sont différentes de celles définies par l'infrastructure. Pour la communication par exemple, on peut vouloir définir qui peut communiquer avec qui, sous quelle forme, dans quelles circonstances, ou encore qui peut intercepter les messages.

Finalement il nous apparaît qu'un environnement de première classe doit être défini pour tout SMA dont l'objet n'est pas un environnement existant. Une des rares situations où l'infrastructure réseau constitue réellement un environnement pour un SMA est quand le SMA traite un problème qui concerne cette infrastructure elle-même (comme par exemple la gestion du trafic sur un réseau).

2.2.2 Différents types d'environnement

Décrire un environnement de manière générale n'est pas une chose aisée car la nature de l'environnement d'un système dépend fortement de la nature du problème pour lequel le système est conçu. Dans [Weyns *et al.*, 2004a], un état de l'art sur l'environnement est présenté qui illustre l'importance de ce composant dans un SMA, mais aussi la difficulté qu'il y a à capturer cette notion et l'absence de consensus dans la communauté. En nous appuyant sur cette synthèse, nous rapportons ici différentes caractérisations qui sont proposées pour un environnement.

Dans [Russell et Norvig, 2003], les auteurs ont présenté un programme générique d'environnement (algorithme 3). Ce programme est trop générique pour capturer toutes les spécificités d'un environnement mais il illustre bien la relation entre les agents et leur environnement. La réalité maintenue par l'environnement est matérialisée par un état (la variable *state*) et l'environnement est décrit comme un processus cyclique en trois phases : (1) les perceptions sont données aux agents, ces perceptions sont construites par la fonction GET-PERCEPT et dépendent de l'état de l'environnement et de l'agent lui-même, (2) les actions des agents sont collectées, un agent est ici représenté par un programme (PROGRAM[AGENT]) qui prend en paramètre des perceptions et retourne des actions, (3) l'état de l'environnement est mis à jour par la fonction UPDATE-FN qui prend en paramètre les actions des agents, les agents et l'état courant de l'environnement.

Algorithme : RUN-ENVIRONMENT(*state*, UPDATE-FN, *agents*, *termination*)

Input : *state*, l'état initial de l'environnement

UPDATE-FN, la fonction de mise-à-jour de l'environnement

agents, un ensemble d'agents

termination, un prédicat pour tester la fin du programme

repeat

foreach *agent* **in** *agents* **do**

 PERCEPT[agent] ← GET-PERCEPT(agent, *state*);

end

foreach *agent* **in** *agents* **do**

 ACTION[agent] ← PROGRAM[agent](PERCEPT[agent]);

end

state ← UPDATE-FN(*actions*, *agents*, *state*);

until *termination*(*state*);

Algorithme 3: Un programme générique de l'environnement

Russell et Norvig distinguent pour un environnement les propriétés suivantes :

- *accessible* versus *inaccessible* : un environnement est dit accessible si les agents ont accès à l'intégralité de son état. L'accessibilité de l'environnement n'est pas une propriété souhaitable pour un SMA où la portée des actions et de la perception est locale. Nous nous intéressons donc aux environnement inaccessible.
- *déterministe* versus *indéterministe* : un environnement est déterministe si un changement de son état est déterminé uniquement par son état courant et les actions des agents. Un environnement indéterministe pourra produire un résultat différent pour une même action.
- *statique* versus *dynamique* : un environnement est dit statique si il ne change d'état que sous l'effet des actions des agents. Au contraire, un environnement dynamique possède ses propres processus d'évolution qui peuvent modifier son état sans intervention des agents. Un environnement dynamique permet de représenter l'activité manifestée par un objet (comme un ballon qui roule par exemple).
- *discret* versus *continu* : on parle d'environnement discret lorsque le nombre de perceptions et d'actions possibles est limité. Si ce n'est pas le cas, l'environnement est alors continu.

Les environnements les plus intéressants pour les SMA sont ceux qui sont inaccessibles, indéterministes et dynamiques.

[Ferber, 1999] considère qu'un environnement peut être représenté comme un système monolithique, c'est-à-dire un environnement centralisé, ou comme un réseau de cellules, c'est-à-dire un environnement distribué. Dans un environnement centralisé tous les agents ont accès à la même structure (mais la perception qu'ils en ont est locale et partielle). Dans un environnement distribué, chaque cellule se comporte comme un environnement centralisé en miniature. Cependant les cellules d'un environnement distribué se distinguent d'un environnement centralisé par les points suivants : (1) l'état d'une cellule dépend de l'état des cellules voisines, (2) la perception d'un agent n'est pas réduit à une unique cellule, (3) les agents peuvent se déplacer de cellule en cellule, il faut donc gérer le lien entre un agent et une cellule et (4) il faut gérer la propagation de signaux entre les cellules.

Les auteurs de [Odell *et al.*, 2002] font la distinction entre l'*environnement physique* et l'*environnement de communication*. L'environnement physique fournit des lois, des règles, des contraintes et des stratégies qui gouvernent et supportent l'existence physique des objets et des agents. L'environnement de communication fournit les principes et les processus qui gouvernent et supportent l'échange d'idées, de connaissance et d'information, et les fonctions et les structures qui sont communément employées pour la communication tels que les rôles, les groupes et les protocoles d'interaction entre les rôles et les groupes.

2.2.3 Enjeux pour la conception d'un environnement

Les auteurs de [Weyns *et al.*, 2004a] classent les enjeux pour la conceptions d'un environnement en deux catégories : les aspects structurels (structuration, ontologie) et les aspects dynamiques (perception, action, processus environnementaux).

Structuration

Un des rôles fondamental de l'environnement est de structurer l'ensemble des entités d'un SMA (les objets et les agents). Pour cela il doit définir les règles qui permettent d'établir et de faire évoluer des relations entre les entités. La structuration définie par un environnement peut être de nature différente. Les figures 2.3(a,b,c) illustrent trois structurations de nature spatiale : (a) modélise une maison qui définit une topologie particulière, (b) est une grille régulière qui permet de localiser les entités dans des cases et (c) modélise un réseau, les entités étant situées sur des nœuds du réseau. La structuration peut aussi être de nature organisationnelle, nous aborderons cet aspect à la section 2.4.1 qui est dédiée aux organisations.

La structure de l'environnement a bien évidemment un influence sur la perception que les agents peuvent en avoir et sur les actions qu'ils peuvent effectuer.

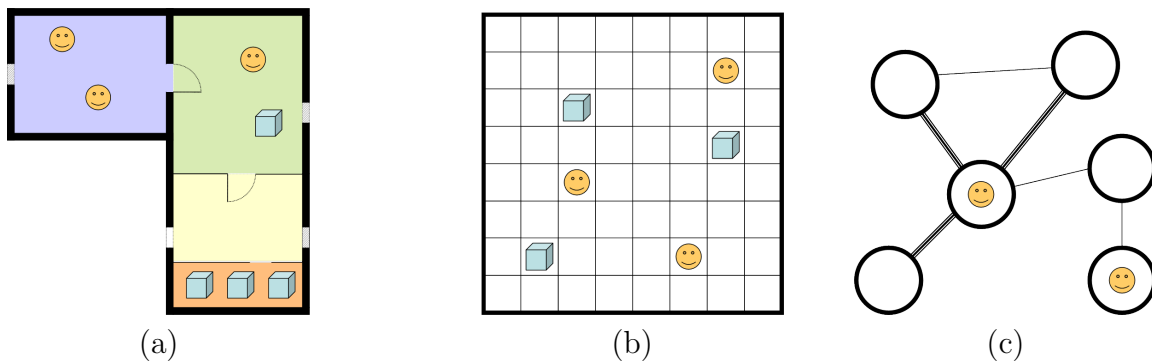


FIG. 2.3 – Structuration physique de l'environnement

Ontologie

Dans [Chang *et al.*, 2004], Chang et ses collègues déclarent : “*agents must be able to understand their environment*”. Pour cela, l'environnement doit définir une ontologie qui fournit une représentation conceptuelle du domaine. Cette ontologie doit permettre de décrire la structure de l'environnement, les propriétés observables des objets, des ressources et des agents et leurs relations.

Une ontologie explicite permet à des agents cognitifs d'interpréter et de raisonner sur l'environnement. Pour des agents réactifs, l'ontologie de l'environnement doit être

encodée dans leur structure interne.

Perception

Dans [Weyns *et al.*, 2004a] la perception est définie comme la capacité d'un agent à observer son voisinage, ce qui se traduit par la construction de percepts de l'environnement qui sont transmis aux agents. Un percept décrit l'environnement sous une forme compréhensible pour les agents. Les agents utilisent les percepts pour mettre à jour leur connaissance du monde ou bien ils les utilisent directement pour prendre leurs décisions. Dans le cas d'agents réels la perception est possible par le biais de périphériques matériels tel qu'une caméra vidéo ou un capteur laser. Dans le cas d'agents artificiel la perception doit être implémentée par une solution logicielle. C'est le rôle de l'environnement de déterminer la perception que peut avoir un agent suivant les propriétés de l'environnement et les capacités de l'agent. La perception peut être soumise à des *lois de perception* [Weyns *et al.*, 2004b]. Une loi de perception permettant d'exprimer, par exemple, qu'un objet derrière un obstacle n'est pas perceptible ou encore d'introduire du bruit dans la perception. De plus, la perception peut être active [Weyns *et al.*, 2004b], dans ce cas elle peut aussi être orientée par l'agent qui décide de se focaliser sur tel ou tel type d'information.

Action

Le traitement de l'action est une responsabilité de l'environnement. C'est l'environnement qui doit établir pour les agents quelles actions ils peuvent effectuer, et quel est l'effet de ces actions. Cet aspect est parfois traité de manière implicite en encodant les actions dans la structure interne des agents. Cette solution pose un problème de conception de manière générale car il n'est pas évident de garantir que les actions définies au niveau local de l'agent respecteront bien les contraintes que l'on souhaite avoir sur l'environnement. De plus cette solution n'est pas exploitable dans le cadre de systèmes ouverts. Nous présenterons pas la suite le principe *influence/réaction* [Ferber *et Müller*, 1996] qui a été proposé pour traiter l'action dans un SMA.

Processus environnementaux

En parallèle de l'activité engendrée par les agents, les objets et les ressources peuvent eux aussi produire de l'activité dans l'environnement. On peut citer par exemple un ballon qui roule ou encore la température de l'environnement qui varie. L'environnement doit donc être en mesure de maintenir ces différentes dynamiques.

2.2.4 Le principe Influence/Réaction

Le traitement de l'action est un aspect fondamental qui incombe à l'environnement. Dans la théorie classique de l'action, une action est considérée comme une transformation d'état de l'environnement. Un agent agit en modifiant directement l'état de son environnement. Alors que cette approche est satisfaisante dans le cadre de l'intelligence artificielle classique où un seul agent agit, elle ne permet pas de traiter correctement l'action dans le cadre d'un SMA, et cela pour 2 raisons :

1. Un agent n'est pas seul à agir. Les activités des autres agents peuvent interférer avec la sienne ; le résultat d'une action ne peut donc pas être déterminé sans considérer l'ensemble des actions qui ont été réalisées au même instant. C'est le problème de la simultanéité de l'action.
2. Un environnement est régi par des lois qui déterminent son évolution. Un agent n'a pas nécessairement connaissance de toutes ces lois, il n'est donc pas en mesure de prévoir toutes les conséquences de ses actes. Et quand bien même un agent aurait connaissance de ses lois, il pourrait décider ne pas les respecter. L'environnement doit donc assurer que l'évolution de son état est conforme aux lois qui ont été pré-établies à sa conception.

Le principe *influence/réaction* a été proposé par [Ferber et Müller, 1996] pour surmonter les limitations de la théorie classique de l'action et rendre compte de l'évolution de l'environnement en fonction de l'activité des agents. Ce principe stipule que les agents ne peuvent agir directement sur l'environnement mais peuvent seulement l'influencer. L'action directe est remplacée par la transmission d'influences ; les influences exprimant la volonté d'un agent de voir l'état de l'environnement évoluer d'une certaine manière. Ensuite l'environnement transforme lui-même son état en appliquant ce que Ferber et Müller appellent les *lois de l'univers*, ces lois prenant en compte les influences produites par les agents. Cette phase durant laquelle l'environnement met à jour son état est appelée *phase de réaction*. La phase de réaction engendre aussi l'émission de percepts à destination des agents. Ainsi les agents ne peuvent que tenter de faire évoluer l'environnement d'une certaine manière puis ensuite constater comment celui-ci a évolué.

Ce principe permet de traiter la simultanéité de l'action car les influences de tous les agents sont considérées pour déterminer la réaction de l'environnement. Il permet aussi de préserver l'intégrité de l'environnement [Michel, 2004] en assurant que l'évolution de l'environnement est bien conforme aux lois de l'univers. Ce dernier aspect est essentiel : il permet d'une part de délimiter l'autonomie des agents en établissant ce qu'ils peuvent faire, et d'autre part de favoriser la sécurité du système car aucun agent ne peut contourner les lois établies par l'environnement.

2.2.5 La séparation entre le corps et l'esprit d'un agent

Magnin a mis en évidence dans sa thèse [Magnin, 1996] la confusion inérante au manque de différenciation entre la composante environnementale d'un agent et son système décisionnel. On retrouve cette idée de séparation explicite entre ces deux composantes d'un agent dans les travaux de Soulié [Soulié, 2001] puis dans la thèse de Michel [Michel, 2004] avec la notion de séparation du corps et de l'esprit d'un agent.

L'esprit d'un agent fait référence à sa structure interne, c'est-à-dire la mémoire et les mécanismes dont dispose l'agent pour raisonner afin de produire son comportement. La structure interne d'un agent lui est propre, celle-ci doit être inaccessible pour tout élément extérieur à l'agent, que ce soit l'environnement ou les autres agents. L'autonomie d'un agent dépend de l'intégrité de son esprit.

Le corps d'un agent fait référence à l'existence de l'agent dans l'environnement. Il définit l'état qui caractérise l'agent dans l'environnement et les capacités d'action et de perception que possède l'agent dans l'environnement. Le corps est donc la manifestation d'un agent qui pourra être perçue par les autres agents. Alors que la conception d'un esprit relève de celle d'un agent, la conception d'un corps relève de celle de l'environnement. D'une part c'est l'environnement qui attribue à un agent ses capacités, d'autre part l'agent ne doit pas être en mesure de contrôler directement les propriétés qui le caractérisent dans l'environnement. Si l'on considère un environnement dans lequel les agents possèdent un niveau d'énergie, un agent ne doit pas avoir la possibilité de modifier directement ce niveau d'énergie. Et si l'on considère un environnement dans lequel les agents peuvent se déplacer, c'est l'environnement qui doit établir les contraintes de déplacement telle que la vitesse de déplacement (qui peut par exemple dépendre d'un niveau d'énergie).

En reliant ce principe avec le principe influence/réaction, on peut voir un agent comme un esprit relié à un corps, un corps étant un objet sous le contrôle de l'environnement. L'esprit a la possibilité d'influencer son corps et le corps transmet à l'esprit des percepts suivant ses capacités de perception. Ces deux principes permettent la conception séparée du corps et de l'esprit d'un agent.

La distinction entre le corps et l'esprit d'un agent amène différents avantages. D'un point de vue génie logiciel, elle permet une implémentation plus modulaire d'un SMA. Du point de vue du paradigme agent, elle permet de préserver à la fois l'intégrité des agents et celle de l'environnement. Enfin elle permet l'intégration naturelle dans un SMA de robots réels et d'agents humains. Dans le cas d'un robot le corps est réel donc seul l'esprit doit

être conçu. Au contraire dans le cas d'un agent humain, la fonction de l'esprit est remplie par l'utilisateur, la conception du corps permet alors de représenter l'utilisateur et de lui attribuer ses capacités dans l'environnement.

2.2.6 Modélisation explicite du dynamisme

[Helleboogh *et al.*, 2007] proposent un formalisme fondé sur le principe *influence/réaction* et sur la séparation du corps et de l'esprit d'un agent pour modéliser un environnement dynamique dans lequel le dynamisme est décrit de manière explicite.

Le dynamisme dans un environnement provient des agents qui agissent de manière autonome ainsi que des objets qui ont leur propre évolution. Le formalisme qui est proposé permet de décrire (1) le dynamisme de chaque entité de l'environnement, (2) comment les agents peuvent manipuler les dynamismes et (3) comment ces dynamismes peuvent interférer.

Pour capturer le dynamisme d'une entité de l'environnement, Helleboogh et al. introduisent la notion d'*activité*. Une activité décrit pour une entité comment son état évolue au cours du temps en l'absence de perturbations extérieures. Une activité définit une stratégie d'évolution qui permet de déterminer l'opération qui doit être appliquée à l'entité pour produire son nouvel état.

Pour décrire comment les agents peuvent intervenir sur l'environnement, Helleboogh et al. s'appuient sur le principe *influence/réaction* et la séparation esprit/corps. Un agent a la possibilité d'envoyer des influences à son corps qui est une entité de l'environnement. L'effet d'une influence est décrit par des *lois de réaction*. Une loi de réaction exprime la transformation qui appliquée aux activités du corps d'un agent en fonction des influences qui sont émises.

Enfin pour décrire les interférences que peuvent produire les différentes activités individuelles, Helleboogh et al. introduisent les *lois d'interférence*. Une loi d'interférence est la donnée d'une condition d'interférence qui décrit dans quelles circonstances une interférence donnée survient, et d'une transformation d'activité. Alors que les lois de réaction produisent des modifications d'activité pour une unique entité, les lois d'interférence transforment les activités de différentes entités ce qui produit l'interaction.

2.2.7 Conclusion

Dans cette section nous avons souligné l'importance de l'environnement en nous appuyant sur les travaux qui se sont focalisés sur cet élément d'un SMA. Nous avons mis en évidence avec le principe influence/réaction et le principe de séparation esprit/corps que le contexte d'interaction qui est défini par l'environnement doit être géré par celui-ci et non par les agents eux-mêmes. Nous avons par ailleurs motivé le fait qu'un environnement de première classe doit être défini pour tout type de SMA, y compris ceux pour lesquels les agents ne sont pas situés. Un environnement de première classe permet d'une part de délimiter l'autonomie des agents en établissant ce qu'ils peuvent faire et ce qu'ils ne peuvent pas faire, et d'autre part de garantir une perspective publique de l'action qui est nécessaire pour les approches sociales que nous présenterons à la section 2.4.1.

2.3 Interaction

Le simple fait de placer un ensemble d'agents dans un même environnement n'est pas suffisant pour définir un SMA. Les différents agents doivent être en mesure d'interagir et de se comprendre mutuellement afin de pouvoir se coordonner et éventuellement coopérer. L'étude des mécanismes d'interaction est donc primordiale dans la conception d'un SMA.

Pour citer Ferber :

“Pour un agent, interagir avec un autre constitue à la fois la source de sa puissance et l'origine de ses problèmes. C'est en effet parce qu'ils coopèrent que des agents peuvent accomplir plus que la somme de leurs actions, mais c'est aussi à cause de leur multitude qu'ils doivent coordonner leurs actions et résoudre des conflits.” [Ferber, 1995, p. 55]

Morin a donné la définition suivante de l'interaction :

“ [...] Les interactions sont des actions réciproques modifiant le comportement ou la nature des éléments, corps, objets, phénomènes en présence ou en influence.” [Morin, 1977, p. 51].

L'interaction entre les agents peut apparaître sous différentes formes comme la *collaboration* quand il s'agit de répartir des tâches entre un ensemble d'agents, la *coordination* quand il s'agit d'organiser les actions de différents agents pour atteindre un but collectif, ou encore la *négociation* quand il s'agit de résoudre des conflits entre les

agents.

On distingue généralement deux modes d'interaction : le mode direct et le mode indirect. Usuellement cette distinction est décrite dans la littérature relativement à l'environnement : l'interaction indirecte est traitée comme une interaction entre les agents et l'environnement alors que l'interaction directe correspond à des échanges entre agents sans passer par l'environnement.

Nous avons défendu dans la section précédente (section 2.2) que toute action, et donc toute interaction, devait passer par l'environnement. Nous distinguons bien l'interaction directe de l'interaction indirecte, mais en utilisant un critère différent de l'environnement.

L'interaction directe correspond à une interaction dirigée comme dans le cas d'envoi de messages. Un agent exploite ses capacités dans l'environnement pour émettre un message vers un ou plusieurs agents désignés dans l'environnement.

L'interaction indirecte quant à elle n'est pas dirigée, elle est au contraire caractérisée par un découplage du nom, du temps et de l'espace [Weyns *et al.*, 2004a]. Ainsi pour interagir de manière indirecte, les agents n'ont pas besoin de connaître explicitement les autres agents, ils n'ont pas besoin d'être situés au même endroit et ils n'ont pas besoin de co-exister en même temps.

Nous présenterons tout d'abord l'interaction indirecte qui a été baptisé *stigmergie*. Ensuite nous présenterons les langages de communication agent et les protocoles d'interaction qui sont les principaux points qui ont été étudiés pour l'interaction directe.

2.3.1 Stigmergie

Le terme fut introduit par le biologiste français Pierre-Paul Grassé en 1959 [Grassé, 1959], en référence au comportement des termites. Il le définit comme la « stimulation des travailleurs par l'œuvre qu'ils réalisent ». Le terme provient des mots grecs *στίγμα* (*stigma*) « marque, signe » et *εργον* (*ergon*) « travail, action », exprimant la notion que les actions d'un agent laissent des signes dans l'environnement, signes perçus par lui-même et les autres agents, et qui déterminent leurs prochaines actions.

L'exemple le plus classique de stigmergie est celui de la phéromone qui est déposée dans l'environnement par les fourmis. Cette trace déposée dans l'environnement permet à l'ensemble des fourmis de se guider dans un vaste environnement alors que celles-ci ne disposent pas de mémoire et que leurs capacités de perception dans l'environnement sont locales.

La stigmergie permet une forme de communication non-symbolique entre les agents, elle peut donc être exploitée par des agents réactifs.

2.3.2 Langage de communication agent

Les langages de communication agent, souvent désignés par l'acronyme ACL (*Agent Communication Language*), ont retenu une grande attention de la part de la communauté multi-agents. La communication est indispensable entre les agents, sans quoi ils se trouvent tous être isolés et incapables de se coordonner pour réaliser des tâches collectivement et résoudre les différents conflits qui proviennent de la concurrence de leurs activités.

A la base de toutes les théories de la communication se trouve celle de Shannon [Shannon et Weaver, 1949]. Dans cette théorie, l'acte de communication correspond à l'envoi d'information par un émetteur vers un destinataire au travers d'un canal de communication. L'information transmise est encodée dans un langage à l'émission et décodée à la réception.

La théorie de Shannon a établi un support pour les aspects techniques de la communication mais elle n'aborde pas les autres aspects de la communication. Par la suite, [Austin, 1962] puis [Searle, 1969] ont développé la théorie des actes du langage qui a influencé tous les travaux suivants sur la communication.

La théorie des actes du langage considère trois aspects dans la communication :

1. L'acte *locutoire* : comment le message est formulé.
2. L'acte *illocutoire* : ce que cela signifie pour l'émetteur ou comment cela est compris par le destinataire.
3. L'acte *perlocutoire* : l'effet produit par l'acte illocutoire sur le destinataire.

Les actes de langages ont été classés suivant le performatif; Searle et Vanderveken en distinguent cinq : *assertif*, *directif*, *promissif*, *expressif* et *déclaratif*. D'autres travaux considèrent des performatifs supplémentaires.

La théorie des actes du langage a fourni le support pour la définition de la sémantique d'un langage en termes des états mentaux des agents [Cohen et Levesque, 1990]. Le langage FIPA-ACL [FIPA, 2001] est l'héritier de cette approche de définition de la sémantique : chaque performatif est associé à des pré-conditions et des post-conditions sur les états mentaux des agents qui communiquent.

Par ailleurs les travaux sur les ACL distinguent trois niveaux dans la structure d'un message [Koning et Pesty, 2001; Eijk, 2002] :

1. *Le niveau du contenu* qui est décrit dans un langage de contenu comme KIF, PROLOG ou FIPA-SL.
2. *Le niveau du message* qui est défini par le performatif du message, une ontologie, le langage, etc.
3. *Le niveau de la communication* qui permet de décrire les méta-données relatives à la communication telles que l'émetteur du message, le destinataire, le protocole utilisé, etc.

Le tableau 2.3.2 présente suivant ces trois niveaux les paramètres qui sont utilisés pour décrire un message FIPA-ACL.

COMMUNICATION	MESSAGE	CONTENU
:sender	:performative	:content
:receiver	:language	
:reply-to	:encoding	
:protocol	:ontology	
:conversation-id		
:reply-with		
:in-reply-to		
:reply-by		

TAB. 2.1 – Les paramètres des messages FIPA-ACL

Les ACL disposant d'une sémantique précise ont remplacé les langages de communication *ad-hoc* ce qui a constitué une évolution fondamentale pour les SMA. Le fait que la sémantique du langage soit connue à l'avance permet de concevoir séparément des agents et d'assurer qu'ils se comprennent mutuellement.

Cependant, l'approche qui consiste à définir la sémantique d'un langage en termes des états mentaux des agents est maintenant largement critiquée [Singh, 1998]. En effet cette approche impose une sérieuse contrainte sur l'autonomie et l'hétérogénéité des agents car elle présuppose que les agents s'expriment comme ils pensent. Elle présuppose la condition de *sincérité* pour les agents.

Par exemple, la pré-condition du performatif **Inform** dans le langage FIPA-ACL est que l'agent croit lui-même la propriété qu'il annonce à un autre agent. Et la post-condition de ce performatif est que le destinataire du message croit alors que l'émetteur croit que la propriété énoncée est vraie.

Cette sémantique est restrictive, elle ne considère pas que les agents puissent mentir. Elle réduit l'hétérogénéité et l'autonomie des agents, mais de plus, elle inadaptée au contexte des systèmes ouverts pour lesquels il n'est pas possible de supposer la sincérité des agents.

Nous verrons à la section 2.4 les nouvelles propositions qui sont faites pour définir la sémantique des langages de communication en exploitant la dimension normative et institutionnelle. Ces propositions sont des alternatives qui ne se fondent pas sur les états mentaux des agents.

2.3.3 Protocole d'interaction

Les protocoles d'interaction ont été proposés pour déterminer la structure des interactions entre les agents et ainsi faciliter leur coordination. Un protocole d'interaction spécifie des règles qui doivent être respectées par les agents durant une conversation, et définit ainsi pour chaque étape les types de messages qui peuvent être envoyés.

En suivant un protocole, un agent interprète les messages d'une conversation un par un, en changeant son propre état à chaque étape, et exploite le protocole pour produire le prochain message de la conversation.

Un protocole d'interaction possède quatre caractéristiques [Ferber, 1995] : (i) une conversation débute par un performatif qui exprime l'intention de l'agent qui débute la conversation, (ii) à chaque étape de la conversation il y a un ensemble fini d'actions possibles, (iii) certains états sont finaux, et quand ils sont atteints la conversation est terminée, (iv) quand un acte de langage est réalisé, l'état de la conversation ainsi que les états mentaux des agents sont changés.

Le réseau contractuel (Contract Net Protocol)

Un protocole des plus classiques est le réseau contractuel (*Contract Net Protocol*) proposé par [Davis et Smith, 1983]. La spécification de ce protocole par FIPA est donnée sur la figure 2.4.

Ce protocole offre une solution au problème d'allocation de tâches qui reprend le fonctionnement du marché. Un agent peut jouer deux rôles par rapport à ce protocole, celui d'initiateur et celui de participant. Le protocole comporte trois étapes. Dans la première, l'initiateur réalise un appel d'offre, c'est-à-dire qu'il annonce à l'ensemble des participants la tâche qu'il souhaite voir réaliser. Dans la deuxième étape, les participants évaluent leurs capacités ainsi que leurs intérêts à réaliser cette tâche. Si ils trouvent leur intérêt à réaliser cette tâche, ils font alors une offre à l'initiateur. Enfin dans la dernière étape, l'initiateur sélectionne un ou plusieurs participants parmi tout ceux qui ont fait une offre et les informe de son choix.

Les protocoles d'interaction, définis comme des automates à états, permettent de donner un cadre à l'interaction entre les agents. Cependant ils limitent l'autonomie des agents

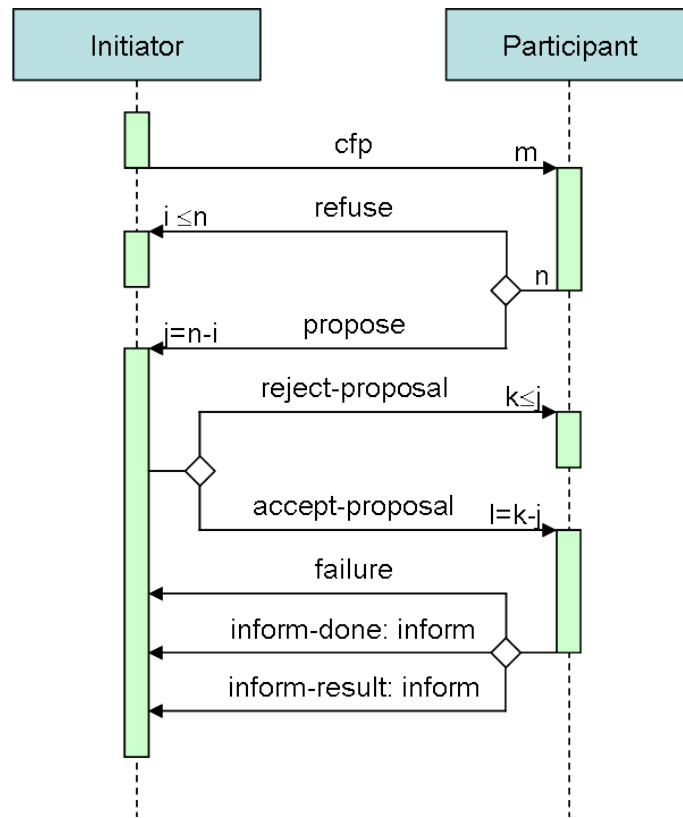


FIG. 2.4 – Le protocole d'interaction FIPA Contract Net

en leur imposant une conduite à suivre.

2.3.4 Conclusion

Pour qu'un SMA puisse atteindre un certain objectif, il est nécessaire qu'au niveau local les agents soient capables de se comprendre, de se répartir le travail, ou encore de résoudre des conflits qui apparaissent entre eux. Ainsi dans cette section nous avons présenté certains mécanismes qui sont proposés pour faciliter la coordination et la coopération entre agents. Cependant, les approches que nous venons de voir présentent un inconvénient important : elles affectent l'autonomie et l'hétérogénéité des agents.

C'est pour pallier à cette limitation que la dimension sociale d'un SMA a été étudiée plus spécifiquement. La prochaine section est consacrée aux travaux qui se sont portés sur cet aspect.

2.4 Dimension sociale

Les aspects sociaux et normatifs ont suscité et suscitent encore une très vive attention de la part de la communauté multi-agents. Un très large panel de concepts a été introduit dans le cadre des SMA dont les organisations, les normes, les relations déontiques, les

engagements sociaux ou encore les institutions. Les travaux dans ce domaine puisent largement leur inspiration dans la métaphore sociale car bon nombre de problèmes qui sont envisagés dans le cadre des SMA se sont posés et ont été étudiés dans le cadre des sociétés humaines. Ainsi les recherches visant à introduire une dimension sociale dans les SMA ont un caractère particulièrement pluridisciplinaire ; ces recherches s’inspirent souvent de travaux effectués dans des disciplines telles que la sociologie, la philosophie ou encore le droit. La particularité du domaine de l’informatique est de devoir fournir des formalisations opérationnelles à ces notions conceptuelles afin de pouvoir les manipuler en machine.

Ces notions ont été étudiés principalement pour apporter des réponses aux quatre problèmes fondamentaux suivants :

1. assurer la coordination dans un SMA ouvert et hétérogène ;
2. la dualité de l’autonomie et du contrôle ;
3. relier le niveau macro et le niveau micro ;
4. sortir de la conception « mentaliste » des SMA.

Avant de présenter les éléments de réponses qui sont proposés, analysons les enjeux de ces différents problèmes.

Assurer la coordination dans un système ouvert et hétérogène

Nous avons déjà abordé le problème de la coordination dans un SMA de manière générale (section 2.3), nous allons le considérer ici dans le cadre d’un système ouvert et hétérogène, ce qui le rend d’autant plus complexe. La citation suivante illustre cela : “Like humans, software agents will not work together just because they happen to be together, but require some external incentives for collaboration” [Dignum *et al.*, 2002c].

Dans un système ouvert, les agents sont développés par des concepteurs différents, d’où leur hétérogénéité, et ils peuvent intégrer à tout moment le système. Ceci pose les problèmes suivants :

- l’architecture interne des agents n’est pas connue à l’avance, il n’est donc pas possible de faire d’hypothèse sur celle-ci ;
- l’hétérogénéité des agents induit une nature très incertaine de l’interaction, et donc un risque de coût important pour assurer la coordination (en temps et en ressources) ;
- les agents étant des inconnus les uns pour les autres, il se pose un problème de confiance qui peut entraver leurs échanges ;
- l’hétérogénéité des buts des différents agents peut être très forte, ce qui peut conduire à de nombreux conflits ;

- on ne peut pas supposer une obéissance quelconque de la part des agents à des directives données par le concepteur du système, les agents doivent être considérés comme potentiellement mal intentionnés.

Il est donc nécessaire de définir des moyens explicites et externes aux agents pour favoriser leur coordination et assurer une certaine stabilité globale du système.

La dualité de l'autonomie et du contrôle

Dans sa thèse de doctorat [Stratulat, 2002] souligne la dualité qu'il existe dans les SMA entre le besoin d'autonomie au niveau des agents et le besoin de contrôle au niveau du système. L'autonomie est l'essence même des agents, c'est parce que les agents agissent de manières autonomes que le système résultant est capable de s'adapter aux situations nouvelles et d'évoluer. C'est aussi l'autonomie qui rend possible l'émergence nécessaire à l'auto-organisation d'un système. Mais plus l'autonomie des agents est importante et plus il est difficile de prévoir le résultat de leurs interactions. Un SMA, comme tout système informatique, est conçu pour répondre à certains objectifs et il est donc nécessaire d'avoir un certain niveau de contrôle pour garantir certaines propriétés au niveau global. Il s'agit de trouver le bon compromis entre le niveau d'autonomie et le niveau de contrôle. Si le niveau de contrôle est absolu alors le système devient totalement prévisible. Or l'intérêt du paradigme multi-agents est justement d'apporter des réponses pour les problèmes dont on n'est pas en mesure de prévoir à l'avance une bonne solution. Mais à l'inverse si l'autonomie des agents est trop importante le système risque alors d'être chaotique et contre performant. Par exemple, si chaque agent utilise son propre langage de communication, aucun n'est en mesure de comprendre les autres. La communication ici n'est qu'un exemple, il va de même pour toutes les formes d'interactions. De manière générale, plus l'autonomie des agents est importante, et plus la complexité de la coordination de leur comportement est importante, ce qui inévitablement réduit la performance globale du système.

En abordant le problème du contrôle du point de vue de l'agent, la seule solution envisageable est la définition de standards (tel que FIPA) qui engendrent une homogénéisation des agents qui n'est pas souhaitable.

De plus la définition de standards n'est pas applicable aux systèmes ouverts car les agents sont conçus séparément. Hors, dans ces systèmes plus que tout autres, il faut être en mesure de se prémunir contre des agents qui pourraient avoir un comportement nuisible au système.

Des solutions peuvent être apportées par l'*enrégimentation* qui consiste à fournir des solutions logicielles ou matérielles pour rendre impossible certains comportements. Un

exemple tiré de la vie courante est celui d'une porte verrouillée à clef; l'effet attendu est d'empêcher l'accès à une pièce à toute personne n'en possédant pas la clé. Dans un SMA l'enrégimentation peut s'effectuer par le biais de l'environnement. Ce type de solutions est intéressant pour les comportements que l'on souhaite rendre réellement impossible, en particulier pour tous les aspects pour lesquels la sécurité est particulièrement sensible. Néanmoins il n'est pas souhaitable d'appliquer ce type de solutions dans toutes les circonstances car cela affecte l'autonomie des agents. L'enrégimentation résulte d'un processus *off-line*, c'est-à-dire réalisé durant la conception du système. Les contraintes qui sont ainsi définies ne peuvent pas être adaptées dynamiquement durant l'exécution du système. L'intérêt se limite donc à empêcher les comportements dont on a la certitude au moment de la conception du système qu'ils seront néfastes en toutes circonstances.

Relier le niveau macro et le niveau micro

Il est particulièrement difficile de prévoir au niveau macro le résultat d'un ensemble d'activités autonomes qui se produisent au niveau micro. De manière duale, il est très complexe de définir un système répondant à certaines exigences uniquement à partir des spécifications des entités du niveau micro. Il est donc nécessaire de se donner des moyens permettant de relier l'activité locale et autonome d'un agent aux propriétés attendues au niveau global du système.

De plus les organisations ont des buts qui leurs sont propres, comme par exemple une entreprise dont le but est de produire un certain produit. Il n'est pas réaliste d'attendre d'un agent autonome qu'il soit concerné par les buts d'une organisation par le simple fait qu'il intègre cette organisation. Si l'on pense à un employé qui intègre une entreprise, il est probable que son but sera uniquement de percevoir un salaire. Il faut donc établir un lien explicite entre le but de l'organisation et l'activité des agents qui intègre une organisation.

Sortir de la conception « mentaliste » des SMA

L'approche mentaliste est une tradition héritée de l'intelligence artificielle classique. Cette approche tend à définir la spécification d'un système à partir des états mentaux de ses participants. La définition de la sémantique des langages de communication pour les agents en est un des exemples les plus représentatifs. Comme cela a été argumenté dans [Singh, 2000], cette conception des choses n'est pas satisfaisante pour la communication. La communication est un phénomène de nature publique, alors que les états mentaux sont intrinsèquement privés. Les états mentaux d'un agent ne peuvent pas être vérifiés sans violer l'intégrité d'un agent. La définition de la sémantique des communications en termes d'états mentaux imposent donc la condition de sincérité pour les agents. Cette condition limite l'autonomie et l'hétérogénéité des agents, et elle n'est pas acceptable

pour un système ouvert.

L'utilisation des états mentaux est tout à fait adaptée à la spécification du comportement des agents, mais elle ne l'est pas pour définir des sémantiques publiques.

Contrôle social

Pour répondre à ces différents problèmes, une partie de la communauté s'est intéressée à la mise en œuvre d'un contrôle social pour les SMA. Le concept de contrôle social dénote la capacité d'un groupe ou d'une société à s'autoréguler ainsi que sécuriser la cohérence et l'unité dans la vie sociale [Martindale, 1978]. Le contrôle social appliqué aux SMA est une réification du concept de contrôle dans le paradigme multi-agent : le contrôle est réalisé de manière décentralisé et autonome par les agents eux-mêmes.

Nous allons présenter dans la suite de cette section les concepts qui ont été proposés dans cette direction : les *organisations*, les *normes*, les *engagements sociaux* et les *institutions*.

2.4.1 Organisation

Dans les premiers temps, la communauté multi-agent s'est principalement focalisé sur la notion d'agent et les relations entre les états internes d'un agent et son comportement. Mais plus récemment la nécessité de prendre en compte la dimension sociale a été mise en évidence [Gasser *et al.*, 1987; Fox, 1988; Demazeau et Costa, 1996; Ferber et Gutknecht, 1998; Jennings, 2000]. Les travaux dans cette nouvelle direction ont introduit les notions d'*organisation*, de *rôle*, de *groupe* et de *tâche* pour les SMA.

Deux approches de conception des SMA ont alors été distinguées [Lemaitre et Excellence, 1998] : les approches centrées agent et les approches centrées organisation.

Approches centrées agent

Les approches centrées agents se focalisent sur la description de la structure interne des agents sans définir de structures collectives au niveau du système. La structure du système est donc entièrement définie au travers des agents : potentiellement un agent peut communiquer avec n'importe quel autre agent, c'est donc de la responsabilité de chacun que de gérer son accessibilité ainsi que de définir les relations, les contrats, etc. qu'il entretient avec les autres agents.

Dans cette situation, Jennings et Wooldridge ont souligné les limites de cette conception pour construire des SMA [Wooldridge *et al.*, 2000] :

“A common misconception is that agent-based systems require no real structuring. Although perhaps true in certain cases, it should not be viewed as the only way of developing agent societies. Many agent systems require considerably more system-level engineering, particularly large-scale systems or those in which the society is supposed to act with some commonality of purpose. In such cases, a means of structuring the society is needed to reduce the system’s complexity, increase the system’s efficiency, and more accurately model the problem being tackled.”

D’un point de vue plus génie logiciel, l’absence de structures collectives est un obstacle à la sécurité, la modularité et la réutilisabilité des éléments du système.

Approches centrées organisation

C’est pour dépasser ces limitations que la métaphore sociale a été proposée. Elle consiste à appliquer pour les SMA des principes qui ont été identifiés par la sociologie dans les sociétés humaines. La motivation de cette démarche est de considérer que les sociétés humaines sont un exemple concret de SMA ouvert qui fonctionne.

D’un point de vue sociologique, Morin a donné la définition suivante de l’organisation :

“L’organisation peut être définie comme un agencement de relations entre composants ou individus qui produit une unité, ou système, dotée de qualités inconnues au niveau des composants ou individus. L’organisation lie de façon inter-relationnelle des éléments ou événements ou individus divers qui dès lors deviennent les composants d’un tout. Elle assure solidarité et solidité relative, donc assure au système une certaine possibilité de durée en dépit des perturbations aléatoires.” [Morin, 1977]

Dans le contexte des SMA « informatiques », [Ferber *et al.*, 2003] définissent trois principes qui caractérisent une approche centrée organisation :

Principe 1 : le niveau organisationnel décrit le « quoi » et pas le « comment ». Il impose une structure sur les schémas d’activité des agents, mais il ne décrit pas comment les agents se comportent.

Principe 2 : le niveau organisationnel ne contient aucune description des états internes des agents. Ainsi une organisation peut aussi bien contenir des agents réactifs que des agents intentionnels. Le niveau organisationnel ne fournit que des descriptions de comportements attendus.

Principe 3 : une organisation fournit un moyen de partitionner le système, chaque partition (ou groupe) définit un contexte d'interaction pour les agents. Les agents membres d'un groupe peuvent interagir librement, mais un groupe est opaque pour tous les agents qui sont externes au groupe.

Pour un SMA les organisations permettent de structurer les entités du système. Lorsqu'un agent est amené à interagir, l'organisation lui indique généralement avec quel autre agent le faire et comment le faire, lui évitant de faire cette recherche par lui-même [Gutknecht, 2001]. De ce point de vue, l'organisation permet de réduire la complexité de l'espace de recherche qu'un agent doit envisager pour atteindre son but. D'un point de vue global, l'organisation permet de limiter le nombre de conflits et donc d'augmenter l'efficacité globale de la société.

D'autre part, la notion d'organisation permet d'accroître la modularité et la réutilisation dans la conception de SMA. Le rôle est la représentation abstraite d'une fonction, d'un point d'interaction ou d'une identification d'un agent au sein d'un groupe particulier [Gutknecht, 2001]. Les rôles permettent de définir des schémas génériques d'interaction : les comportements sociaux ne sont pas seulement décrits au niveau des agents mais aussi au niveau du système.

On peut distinguer pour une organisation la *structure organisationnelle* de l'*organisation concrète* [Ferber et al., 2003]. La structure organisationnelle est ce qui persiste quand les composants ou les individus rejoignent ou quittent une organisation, c'est-à-dire les relations qui font d'un agrégat d'éléments un tout.

Une organisation concrète est une instanciation d'une structure organisationnelle : elle est la donnée des entités qui prennent effectivement part dans le tout et des relations qu'ils entretiennent à tout moment.

Une organisation comporte un aspect *structurel* et un aspect *fonctionnel*. L'aspect structurel est lié à la structure de partition et à la structure de rôle, c'est-à-dire les relations qui sont établies entre les groupes et entre les rôles. L'aspect fonctionnel exprime comment l'activité des individus est reliée à l'activité de l'organisation. Il fait intervenir des notions de *tâche*, d'*objectif* et de *contrat*.

AGR (Agent-Groupe-Rôle)

Le modèle AGR [Ferber et al., 2003] (anciennement appelé Aalaadin [Ferber et Gutknecht, 1998]) est un modèle organisationnel minimal qui repose sur trois concepts primitifs : l'*agent*, le *groupe* et le *rôle* qui sont structurellement connectés et qui ne peuvent pas

être définis par d'autres primitives.

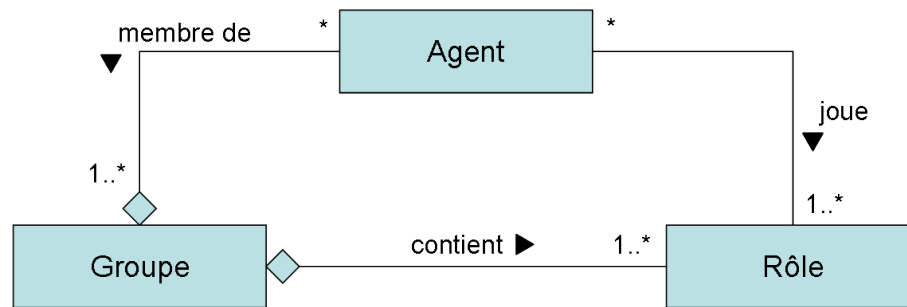


FIG. 2.5 – UML : Le méta-modèle AGR

Ces trois concepts primitifs sont unis par un ensemble d'axiomes :

Agent : un agent est une entité active et communicante qui joue des *rôles* dans des *groupes*. Un agent peut jouer plusieurs rôles et être membre de plusieurs groupes. Aucune contrainte n'est imposée sur l'architecture interne de l'agent ou sur ses facultés mentales. Un agent peut être aussi réactif qu'une fourmi ou bien aussi intelligent qu'un humain.

Groupe : un groupe est un ensemble d'agents partageant un ensemble de caractéristiques. Un groupe fournit un contexte pour des *patterns* d'activité. De plus les groupes sont utilisés pour partitionner des organisations. Pour communiquer, deux agents doivent appartenir à un même groupe, mais un agent peut appartenir à différents groupes. Cette caractéristique permet de définir des structures organisationnelles.

Rôle : un rôle est la représentation abstraite de la fonction d'un agent dans un groupe. Un agent doit jouer au moins un rôle dans un groupe, mais il peut en jouer plusieurs. Les rôles sont locaux aux groupes et sont attribués à un agent sur sa demande. Un même rôle peut être joué par plusieurs agents.

La représentation UML du méta-modèle AGR est donnée sur la figure 2.5.

Une notation est associée au méta-modèle AGR pour décrire une organisation concrète (figure 2.6).

D'autre part, la plateforme MadKit [Gutknecht et Ferber, 2000], qui est elle-même construite sur le modèle AGR, permet de concevoir des SMA suivant ce modèle.

MOISE

Le modèle *MOISE* (*Model of Organization for multi-agent System*) [Hannoun et al., 2000] et ses extensions *MOISE⁺* [Hübner et al., 2002] et *MOISE^{Inst}* [Gâteau et al., 2005] sont des modèles organisationnels qui prennent en compte aussi bien les aspects

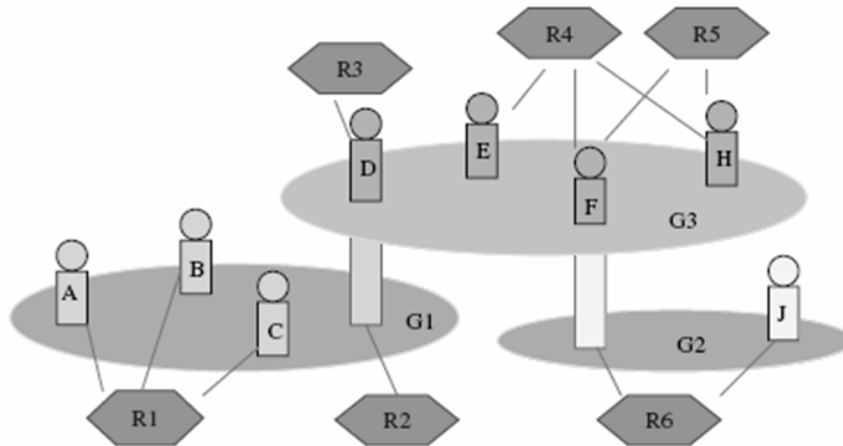


FIG. 2.6 – UML : Le méta-modèle AGR

structurels que les aspects fonctionnels.

Trois niveaux de spécification sont distingués :

Spécification structurelle La spécification structurelle d'une organisation est exprimée en termes de groupes, de rôles et de relation entre rôles. Les groupes sont composés de sous-groupes et de rôles. L'ensemble des rôles sont définis selon une hiérarchie avec une relation d'héritage. Enfin les liens suivants peuvent être établis entre deux rôles :

- *lien d'acointance* : pour structurer la représentation des autres agents ;
- *lien de communication* : pour structurer les échanges d'informations ;
- *lien d'autorité* : pour définir une structure de contrôle ;
- *lien de compatibilité* : pour définir les rôles qui peuvent être joués par un même agent.

Spécification fonctionnelle La spécification fonctionnelle permet d'exprimer comme des *buts collectifs* sont décomposés par des *plans* et répartis sur les agents par des *missions*.

Spécification normative ou déontique La spécification de normes est utilisée dans MOISE pour relier la spécification structurelle et la spécification fonctionnelle. Les normes permettent de décrire pour les agents, en fonction de leurs rôles, quelles sont les missions pour lesquelles ils peuvent s'engager.

L'intérêt de ce modèle est de permettre de relier explicitement les objectifs d'une organisation avec l'activité de ses membres en exploitant la structure de l'organisation.

2.4.2 Norme

La notion de norme a été étudiée dans de multiples disciplines telles que la philosophie, les sciences sociales ou encore le domaine juridique. Une expression normative décrit comment les choses doivent être ou ce qui doit être fait, par opposition à une expression descriptive qui décrit comment les choses sont. La notion de normativité sous-entend celle d'idéal : elle conduit dans une société à porter des jugements de valeur, à décrire ce qui est bien et ce qui est mal.

Plus récemment cette notion a fait l'objet de nombreuses études dans le domaine des SMA. Dans cette section nous allons présenter quels sont les intérêts à intégrer des normes dans les SMA et différents travaux qui ont été proposés pour aller dans cette direction.

Une norme est la description syntaxique d'un comportement (ou d'une situation) idéal dans une société. Une norme se présente à l'individu sous la forme d'un ensemble de règles ou de modèles dont il doit s'inspirer pour orienter son action dans la société.

Dans les SMA, la norme peut devenir un moyen privilégié pour décrire les objectifs poursuivis par le système et les modalités pour le satisfaire par le biais des agents. Pour cela, il faut que les normes soient des composantes externes aux agents ; qu'elles puissent être communiquées aux agents et que ceux-ci soient capables de les interpréter.

L'intérêt principal de la norme est de constituer un compromis entre le besoin d'autonomie et le besoin de contrôle. Une norme sert à prescrire aux agents un comportement à adopter dans certaines situations, ce qui permet de les guider dans leur comportement. Cependant une norme n'est pas une contrainte absolue, ce qui permet de préserver l'autonomie des agents : un agent peut en toutes circonstances décider de *violier* la norme. Evidemment la seule présence de normes dans le système est insuffisante pour produire un effet sur le comportement des agents, il faut que celles-ci soient assorties de sanctions pour avoir un caractère coercitif. Une sanction peut être positive, dans ce cas elle correspond à une rétribution qui sera donné à l'agent qui se conforme à la norme, ou bien négative, et dans ce cas une pénalité sera appliquée à l'agent qui ne respecte pas la norme.

Finalement les normes vont constituer un moyen d'influencer les agents dans leur comportement tout en préservant leur libre arbitre en dernier recours. En situation, un agent décidera quelles actions il doit réaliser d'après sa rationalité, en prenant en compte son environnement, les normes auxquelles il est soumis, et les sanctions qu'il peut encourir si il les viole.

L'usage de normes induit aussi un impact positif sur la coordination en réduisant

l'incertitude des interactions dues à l'autonomie des agents. Les normes permettent dans certaines situations pour un agent d'anticiper ou d'attendre certains comportements de la part des autres agents. Le code de la route est un très bon exemple de normes qui permettent d'améliorer significativement la coordination et ainsi l'efficacité du système. Imaginez seulement combien il serait complexe de se déplacer sur la route si la convention de rouler à droite n'était pas adoptée. On voit très bien sur cet exemple, que si la norme est avant tout une contrainte ou une pression sur l'activité d'un agent, elle peut tout à fait être bénéfique pour l'agent qu'elle contraint en réduisant le coût qu'engendre la coordination.

Les normes peuvent être élaborées à différents niveaux :

- elles peuvent être définies par le concepteur du système, elles résultent dans ce cas d'un processus de conception *off-line* ;
- elles peuvent être établies et modifiées durant l'activité du système par certains agents qui en ont la responsabilité (on parle alors d'autorité) ;
- elles peuvent être déterminées par émergence de part l'activité de tous les agents qui sont dans le système.

On parle de *systèmes normatifs* lorsqu'un système utilise des normes explicites pour réguler les interactions [Stratulat, 2002].

Nous avons maintenant une meilleure idée générale de ce que peut apporter le concept de normes pour la conception de SMA ; mais pour être en mesure de profiter de ces normes dans un système informatique, il faut être en mesure de leur donner un caractère opérationnel qui va permettre de les manipuler. Nous allons donc présenter les différents travaux qui ont été effectués dans ce sens.

Classification des normes

Il est tout d'abord nécessaire de caractériser plus précisément cette notion de norme. Tuomela a proposé une classification des normes dans laquelle quatre types sont distingués [Tuomela, 1995] :

1. les *r-normes* sont des règles explicites établies dans une société par une autorité. Elles sont souvent associées à des sanctions. On y obéit car on les accepte, cette acceptation est relative à l'adhésion même de la société ;
2. les *s-normes* sont les normes sociales à proprement dites, elles sont basées sur des croyances et des acceptations mutuelles. Elles concernent les conventions qui s'ins-

tallent entre les membres d'une société ou d'un groupe spécifique. On y obéit car les autres attendent que l'on s'y conforme ;

3. les *m-normes* sont des normes morales ; on y obéit non pas pour répondre à une attente sociale mais à cause de sa conscience individuelle ;
4. les *p-normes* sont les normes de prudence, on y obéit car c'est une chose rationnelle à faire.

Les positions normatives

Hohfeld est le premier à avoir fourni une théorie des droits [Hohfeld, 1917]. Cette théorie décrit ce qu'il appelle les *notions juridiques fondamentales* : le droit, la liberté (ou privilège), le pouvoir, l'immunité, l'absence de droit, le devoir, l'incapacité et la responsabilité. Chaque notion possède sa notion opposée (table 2.2) et sa notion duale (table 2.3). Hohfeld a étudié ces notions de droit dans la perspective relationnelle en les axant sur la relation qu'il peut y avoir entre deux agents. Par exemple, si *A* a un droit sur *B*, cela signifie que *B* a un devoir envers *A*. Si *B* n'a pas de devoir, alors *B* est libre. Et si *B* est libre, alors *A* n'a pas le droit de l'empêcher de faire quelque chose.

droit	absence de droit
liberté	devoir
pouvoir	incapacité
immunité	responsabilité

TAB. 2.2 – Les notions juridiques opposées

droit	devoir
liberté	absence de droit
pouvoir	responsabilité
immunité	incapacité

TAB. 2.3 – Les notions juridiques duales

Le but de cette théorie était d'une part de lever l'ambiguïté sur les notions juridiques distinctes et d'autre part d'établir des bases pour un raisonnement juridique précis. Cependant Hohfeld ne disposait pas des outils de formalisation logique qui ont été développés par la suite, sa théorie reste donc à un niveau conceptuel. Ce n'est que plus tard qu'une formalisation logique de cette théorie a été définie [Kanger, 1957].

Logique déontique

La représentation des normes fait intervenir des expressions que l'on appelle *déontiques* telles que l'*obligation*, la *permission*, l'*interdiction*, le *droit*, le *pouvoir*, le *devoir*, etc. La formalisation de ces notions a donné naissance à la *logique déontique*. On reconnaît à Von Wright le mérite d'avoir énoncé la première axiomatique solide pour définir une logique déontique [von Wright, 1951]. Le système présenté initialement par von Wright est construit dans une logique de premier ordre, mais la plupart des travaux qui ont suivi ont

développé la logique déontique comme une sous-branche de la logique modale. Cela s'explique par le fait qu'il y a une forte similarité entre les notions d'obligation, d'interdiction et de permission, et celles de nécessité, d'impossibilité et de possibilité.

Une logique modale est construite comme une extension de la logique propositionnelle classique à laquelle on ajoute un ou plusieurs opérateurs modaux. Dans le cas de la logique déontique, les opérateurs modaux qui sont introduit sont : O pour l'obligation, P pour la permission et I pour l'interdiction. Ainsi si α est une formule propositionnelle, alors $O\alpha$, $P\alpha$ et $I\alpha$ sont des formules qui peuvent être exprimées dans le langage naturel respectivement par : « il est obligatoire que α », « il est permis que α » et « il est interdit que α ».

Les relations entre les opérateurs sont décrites par les définitions suivantes :

$$I\alpha =_{def} O\neg\alpha \quad (2.1)$$

$$P\alpha =_{def} \neg O\neg\alpha \quad (2.2)$$

La sémantique généralement utilisée est celles des mondes possibles de Kripke [Kripke, 1963].

Pour plus d'information sur les logiques déontiques le lecteur peut se référer aux ouvrages suivants : [Chellas, 1980; Bailhache, 1991].

Normes et organisations

Les normes ont souvent été étudiées dans le contexte d'organisation. Elles sont alors utilisées pour décrire le comportement qui est attendu d'un agent en fonction du rôle qu'il joue dans une organisation [Dignum *et al.*, 2002a; Dignum *et al.*, 2002b; Hübner *et al.*, 2002; Tranier *et al.*, 2003].

2.4.3 Agent normatif

Nous avons discuté de l'intérêt de mettre en place des normes dans un SMA ainsi que des outils proposés pour les formaliser. Mais pour que les normes puissent avoir un impact sur le comportement d'un agent, il est nécessaire qu'il soit capable de les prendre en compte au moment de décider du comportement à adopter. Il faut donc proposer de nouvelles architectures internes d'agent permettant d'intégrer les normes dans le raisonnement des agents.

Des architectures ont été proposées pour définir des agents qui se soumettent totalement aux normes. De notre point de vue ce type d'approches dénature l'intérêt

profond des normes qui réside dans la possibilité de violer les normes. L'agent doit prendre en compte les normes dans un sens utilitaire ou rationnel, si le comportement qui est prescrit par une norme apparaît comme plus néfaste pour l'agent que la sanction qu'il encourt, alors celui-ci devrait violer la norme.

Nous voulons faire remarquer qu'il n'est pas nécessaire que tous les agents qui intègrent un système normatif aient des capacités cognitives qui leur permettent d'appréhender les normes. L'activité d'agents réactifs peut être encadrée par d'autres agents qui eux intègrent la dimension normative du système. Pour faire un parallèle avec nos propres sociétés, peu de personnes sont capables de comprendre l'intégralité de nos lois. D'autre part certaines de nos lois ont pour objet les animaux : prenons comme exemple une loi qui interdirait le pâturage de moutons dans une certaine zone ; il sera alors de la responsabilité du berger de s'assurer que les moutons n'entrent pas sur la zone en question.

Extension de BDI pour prendre en compte les normes

Dans [Dignum *et al.*, 2000], Dignum et al. proposent une extension de l'interpréteur BDI afin de prendre en compte un cadre normatif dans le raisonnement de l'agent. Pour cela ils considèrent deux types de normes : les normes qui interdisent la réalisation de certaines actions (ex. il est interdit de fumer au restaurant), et les normes qui impose que certaines actions soient réalisées ou qu'un certain état soit atteint (ex. l'obligation de payer avant de partir du restaurant).

Par rapport à l'algorithme BDI que nous avons présenté (algorithme 1), les fonctions `options` et `filter` doivent être adaptées :

- La fonction `options(B,I,oblEvents)` prend en compte non seulement les croyances (`B`) et les intentions (`I`) mais aussi les obligations (`oblEvents`) pour réaliser la prise de décision.
- La fonction `filter(B,D,I,oblEvents)` inclue maintenant les normes restrictives (`oblRestr`) afin d'assurer que l'intention que l'agent va finalement choisir est bien légale.
- La fonction `reconsider(I,B,oblEvents)` doit aussi inclure les obligations (`oblEvent`) car durant l'exécution d'un plan, un agent doit aussi considérer l'apparition de nouvelles obligations qui seraient plus importantes que l'intention actuelle.

2.4.4 Engagement social

La notion d'engagement social a été proposée il y a quelques années dans le cadre des SMA pour rendre possible la communication et l'interaction entre agents hétérogènes. Un

```

Algorithme : BDI-interpret
B := B_init;
I := I_init;
while true do
  get_perception(perc);
  B := belief_revision(B,perc);
  D := options(B,I,oblEvents);
  I := filter(B,D,I,oblRestr);
  plan = find_plan(B,I);
  while NOT (empty(plan) OR succeeded(I,B) OR impossible(I,B)) do
    action = next_action(plan);
    execute(action);
    get_perception(perc);
    B := belief_revision(B,perc);
    if reconsider(I,B,oblEvents) then
      D := options(B,I,oblEvents);
      I := filter(B,D,I,oblRestr);
    end
    if not(sound(plan,I,B)) then
      plan = find_plan(B,I);
    end
  end
end

```

Algorithme 4: L'interpréteur BDI étendu

des objectifs est de proposer une alternative à l'approche mentaliste de la communication.

Nous allons voir que la notion d'engagement social est complémentaire de la notion de norme. Alors que la norme est impersonnelle, les engagements sociaux vont porter sur des agents et seront directement manipulés par ceux-ci. Avant d'aller plus loin, il est important de préciser que les engagements sociaux ne doivent pas être confondus avec les engagements individuels ou internes qui sont en rapport avec la persistance de l'intention d'un agent [Cohen et Levesque, 1990]. De plus les engagements sociaux ne peuvent pas être réduits aux engagements internes [Singh, 1991].

Castelfranchi définit les engagements sociaux comme des engagements contractés par un agent (le débiteur) envers un autre agent (le créancier), soulevant l'attente que le débiteur va agir en conséquence pour satisfaire ses engagements [Castelfranchi, 1995]. De nombreuses formalisations ont été proposées pour exprimer cette notion dont [Singh, 1997; Singh, 1999; Castelfranchi, 1995; Fornara et al., 2007; Pasquier et al., 2004]. Nous allons tout d'abord présenter une formalisation la plus générale possible en nous inspirant des notations utilisées dans [Fornara et al., 2007] et nous discuterons ensuite des différentes

spécificités que l'on peut trouver dans la littérature.

Formalisation générale

Un engagement social peut être représenté de la manière suivante :

$$Comm(state, debtor, creditor, content)$$

où *state* désigne l'état dans lequel se trouve l'engagement, *debtor* désigne le débiteur (l'agent qui est engagé), *creditor* désigne le créancier (l'agent envers qui l'engagement est souscrit) et *content* décrit le contenu de l'engagement (l'expression de l'engagement qui est pris par le débiteur envers le créancier).

Bien évidemment le point crucial de la formation d'un engagement social est la description de son contenu. Nous allons présenter le formalisme des *propositions temporelles* proposé dans [Fornara et Colombetti, 2003; Colombetti et al., 2003].

Une proposition temporelle permet de relier un énoncé avec un intervalle de temps. L'aspect temporel est essentiel pour les engagements : un engagement doit être tenu sur une certaine période ou bien accompli avant une certaine date limite. L'énoncé d'une proposition temporelle est une proposition logique qui peut décrire soit un état des choses, soit l'exécution d'une action spécifique. Une proposition temporelle est aussi caractérisé par un mode qui peut être \forall ou \exists . Le mode \forall indique que l'énoncé de la proposition temporelle doit être vérifié sur tout l'intervalle de temps. Le mode \exists qui est dual du premier indique que l'énoncé de la proposition temporelle doit avoir été validé au cours de l'intervalle de temps. Enfin une proposition temporelle possède un état qui à tout instant est soit indéfini (noté \perp), validé (noté *true*) ou invalidé (noté *false*).

La forme générale d'une expression temporelle s'exprime de la manière suivante :

$$TP(statement, [t_{start}, t_{end}], mode, state)$$

Lorsqu'un engagement est contracté, la proposition temporelle qui lui correspond est dans l'état indéfini (\perp). Il sera de la responsabilité d'une autorité de transformer cet état en validé (*true*) ou invalidé (*false*) en suivant les règles suivantes :

- si *mode* = \forall , l'état (*state*) doit prendre la valeur invalidé (*false*) si l'énoncé *statement* est falsifié durant l'intervalle de temps $[t_{start}, t_{end}]$; sinon l'état doit passer à valider (*true*) dès que l'intervalle de temps $[t_{start}, t_{end}]$ est révolu;
- si le *mode* = \exists , l'état (*state*) doit prendre la valeur validé (*true*) si l'énoncé *statement* est vérifié durant l'intervalle de temps $[t_{start}, t_{end}]$; sinon l'état (*state*)

doit passer à invalidé (*false*) dès que l'intervalle de temps $[t_{start}, t_{end}]$ est révolu.

Il faut noter que les bornes de l'intervalle de temps peuvent être infinies, ce qui permet d'exprimer un engagement permanent.

Par exemple la proposition « l'enchère doit être ouverte dans les 10 minutes » se représente de la manière suivante :

$$TP(openAuction(id), [now, now + 10min], \exists, \perp)$$

État d'un engagement social

A tout moment un engagement social est caractérisé par son état qui peut prendre les valeurs suivantes : *inactif* (*unset*), *actif* (*pending*), *annulé* (*cancelled*), *violé* (*violated*) ou *accompli* (*fulfilled*). La figure 2.7 présente le cycle de vie d'un engagement social, les arcs indiquant les transitions possibles entre tous les états.

Un engagement est inactif lorsqu'il vient d'être créé mais qu'il n'est pas encore en vigueur. Du moment où l'engagement entre en vigueur son état devient actif. Un engagement actif peut être annulé dans certaines circonstances par l'une des deux parties ou par une intervention extérieure ; un engagement annulé ne contraint plus son débiteur. Enfin un engagement acquiert l'état accompli lorsque son débiteur en a rempli les clauses, ou au contraire il acquiert l'état violé lorsque le débiteur n'a pas respecté une des clauses de l'engagement.

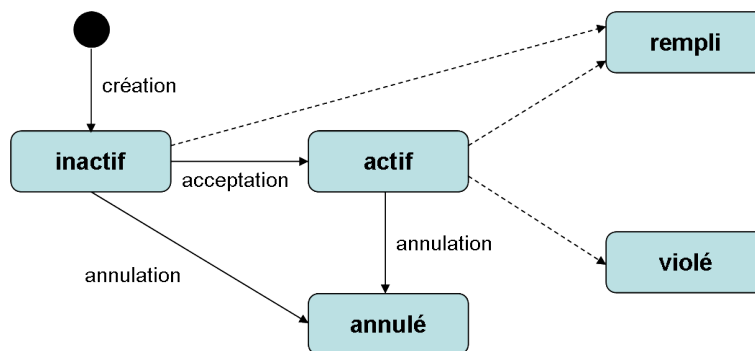


FIG. 2.7 – Cycle de vie d'un engagement social

Opérations sur les engagements sociaux

Pour réaliser ces changements d'états des opérations sont définies sur les engagements sociaux. On peut distinguer les opérations qui peuvent être effectuées directement par les agents, on parle alors d'*action sociale*, des opérations qui seront automatiquement effectuées par le contexte.

On trouve dans la littérature les actions sociales classiques suivantes :

- *Création* : opération qui permet d'établir un nouvel engagement social dans l'état inactif.
- *Acceptation* : opération qui permet de transformer un engagement social inactif en engagement social actif. Cette opération est parfois omise; dans ce cas c'est l'opération de création qui établit directement un engagement social dans l'état actif.
- *Annulation* : opération qui permet de transformer un engagement social actif ou inactif en engagement social annulé.

Les opérations qui permettent de transformer l'état en accompli ou violé ne sont généralement pas traitées comme des actions sociales; c'est le contexte qui permet de l'établir. Cependant il n'est pas toujours possible de déterminer de manière objective si un engagement est rempli ou non. Ces situations nécessitent alors de mettre à la disposition d'une autorité les opérations nécessaires pour transformer l'état d'un engagement social actif.

D'autre part Singh propose de considérer les actions sociales supplémentaires suivantes [Singh, 1999] :

- *Délégation* : opération qui permet de changer le débiteur de l'engagement social.
- *Transfert* : opération qui permet de changer le créancier de l'engagement social.
- *Libération* : Singh distingue la libération de l'annulation : alors que l'annulation peut induire des pénalités pour le débiteur de l'engagement, la libération elle a pour seul effet de donner l'état annulé à l'engagement.

Politiques sociales

Afin de règlementer l'utilisation des opérations sur les engagements sociaux, il faut définir un ensemble de *politiques sociales*. Une politique sociale sert à décrire qui a la possibilité, et dans quelle situation, de réaliser une opération sur un engagement social (comme sa création, son acceptation, ou son annulation) et quels sont les effets de cette opération. Une politique sociale est donc une expression conditionnelle qui met en relation

les engagements et les opérations sur les engagements. Différentes politiques sociales sont possibles suivant la nature que l'on souhaite donner aux engagements sociaux, en voici quelques exemples classiques :

$$\begin{aligned}
 REQUEST. \quad & \forall x, y \in Ag \quad request(y, x, content) \Rightarrow Comm(x, y, content, unset) \\
 CREATE. \quad & \forall x, y \in Ag \quad create(x, y, content) \Rightarrow Comm(x, y, content, pending) \\
 CANCEL. \quad & \exists c = Comm(x, y, content, pending) \wedge cancel(y, c) \\
 & \Rightarrow c = Comm(x, y, content, cancelled)
 \end{aligned}$$

La politique *REQUEST* décrit qu'un agent peut solliciter un engagement de la part d'un autre agent, ce qui a pour effet de créer un engagement inactif ; cet engagement devra être accepté par son débiteur pour devenir actif. La politique *CREATE* décrit qu'un agent peut s'engager de sa propre initiative, ce qui a pour effet de créer un engagement actif. Enfin la politique *CANCEL* décrit que seul le débiteur d'un engagement actif peut décider de son annulation.

Extensions

Nous allons maintenant présenter quelques extensions à cette notion de base d'engagement que l'on peut trouver dans la littérature.

Castelfranchi introduit la notion de *témoin* [Castelfranchi, 1995] : un engagement social doit être contracté dans une perspective publique.

Dans [Singh, 1997], Singh propose d'inclure explicitement des clauses d'annulation pour un engagement social. L'annulation d'un engagement social n'est plus uniquement régi par les politiques sociales mais peut aussi dépendre de clauses spécifiques déterminées à la création de l'engagement.

Dans [Singh, 1999], Singh associe à un engagement social le groupe dans lequel il est contracté qui sert de contexte. Cette notion de contexte englobe la notion de témoin proposée par Castelfranchi. D'autre part, Singh propose que les engagements puissent être contractés envers un groupe. L'expression du créancier peut donc désigner un agent unique ou un groupe d'agents. Enfin Singh propose de traiter les engagements sociaux comme des objets de première classe qui peuvent être nommés. De cette manière les engagements peuvent être référencés et l'objet d'un engagement peut porter sur un autre engagement, introduisant ainsi la notion d'engagement d'*ordre supérieur*. En adoptant cette approche, Singh considère les politiques sociales comme des méta-engagements, ou des engagements d'ordre 1.

Dans [Pasquier *et al.*, 2004] les auteurs associent explicitement à un engagement social les pénalités qui peuvent être appliquées si cet engagement n'est pas respecté par

son débiteur. Cette approche permet de négocier les pénalités d'un engagement à sa création et permet donc plus de flexibilité à l'exécution du système.

Conclusion

Pour conclure sur les engagements sociaux nous allons présenter les différents contextes pour lesquels ils ont été étudiés.

Tout d'abord les engagements sociaux permettent de favoriser la coordination car ils rendent les comportements des agents plus prévisibles. Un engagement permet de faire bénéficier à son créancier de la pression sociale qui peut être exercée par le groupe ou la société, ce qui renforce la confiance au sein d'un système ouvert. Mais contrairement à l'enrégimentation, les engagements sociaux n'imposent pas un comportement aux agents, c'est donc une solution plus souple pour préserver l'autonomie des agents.

Les engagements sociaux ont été exploités pour définir une sémantique sociale aux langages de communication [Singh, 2000; Flores et Kremer, 2002; Colombetti *et al.*, 2003; Chaib-draa *et al.*, 2006; Fornara *et al.*, 2007]. Ces approches consistent à considérer les performatifs comme des actions sociales qui engendrent des engagements sociaux. Par exemple la sémantique du performatif *Inform* n'est plus définie en termes d'états mentaux, mais en termes d'engagement social : lorsqu'un agent informe un autre agent, il s'engage au près de celui-ci sur la véracité de l'information énoncée. La condition de sincérité n'est plus nécessaire, un agent peut mentir, cela revient à ne pas respecter un engagement.

Les engagements sociaux ont aussi été utilisés pour définir des protocoles d'interaction [Venkatraman et Singh, 1999; Yolum et Singh, 2001; Yolum et Singh, 2002; Fornara et Colombetti, 2003; Pasquier *et al.*, 2004]. Ces approches permettent de définir des protocoles flexibles car ils n'enrégimentent pas le comportement des agents qui les utilisent.

[Carabelea et Boissier, 2006] ont proposés d'utiliser les engagements sociaux dans le cadre d'organisations . Le principe est d'associer des politiques sociales à la spécification des rôles d'une organisation. L'acquisition d'un rôle par un agent est alors considéré comme la souscription d'un certain nombre d'engagements sociaux envers l'organisation. Les engagements sociaux permettent donc de décrire ce qui est attendu d'un agent qui joue un certain rôle.

2.4.5 Contrôle et sanction

Tout le pouvoir coercitif d'une dimension normative repose sur les moyens qui sont mis en place pour contrôler l'activité des agents et leur appliquer les sanctions prévues lorsque cela est nécessaire. Que ce soit dans le cas des normes ou dans le cas des engagements sociaux, ce sont les sanctions encourues qui permettent d'influencer les agents dans leur comportement.

Contrôle

Pour que le contrôle social soit possible, l'activité des agents doit pouvoir être surveillée, et cela nécessite de garantir une perspective publique à l'action. C'est le rôle de l'environnement que de rendre observable le comportement des agents. Ainsi, pour qu'un contrôle social puisse être mis en œuvre, toutes les actions doivent être effectuées dans l'environnement et laisser des traces dans celui-ci, ce qui assure que l'interaction ne peut pas se produire « en cachette ».

[Kaminka *et al.*, 2002] ont proposé le principe de l'*écoute flottante*, qui est une technique de surveillance non-intrusive et qui repose sur l'environnement. Kaminka *et al.* définissent l'écoute flottante de la manière suivante : “*Overhearing is an indirect interaction whereby an agent receives information for which it is not address*”. Le caractère non-intrusif est essentiel pour préserver l'intégrité des agents.

Sanction

Dans [Pasquier *et al.*, 2004], les auteurs présentent une ontologie de la sanction dans laquelle ils caractérisent une sanction par sa direction, son type et son style.

Il existe deux directions possibles pour une sanction :

- *Positive* : la sanction correspond à une récompense donnée pour encourager un certain comportement.
- *Négatives* : la sanction est une pénalité appliquée pour décourager un certain comportement ; cette direction est utilisée dans les cas de violation de normes.

Quatre types de sanctions sont distingués :

- *Automatique* : la sanction est automatiquement imposée à celui qui réalise une certaine action ; l'action transporte elle-même sa propre sanction. C'est par exemple le cas lorsqu'un robot percute un mur ou encore une action qui nécessitait une

coordination qui a échoué.

- *Matérielle* : les sanctions matérielles incluent les sanctions physiques (violence ou soin) et les sanctions financières comme dans le cas d’une amende. Ce type de sanction peut être appliqué immédiatement ou au contraire différé dans le temps.
- *Sociale* : Les sanctions sociales affectent la réputation ou la crédibilité d’un agent au sein d’un groupe. Comme c’est indiqué dans [Posner et Rasmusen, 1999], les sanctions sociales correspondent généralement à une action qui transporte une information sur le contrevenant qu’il aurait préféré que les autres ne connaissent pas. Par exemple, le fait qu’un agent annule fréquemment ses engagements sans raison explicite pourra être pris en compte par les autres agents pour évaluer sa réputation ou la confiance qu’ils lui portent.
- *Psychologique* : les sanctions psychologiques nécessitent des architectures internes d’agents qui manipulent des émotions. Des sanctions psychologiques vont pouvoir induire des sentiments tels que la culpabilité ou la honte.

Les auteurs distinguent aussi le *style* d’une sanction. Une sanction peut être *implicite* ou *explicite*. Une sanction implicite est déterminée de manière autonome et unilatérale par les agents. Les sanctions de cette nature posent une difficulté : elles ne sont pas connues publiquement et il est donc nécessaire de découvrir si l’on sera puni et comment. Les sanctions explicites quant à elles sont exprimées au préalable ; elles sont publiquement connues de tous les agents qui interagissent.

Enfin pour concevoir un système de contrôle social il faut mettre en œuvre des stratégies d’application des sanctions. D’après les théoriciens du contrôle social, il existe cinq philosophies de la punition à partir de laquelle toute stratégie d’application des sanctions peut être définie [Vold *et al.*, 2002] :

- *Dissuasion* : c’est une philosophie qui considère que l’objectif de la sanction est de dissuader. Pour être efficaces, les sanctions doivent être appliquées rapidement, de manière certaine et sévère. Les sanctions étant importantes, ce type de stratégies réduit fortement la flexibilité.
- *Vengeance (retribution)* : cette philosophie considère que la violation doit être réparée par une pénalité aussi sévère que le dommage qui a été causé.
- *Invalidation* : cette philosophie considère le handicap ou la restriction des capacités d’un agent comme le moyen de l’empêcher de commettre une nouvelle violation. Dans les sociétés humaines l’exclusion et l’emprisonnement sont les méthodes d’invalidation les plus utilisées.
- Les deux dernières philosophies sont la *réhabilitation* et la *restauration*. Nous n’en discuterons pas ici car elles ne sont pas applicables à l’heure actuelle pour les SMA.

2.4.6 Institution

Plus récemment la notion d'*institution* a été proposée pour définir des systèmes normatifs. On retrouve cette notion dans la littérature sous les appellations *institution électronique*, *e-institution*, ou encore *institution artificielle*. Le terme institution en lui-même n'est pas clair dans la langue naturelle. On l'utilise parfois pour parler d'organisations établies (typiquement celles qui ont un caractère public) avec un code de loi, comme par exemple un hôpital ou une université. On l'utilise aussi pour parler d'un ensemble de concepts qui existent uniquement parce qu'il y a un accord collectif au sein d'une société, comme c'est le cas pour la monnaie, la propriété ou le mariage.

Dans la communauté multi-agents il y a aussi différentes acceptions de la notion d'institution. Elle est utilisée pour parler d'organisations spécifiques [Esteva *et al.*, 2001], ou de schémas abstraits destinés à réguler les interactions entre les agents [Vázquez-Salceda et Dignum, 2003], ou encore pour se référer à la description de concepts partagés et de règles qui établissent et régulent une réalité institutionnelle [Báez-Barranco *et al.*, 2006; Fornara *et al.*, 2007].

Comme les normes et les engagements sociaux, la notion d'institution provient de l'analyse des sociétés humaines. Dans son livre *Institutions, Institutional Change and Economic Performance* [North, 1990], l'économiste North motive les institutions humaines par la nature incertaine des interactions humaines ce qui entraîne des coûts de transaction dans les échanges. Les institutions sont un moyen de réduire cette incertitude afin de permettre une meilleure efficacité dans les échanges. Les institutions peuvent être informelles (normes de comportement ou règles de conduites) ou formelles (lois, règles). Ces deux formes d'institutions nécessitent une mise en application des règles qu'elles prescrivent, et plus les sociétés deviennent complexes, plus il y a un besoin pour que cette mise en application soit réalisée par une tierce partie. Dans nos sociétés humaines, ce rôle est rempli par l'état avec sa force coercitive. Les institutions fournissent la structure de base par laquelle les êtres humains peuvent créer un ordre social et essayer de réduire l'incertitude dans l'échange.

North distingue la notion d'institution de celle d'organisation, il est important d'insister sur ce point car la distinction n'est pas toujours bien établie dans la littérature. Une institution se réfère à un système de règles abstraites destinées à réguler les échanges. Au contraire une organisation est concrète; une organisation est toujours associée à certains objectifs, et elle exploite les possibilités données par les institutions pour maximiser leur intérêt. Par exemple, les sociétés exportatrices sont des organisations qui exploitent les possibilités données par l'institution du commerce international.

ISLANDER

ISLANDER [Rodríguez-Aguilar, 2001; Esteva *et al.*, 2002] est un outil qui a été proposé pour spécifier et vérifier des institutions électroniques multi-agents. L'objectif principal de cet outil est de permettre la conception d'infrastructure pour les SMA ouverts.

Dans ce travail, une institution est décrite à partir de quatre éléments de base : *cadre dialogique*, *scene*, *structure performative* et *norme*.

Cadre dialogique : Un cadre dialogique fournit le contexte pour l'interaction, en termes de concepts du domaine (ontologie) et de langage de communication (illocutions). De plus un cadre dialogique définit les rôles que les participants peuvent avoir dans une institution et les relations entre ces rôles. Chaque rôle définit un motif de comportement dans l'institution, et il est exigé que tout agent dans l'institution adopte certains de ces rôles.

Scene : Les scènes sont des protocoles d'interaction bien définis entre différents rôles. Une scène correspond à la modélisation d'une conversation possible dans l'institution. Les agents ont la possibilité d'entrer et de sortir d'une scène à certains états de la conversation et en fonction des rôles qu'ils jouent.

Structure performative : Les structures performatives permettent de connecter les scènes en un réseau de scènes pour décrire des activités complexes. De plus, elles permettent de décrire comment un rôle permet de passer d'une scène à une autre.

Norme : Les communications d'un agent dans une scène peuvent avoir pour conséquences d'imposer des obligations à l'agent et de modifier ses possibilités dans la structure performative. Cet aspect est décrit par des normes, et les normes elles-mêmes sont décrites par un ensemble de conditions qui active les obligations et un ensemble de conditions qui les désactive.

Dans ISLANDER, la structure performative est une structure enrégimentée qui établit les comportements possibles pour les agents. Les transitions définies dans les scènes expriment, à partir d'un état donné, quelles sont les possibilités d'un agent. Enfin les normes permettent de décrire dans des situations données, le comportement qui est attendu d'un agent.

Cette approche est intéressante pour la conception d'un système ouvert car elle permet d'une part de décrire un contexte d'activité qui s'impose aux agents sans faire de présupposition sur leur architecture interne, et d'autre part elle permet une certaine forme de régulation des activités avec les normes.

Cependant, nous pensons que le formalisme qui est proposé ne capture pas toute la complexité de la notion d'institution car il y a une assimilation du contexte d'activité (ou environnement) et de l'institution. Ainsi, il n'est pas possible de distinguer dans ce formalisme la notion de *possibilité* et la notion de *pouvoir institutionnel*.

Nous allons maintenant présenter une approche différente pour décrire les institutions.

2.4.7 Searle : construction de la réalité sociale

Les travaux de Searle sur la construction de la réalité sociale [Searle, 1997; Searle, 2005] deviennent très influents dans les recherches sur les SMA [Jones et Sergot, 1996; Artikis *et al.*, 2002; Boella et Torre, 2004; Vázquez-Salceda *et al.*, 2005; Báez-Barranco *et al.*, 2006]. Dans cette section, nous allons tout d'abord présenter l'analyse des institutions et des mécanismes sous-jacents qui est proposée par Searle. Ensuite nous présenterons deux modèles qui sont directement inspirés de cette conception des institutions.

Réalité brute et réalité institutionnelle

Afin d'appréhender le fonctionnement des institutions, Searle fait tout d'abord la distinction entre les *faits bruts* et les *faits institutionnels*. Un fait brut est un énoncé qui peut être vérifié dans la nature, il correspond à un phénomène observable : « *Cette pierre est plus grosse que telle autre* » ou bien « *la Terre est à une distance (moyenne) de 150 000 000 km du soleil* » sont des exemples de faits bruts. Au contraire, un fait institutionnel nécessite la présence d'institutions pour pouvoir être établi. « *je suis citoyen français* » ou encore que « *ce bout de papier est un billet de 5 €* » sont des exemples de faits institutionnels. L'ensemble des faits bruts sont donnés par ce que l'on pourrait appeler la réalité brute, et l'ensemble des faits institutionnels relatifs à une institution constitue une réalité institutionnelle.

Si l'on considère un match de football, au niveau brut nous ne pouvons simplement constater que plusieurs corps et un ballon se trouvent sur un espace physique particulier, et que chacun évolue selon certaines trajectoires. Maintenant au niveau de la réalité institutionnelle nous pouvons parler de joueurs de football, de buts, de hors-jeu, etc.

Pour Searle, trois notions primitives sont nécessaires pour expliquer l'apparition de faits institutionnels :

1. *L'intentionnalité collective* :

De la même manière que l'on parle de l'intentionnalité d'un individu, on peut parler

d'intentionnalité pour un groupe d'individus. Searle prend pour exemple une meute de loup qui chasse pour se nourrir ou encore les membres d'un même parti politique qui désirent la victoire de leur candidat. Dans ces deux exemples il existe une intentionnalité de groupe et cette intentionnalité ne peut pas se réduire à la somme des intentionnalités individuelles.

2. *L'attribution de fonction :*

Les êtres humains ont cette capacité (ainsi que certaines espèces d'animaux) d'attribuer des fonctions à des personnes ou des objets qui n'ont pas intrinsèquement cette fonction. L'exemple type est celui des outils : nous donnons à tel type d'objet la fonction de « marteau ».

3. *Les fonctions de statut :*

C'est une sorte de fonction particulière qui peut être attribuée à une personne ou un objet, mais que cette personne ou cet objet n'est pas capable de remplir uniquement en vertu de sa structure physique. Les fonctions de statut vont pouvoir être remplies uniquement en vertu du fait qu'il y a une acceptation collective que l'objet ou la personne visée a bien un certain statut. La monnaie en est une illustration : un billet de banque n'est en soi qu'un morceau de papier sans valeur, mais il y a une acceptation collective que ce morceau de papier est un billet de banque (statut), ce qui transforme notre attitude vis-à-vis de cet objet et lui permet de remplir sa fonction monétaire.

Finalement pour Searle, les faits institutionnels résultent de la combinaison de ces trois notions primitives, à savoir *l'attribution collective d'un certain statut*.

Règle constitutive

Searle a formalisé les attributions collectives de statut par des *règles constitutives* qui sont de la forme :

$$X \text{ counts-as } Y \text{ in context } C$$

où X désigne une caractéristique d'un objet, d'une personne ou un état des choses, Y désigne statut particulier qui est attribué à cet objet, cette personne ou cet état des choses, et C désigne un contexte particulier. Cette formulation « *X counts as Y in context C* » peut se traduire par « *X compte comme Y dans le contexte C* » ou encore « *X a valeur de Y dans le contexte C* ».

Les règles constitutives permettent de faire un passage des faits bruts aux faits institutionnels, et ainsi de construire une réalité institutionnelle à partir de la réalité brute. En voici quelques exemples informels :

Exemple 2.1 – *Counts-as*.

- au football telle et telle série de mouvements compte pour un but ;
- telle et telle procédure compte pour l'élection du président de la république ;
- telle et telle position aux échecs compte pour un échec-et-mat.

Nous pouvons maintenant donner la définition de Searle pour une institution [Searle, 2005] :

“An institution is any collectively accepted system of rules (procedures, practices) that enable us to create institutional facts. These rules typically have the form of X counts as Y in C where an object, person, or state of affairs and lower and lower X is assigned a special status, the Y status, such that the new status enables the person or object to perform functions that it could not perform solely in virtue of its physical structure but requires as a necessary condition the assignment of the status.”

Règle constitutive et règle régulative

Il y a une différence fondamentale entre ce que Searle appelle les *règles constitutives* et les *règles régulatives*.

Les règles régulatives permettent de réguler une activité existante, alors que les règles constitutives sont à l'origine de la possibilité même d'une nouvelle activité. Par exemple l'obligation de rouler à droite sur la route est une règle qui régule l'activité de rouler à droite. Maintenant si l'on considère les règles du jeu d'échecs, leur fonction n'est pas de réguler cette activité, mais au contraire de donner la possibilité de jouer aux échecs.

Pour faire la distinction entre ces deux types de règles, il suffit de considérer les conséquences de la suppression des règles. Si l'on supprime les règles du jeu d'échecs, il n'est tout simplement plus possible de jouer aux échecs. Ces règles ne sont pas là pour contraindre l'activité des joueurs, mais au contraire pour leur fournir le moyen de construire une réalité dans laquelle il est possible de jouer.

Par contre, si l'on supprime la règle qui oblige de rouler à droite sur la route il est toujours possible de rouler.

Searle formalise les règles régulatrices par une expression de la forme suivante :

Do X

Pouvoir institutionnel

Une notion très importante qui se dégage de cette analyse est celle de *pouvoir institutionnel*. Un statut confère à l'objet ou la personne à qui il est attribué des *pouvoirs institutionnels*. Un pouvoir institutionnel est une capacité à établir certains faits institutionnels, et par conséquent à transformer la réalité institutionnelle. Ainsi un joueur de football acquiert la capacité de marquer un but, un joueur d'échecs de faire échec-et-mat, et une personne qui a des billets de banque en poche acquiert du pouvoir d'achat. Sans institution, rien de tout cela ne serait possible.

Capacité matérielle, permission, pouvoir institutionnel

Nous pouvons maintenant distinguer trois notions qui sont fondamentalement différentes : la *capacité matérielle* (ou possibilité), la *permission* et le *pouvoir institutionnel*.

Une capacité matérielle est une capacité d'action dans la réalité brute ; elle s'impose aux agents. Par exemple, les être humains ont la capacité de marcher mais pas celle de voler par leurs propres moyens. Cela relève des lois de la nature et les êtres humains ne peuvent rien y changer.

La permission est une notion déontique qui correspond à une absence d'interdiction. L'interdiction de réaliser une action α n'empêche en rien de réaliser α . Par contre cela peut conduire à encourir une sanction pour avoir violé une règle.

Enfin, un pouvoir institutionnel est encore une notion différente qui permet comme nous l'avons dit de modifier une réalité institutionnelle. L'usage d'un pouvoir institutionnel s'effectue en réalisant les actions matérielles décrites par une règle constitutive. Par exemple, dans un certain contexte, le fait de lever la main dans la réalité brute permet d'exprimer un vote dans une réalité institutionnelle. Lever la main est une capacité matérielle alors qu'exprimer un vote est un pouvoir institutionnel.

Les pouvoirs institutionnels sont conférés par des statuts institutionnels. Par exemple lorsqu'un individu qui possède le statut de maire réalise un certain protocole, cela a pour effet de marier deux personnes, c'est-à-dire de leur attribuer un nouveau statut. Si un individu qui ne possède pas le statut de maire réalise le même protocole, cela n'aura tout

simplement aucun effet sur la réalité institutionnelle.

Il est important de remarquer que dans certaines situations une personne peut avoir un pouvoir institutionnel mais pas la permission de l'utiliser. Prenons comme exemple l'inscription à l'université. Les membres du personnel administratif qui sont responsables des inscriptions ont le pouvoir d'attribuer le statut d'étudiant, mais ils n'ont la permission de l'utiliser que pour des personnes qui remplissent les conditions requises par l'université. En pratique une personne qui ne remplit pas les conditions peut très bien se voir attribuer le statut d'étudiant (qui est matérialisé par la carte d'étudiant). Dans ce cas il y a une violation des règles régulatrices de la part du personnel administratif qui peut engendrer des sanctions à leur encontre, mais l'attribution de statut est bien effective, et la personne désignée pourra être reconnue comme étudiant. Ainsi l'usage abusif d'un pouvoir institutionnel ne remet pas en question les effets engendrés sur une réalité institutionnelle. Pour rétablir une telle situation, il est nécessaire d'utiliser un « contre-pouvoir » institutionnel pour supprimer le statut précédemment attribué.

Formalisation de l'opérateur *counts-as*

Jones et Sergot [Jones et Sergot, 1996] ont proposé une formalisation logique du pouvoir institutionnel en s'inspirant de l'opérateur *counts-as* de Searle. Ils proposent l'énoncé suivant pour capturer la notion de pouvoir institutionnel :

$$E_x A \Rightarrow_s E_y F \quad (2.3)$$

où \Rightarrow_s désigne la notion de conséquence exprimée par la relation *counts-as*, E_x est l'opérateur modal de l'action, x et y sont deux agents quelconques désignés au sein d'une institution s pour réaliser la tâche spécifiée. L'énoncé 2.3 exprime que si l'agent x réalise A par a pour conséquence que l'état des choses décrit par F est établi pour l'institution s , et tout se passe comme si l'agent y était responsable de l'apparition de F .

Jones et Sergot soulignent l'importance de relativiser l'opérateur \Rightarrow à une institution car les faits qui sont établis pour une institution ne le sont pas nécessairement pour une autre.

Le terme y peut être égal au terme x , mais ce n'est pas nécessairement le cas. De plus le terme y peut dénoter l'institution tout entière. Cette formalisation permet d'exprimer le fait d'agir au nom de quelqu'un. On peut considérer que lorsque le maire réalise le protocole consacré pour marier deux personnes, le maire agit au nom de toute l'institution du mariage, le résultat produit étant que l'institution du mariage confère à ces deux personnes le statut d'époux. Une autre situation représentative est celle de la délégation de pouvoir qu'il a lorsqu'au sein d'un département s , la signature d'une secrétaire x compte pour la signature de son chef y .

Jones et Sergot remarquent qu'il est possible de distinguer trois types de conséquences : la conséquence logique, la conséquence causale (de type *counts-as*) et la conséquence déontique. Premièrement il y a des conséquences logiques qui sont indépendantes de toute institution comme par exemple $E_y E_x A \rightarrow E_x A$. Deuxièmement toute institution est soumise à des contraintes causales, comme par exemple dans une institution s le fait qu'un agent x assure l'apparition de A cause le fait qu'un agent y cause l'apparition de F (ce qui peut ne pas être vérifié pour une autre institution). Troisièmement une institution s peut être soumise à des contraintes déontiques : il peut par exemple être vrai pour l'institution s (mais pas nécessairement pour les autres) que lorsqu'un agent x assure l'apparition de A alors un agent y a *le devoir* d'assurer l'apparition de C . Ces trois types de conséquence correspondent respectivement aux notions de capacité matérielle, de pouvoir institutionnel et de permission dont nous avons parlé précédemment. Pour capturer ces types de conséquence, Jones et Sergot proposent d'introduire un nouvel opérateur modal plus général, D_s , à partir duquel \Rightarrow_s est exprimé.

L'expression $D_s(A \rightarrow B)$ signifie « c'est une contrainte de l'institution s que si A alors B ». Ainsi l'opérateur D_s permet de décrire la réalité relative à l'institution s .

Nous ne donnons ici que la liste des axiomes utilisés pour définir l'opérateur D_s , le lecteur pouvant se référer à l'article [Jones et Sergot, 1996] pour des informations complémentaires.

$\Rightarrow_s D.$	$(A \Rightarrow_s B) \rightarrow D_s(A \rightarrow B)$
$DK.$	$D_s(A \rightarrow B) \rightarrow (D_s A \rightarrow D_s B)$
$DD.$	$D_s A \rightarrow \neg D_s \neg A$
$DP.$	$\neg D_s \perp$
$\Rightarrow_s DK.$	$(A \Rightarrow_s B) \rightarrow (D_s A \rightarrow D_s B)$
$Const.$	$(A \Rightarrow_s B) \rightarrow A \rightarrow D_s A$
$Det \Rightarrow_s .$	$(A \Rightarrow_s B) \rightarrow (A \rightarrow D_s B)$
$DEx.$	$(E_x A \Rightarrow_s E_y F) \rightarrow (E_x A \rightarrow D_s F)$

Sémantique institutionnelle du langage

Une application particulièrement intéressante de la théorie de Searle aux SMA est la définition d'une sémantique institutionnelle pour les langages de communication qui a été exposée dans [Fornara *et al.*, 2007]. Cette approche utilise la notion d'engagement social en considérant qu'un engagement social est un élément d'une réalité institutionnelle car ils relèvent d'une acceptation collective au sein d'une société. De ce point de vue, la création et les différentes opérations réalisables sur les engagements sociaux sont définies en termes

de règles constitutives. Ces opérations sont donc associées à des pouvoirs institutionnels. Par exemple « faire une promesse » (un acte de langage assertif) compte comme la création d'un engagement social.

L'intérêt de cette approche est de définir une sémantique commune à tous les agents mais sans faire intervenir leurs états mentaux. Chaque agent qui reconnaît l'institution du langage sait qu'une promesse est un engagement, mais cela ne lui permet pas de connaître les états mentaux de ses interlocuteurs, il ne peut que s'en faire des représentations.

Dans [Fornara *et al.*, 2007], les auteurs se limitent au cas de la communication : les règles constitutives leur permettent d'exprimer le passage d'un acte de langage à une action institutionnelle. Nous pensons qu'il n'y a pas de raison de se limiter à la communication, les actes de langage ne sont qu'un type particulier d'action parmi tous ceux possibles ; toute action dans l'environnement peut servir de support à une règle constitutive, et donc établir ou modifier un engagement social. Cette généralisation permet d'exploiter les engagements sociaux aussi bien dans les systèmes basés sur la communication que dans les systèmes qui font une représentation explicite d'une composante physique.

Le modèle AGREEN

Le modèle AGREEN [Báez-Barranco *et al.*, 2006] est une extension du modèle AGRE (AGR + Environnement) [Ferber *et al.*, 2004] qui lui-même est une extension du modèle AGR [Ferber *et al.*, 2003].

Alors que AGR ne permet que de décrire des environnements sociaux (avec la notion d'organisation), AGRE permet de décrire des environnements sociaux et des environnements physiques. Ces deux types d'environnement sont généralisés par la notion d'*espace*. Les agents sont situés dans des espaces, et ils peuvent agir et percevoir dans ces espaces au travers des *modes*. Un mode correspond à la manifestation d'un agent dans un espace, il permet son existence et décrit son apparence dans l'espace. Un mode dans un espace physique est un *corps*, et un mode dans un espace social est un *rôle*. AGRE permet de prendre en compte l'interaction sociale et l'interaction physique, cependant il n'y a pas de possibilité dans ce modèle de relier ces deux types d'interaction.

Le modèle AGREEN étend le modèle AGRE pour relier les espaces sociaux et les espaces physique en s'appuyant sur la notion d'institution. AGREEN permet de définir des règles constitutives (de la forme $X \text{ counts-as } Y$) entre un espace physique et un espace social ou entre deux espaces sociaux. Ces règles constitutives permettent de décrire comment l'activité qui se produit dans un espace physique ou social peut engendrer de manière causale des modifications dans un autre espace social.

Nous présenterons au chapitre 4 le méta-modèle MASQ qui exploite différemment les travaux de Searle en distinguant le niveau institutionnel du niveau social.

Conclusion

Nous pensons que cette vision des institutions proposée par Searle fournit un excellent support à la dimension normative d'un SMA. L'institution fournit non seulement une structure dans laquelle les normes peuvent être définies (ou les règles régulatrices dans la terminologie de Searle), mais elle ouvre aussi de nouvelles perspectives. Alors que les études portant sur les normes se sont focalisées sur les notions d'obligation, d'interdiction et de permission, les institutions permettent aussi de capturer les positions normatives par le biais de la notion de pouvoir. D'autre part, comme Searle l'a souligné, le rôle des institutions ne se limite pas à la régulation ; elles permettent surtout d'établir des réalités institutionnelles qui sont l'expression même d'une interprétation collective. Cette interprétation collective va permettre de renforcer la compréhension mutuelle dans un système ouvert, et cela sans poser de pré-requis sur l'architecture interne des agents. L'institution décrit les règles d'interprétation commune que se donne une société indépendamment de l'interprétation individuelle qui peut être faite par chaque agent. Ainsi en toutes circonstances, un agent peut avoir sa propre interprétation des choses (ce qui relève de son autonomie), il peut même en avoir une interprétation contradictoire de celles données par ses institutions, mais il a conscience de cette interprétation institutionnelle et il va devoir la prendre en compte du fait de la pression sociale.

2.5 Conclusion

Dans ce chapitre nous avons présenté un état de l'art sur les concepts d'agent, d'environnement, d'interaction, d'organisation, de norme, d'engagement social et d'institution. Nous avons vu que chacun de ces concepts étaient nécessaires pour la conception d'un SMA.

Agent Les agents sont évidemment fondamentaux pour la conception d'un SMA. Les travaux sur les modèles d'agents proposent des solutions pour définir des comportements autonomes.

Environnement L'environnement est un élément essentiel d'un SMA. D'une part il permet de délimiter l'autonomie des agents en établissant les comportements qui sont possibles et ce qui ne le sont pas. D'autre part il permet de donner une perspective publique à l'interaction, c'est-à-dire de rendre observable le comportement des

agents. Pour que l'environnement puisse remplir ces fonctions nous avons vu qu'il devait être défini comme une entité de première classe. Nous avons aussi vu le principe de séparation esprit/corps et le principe influence/réaction qui clarifient la relation qu'entretiennent les agents avec leur environnement.

Interaction Les travaux sur l'interaction fournissent différents éléments de réponses pour permettre aux agents d'interagir avec succès.

Organisation Le concept d'organisation permet de structurer un SMA et de décrire de manière abstraite le niveau micro. Les organisations renforcent la sécurité du système en cloisonnant l'interaction. Les organisations permettent aussi de réduire la complexité de conception du système. Enfin les organisations permettent de favoriser la coordination en guidant les agents dans leurs activités.

Norme et engagement social Les concepts normatifs que sont les normes et les engagements sociaux sont nécessaires pour réguler les activités dans un SMA. Ils permettent d'influencer les agents dans leur comportement tout en préservant leur autonomie. Le comportement des agents sont alors plus prévisibles, ce qui favorise la coordination et renforce la confiance entre les agents, mais le système global préserve la possibilité de s'adapter à des situations imprévues.

Institution Le concept d'institution tel qu'il a été présenté par Searle permet de donner une sémantique publique à l'interaction, c'est-à-dire une sémantique qui ne repose pas sur les états internes des agents. Les institutions permettent ainsi de renforcer la compréhension mutuelle entre les agents, et donc de favoriser la coordination.

Ces différents concepts ont chacun fait l'objet de nombreuses études, mais à notre connaissance il n'existe pas à ce jour d'approche capable de tous les prendre en compte. Relativement à chacun de ces concepts de nombreux modèles ont été proposés, mais en l'absence de cadre conceptuel commun il est difficile d'intégrer de manière cohérente des modèles traitant de concepts différents.

C'est dans le but de clarifier les concepts manipulés pour les SMA et de rendre possible leur intégration que nous proposons un nouveau cadre conceptuel pour les SMA qui inspire des *quadrants* de la théorie de la *vision intégrale* proposée par Wilber.

Chapitre 3

Vers une vision intégrale des systèmes multi-agents

DANS ce chapitre, nous proposons un nouveau cadre conceptuel pour l'analyse et la conception de SMA qui est directement inspiré des *quadrants* de la théorie de la *vision intégrale* développée par Wilber. Nous présenterons tout d'abord en quoi consiste cette théorie puis nous motiverons l'intérêt de la considérer dans le cadre des systèmes multi-agents. Sur la base de cette théorie, nous proposerons un nouveau cadre conceptuel reposant sur quatre dimensions fondamentales pour analyser et concevoir un système multi-agent. Nous montrerons que ce cadre conceptuel fournit un support qui permet de clarifier les concepts d'agent, d'environnement, d'interaction, d'organisation, de norme et d'institution, ainsi que les modèles qui traitent de ces concepts. Ensuite nous illustrerons l'intérêt d'utiliser une approche intégrale pour analyser des problèmes d'un point de vue multi-agent. Enfin nous proposerons une approche de conception adaptée aux systèmes ouverts qui découle de ce cadre conceptuel.

Contents

3.1	La théorie de la vision intégrale	70
3.2	Pourquoi une approche intégrale pour les systèmes multi-agents ?	74
3.3	Une carte intégrale pour les systèmes multi-agents	75
3.4	Analyser au travers des 4 quadrants	81
3.5	Approche de conception pour les systèmes multi-agents	83

Dans cette thèse, nous adoptons pour les SMA une démarche fédératrice qui est fondée sur la théorie de la *vision intégrale* [Wilber, 2001]. Comme nous allons le voir par la suite, cette théorie dépasse largement le cadre de cette thèse et il est important de préciser que nous ne cautionnons ni son fondement ni les applications qui sont proposées par son auteur. Toutefois nous nous efforcerons de montrer que cette théorie constitue un cadre de réflexion intéressant en ce qui concerne l'analyse et la conception de SMA.

Dans un premier temps nous introduirons rapidement cette théorie pour ensuite présenter le nouveau cadre conceptuel que nous proposons pour les SMA.

3.1 La théorie de la vision intégrale

La vision intégrale est une théorie qui a été développée par Wilber dans le but d'intégrer la plus grande partie possible de la connaissance.

Le développement perpétuel de la connaissance humaine a pour objectif, au sens large, de mieux comprendre le monde dans lequel nous vivons. Dans la poursuite de cet objectif, un très grand nombre de théories ont été développées telle que les sciences physiques, la biologie, la théorie du chaos, la psychologie du développement, ou encore la théorie des systèmes. Chacun de ces champs propose sa manière d'approcher ou de décrire une vérité.

Face à cette diversité de la connaissance, Wilber part du postulat qu'il n'existe pas une vérité unique qui pourrait être vue au travers d'une seule discipline. Il ne cherche donc pas à déterminer laquelle serait la plus correcte, mais quel serait le système cohérent qui incorporerait le plus grand nombre de ces vérités. Sa démarche n'est pas de se demander qui a tort ou raison, mais de présupposer que toutes ces théories sont valides (au moins partiellement), et de se demander alors comment elles peuvent l'être toutes en même temps.

“I don't believe that any human mind is capable of 100 percent error. So instead of asking which approach is right and which is wrong, we assume each approach is true but partial, and then try to figure out how to fit these partial truths together, how to integrate them, not how to pick one and get rid of the others.” [Wilber, 2001]

L'objectif principal de Wilber est donc d'établir un cadre général qui permettrait de comprendre aussi bien les systèmes physiques, biologiques, économiques, sociologiques, psychiques, spirituels, etc. et d'analyser les liens qui existent entre eux.

C'est en travaillant dans cette direction que Wilber a développé ce qu'il appelle lui-même la « *vision intégrale* », ce qui signifie appréhender quelque chose au travers de tout les points de vue, le long de toutes les dimensions d'analyse possibles et dans tous les niveaux d'évolution. Le terme intégral est à prendre ici dans le sens d'intégrer, de rassembler, de mettre en relation ; l'objectif n'est pas de tout uniformiser, mais d'arriver à une unité dans la diversité.

Dans ce travail nous n'avons exploité qu'une sous-partie de la théorie de la vision intégrale : *les quadrants*. Nous revoyons donc le lecteur aux travaux de Wilber pour une présentation de l'ensemble de cette théorie [Wilber, 2001].

Les quatre quadrants

La théorie de la vision intégrale propose de décomposer toute analyse suivant deux axes :

1. la perspective *individuelle* vs. la perspective *collective* ;
2. la perspective *interne* (subjectivité, états mentaux, représentations, ...) vs. la perspective *externe* (objectivité, objet, comportement manifesté, ...)

La mise en commun de ces deux axes d'analyse constitue une carte en *quatre quadrants* (figure 3.1) où chaque quadrant correspond à un point de vue spécifique suivant lequel un individu, une situation ou un système social peut être analysé.

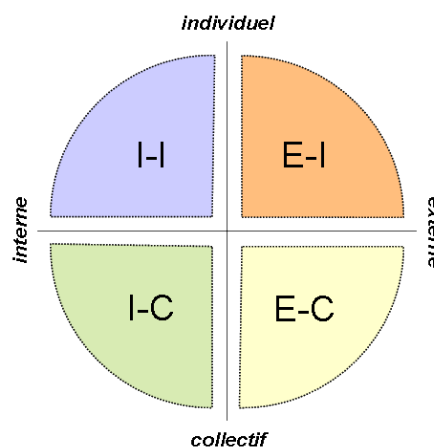


FIG. 3.1 – Les 4 quadrants

Les quatre quadrants sont : le quadrant interne – individuel (I-I), le quadrant externe – individuel (E-I), le quadrant externe – collectif (E-C) et le quadrant interne – collectif (I-C).

Interne – Individuel

Le quadrant I–I est celui de la *subjectivité*. De ce point de vue, on considère le vécu et le ressenti d'un individu qui se traduit par ses états internes, comme par exemple ses émotions, ses intentions, ses désirs, ses croyances, etc.

Externe – Individuel

Le quadrant E–I est celui de l'*objectivité*. De ce point de vue, on considère « la chose » telle qu'elle se manifeste dans le monde, indépendamment du contexte ou des interconnexions qui la relie au reste du monde. On considère donc individuellement un objet, un corps, un processus ou un comportement.

Externe – Collectif

Le quadrant E–C est celui de l'*interobjectivité*. De ce point de vue, on considère les relations complexes et interdépendantes qu'il existe entre les parties d'un système. Ce point de vue étant externe, on se focalise toujours sur ce qui est objectif, mesurable dans un système. De ce point de vue on considère par exemple l'environnement qui met en relation différents individus ou encore les structures sociales qui ont une manifestation dans l'environnement.

Interne – Collectif

Le quadrant I–C est celui de l'*intersubjectivité*. De ce point de vue, on considère les éléments subjectifs qui sont propres non pas à un individu mais à un groupe d'individus ou une société. On considère donc pour un groupe, la morale, les valeurs, les interprétations collectives, les normes sociales, les ontologies ou encore les connaissances partagées. L'ensemble des éléments subjectifs partagés par les individus d'un groupe constitue la *culture* de ce groupe.

Les quatre quadrants représentent finalement l'intérieur et l'extérieur d'un individu et d'un collectif, soit quatre perspectives différentes suivant lesquelles on peut analyser un individu, une situation ou un système social. Ces quatre perspectives d'analyse sont synthétisées sur la figure 3.2.

Pour désigner ces quatre perspectives, Wilber utilise la terminologie suivante : perspective *intentionnelle* (quadrant I–I), perspective *comportementale* (quadrant E–I), perspective *sociale* (quadrant E–C) et perspective *culturelle* (quadrant I–C). Chaque quadrant peut aussi être associé au point de vue désigné par un pronom personnel : le quadrant I–I correspond au « JE », le quadrant E–I correspond au « IL », le quadrant

E–C correspond au « ILS » et le quadrant I–C correspond au « NOUS ».

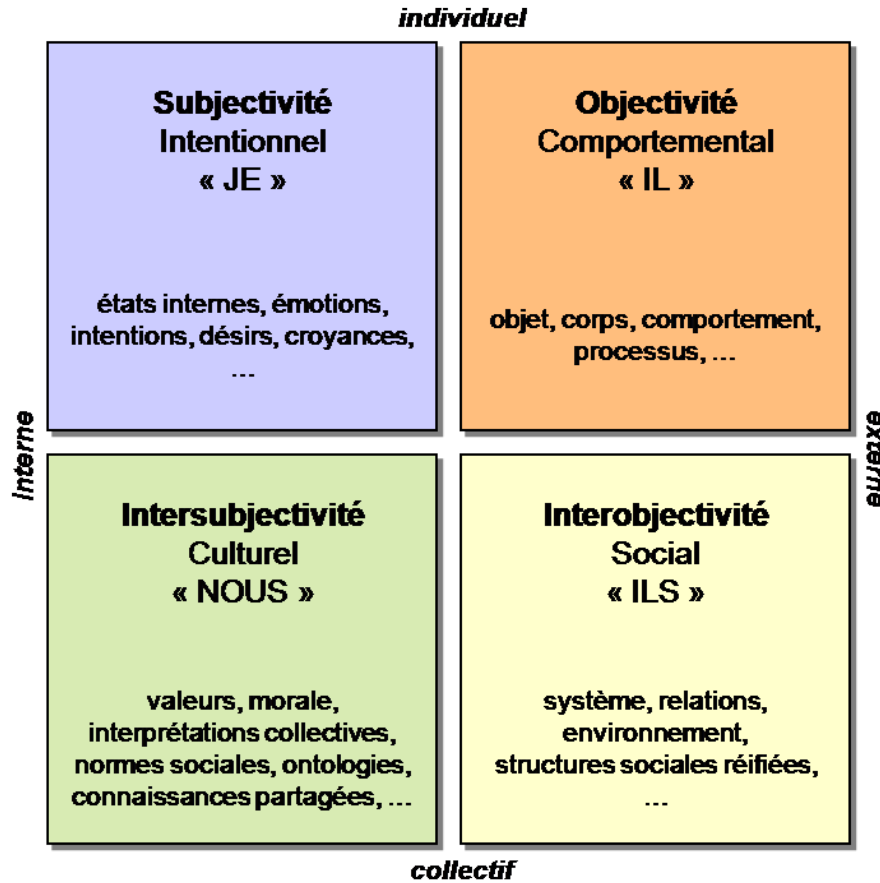


FIG. 3.2 – Les quatre quadrants : quatre points de vue

Approche intégrale

Une approche intégrale, dans le sens donné par Wilber, nécessite d'effectuer toute analyse au moins suivant les quatre points de vue correspondant aux quatre quadrants. En effet, Wilber considère que toute chose possède au moins les quatre dimensions correspondant aux quadrants, et qu'il est donc nécessaire de considérer ces quatre dimensions pour aboutir à une compréhension intégrale. La théorie de la vision intégrale ne se limite pas aux quatre quadrants, mais nous n'aborderons pas ici les autres aspects car ils n'ont pas été exploités dans le cadre de cette thèse.

Dans la suite de ce chapitre, nous présentons ce que l'on pourrait appeler une « approche intégrale pour les SMA », notre objectif étant de définir un nouveau cadre conceptuel qui permettent d'intégrer de manière cohérente les concepts d'agent, d'environnement, d'organisation et d'institution.

3.2 Pourquoi une approche intégrale pour les systèmes multi-agents ?

Les SMA sont des systèmes sociaux pour lesquels on retrouve intrinsèquement les perspectives interne et externe, individuelle et collective, une approche intégrale permet donc de considérer tous ces aspects d'un SMA. Relativement à la terminologie usuelle pour les SMA, une approche intégrale permet de considérer à la fois les aspects agent, environnement, organisation et institution, et cela dans un même cadre conceptuel.

Plus précisément, nous pouvons dégager les intérêts suivants à considérer une « vision intégrale des SMA » :

1. **Situer les travaux existants** : si l'on considère les quadrants comme une carte, tout modèle peut être « situé » sur cette carte. Ce qui permet de mettre en évidence les perspectives qui sont adressées par un modèle et celles qui ne le sont pas. Cette vision permet donc d'appréhender beaucoup plus clairement le champ d'application d'un modèle.
2. **Un cadre d'intégration** : tout en situant les modèles relativement aux quadrants, cette approche permet aussi de mieux comprendre comment relier des modèles qui apportent des solutions sur des plans différents. La carte intégrale permet de réconcilier des approches complémentaires, comme par exemple celles qui se focalisent sur les états internes d'un agent avec celles qui se focalisent sur la définition d'organisations multi-agents. La carte intégrale ne sert pas à dire quel modèle est plus intéressant que tel autre, mais elle permet de mettre en relation différents modèles pour atteindre une solution plus complète.
3. **Un cadre d'analyse** : quelque soit le problème spécifique qui soit considéré pour un SMA, une approche intégrale préconise de mettre en œuvre des solutions dans la perspective des quatre quadrants. Cette démarche permet de compléter une approche partielle du point de vue des quadrants en la reliant à d'autres approches qui couvrent le ou les quadrants manquants.
4. **Une approche de conception** : dans l'esprit de l'approche intégrale, nous proposons une approche de conception pour les SMA qui repose avant tout sur une distinction et une représentation explicite des dimensions qui correspondent aux quadrants.

3.3 Une carte intégrale pour les systèmes multi-agents

Dans cette section nous allons voir en quoi les quadrants définissent une carte sur laquelle il est possible de situer les différentes notions et modèles qui ont été proposés pour les SMA. Tout d'abord reconsidérons suivant les quadrants les principaux concepts qui ont été présentés dans le chapitre 2 : agent, environnement, interaction, organisation, norme, engagement social et institution.

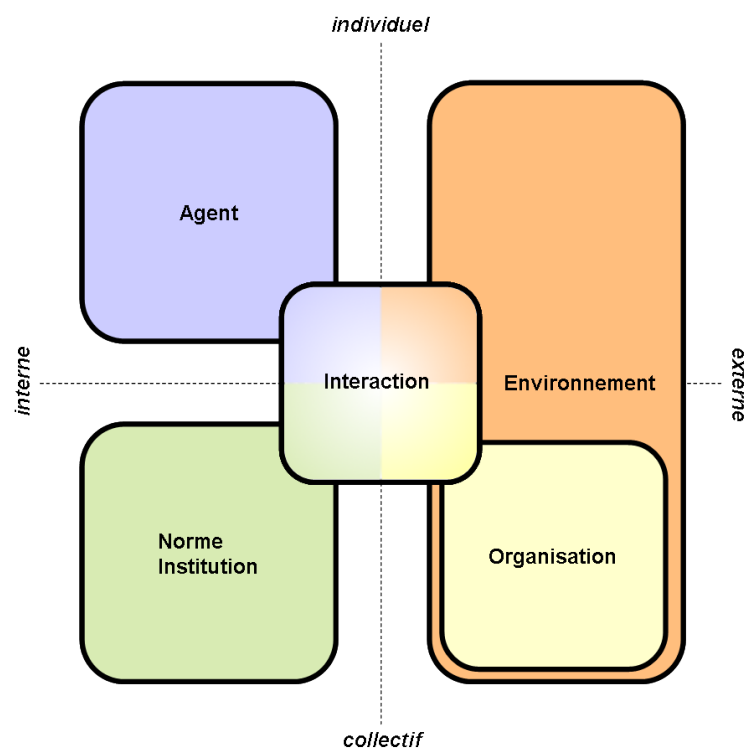


FIG. 3.3 – Les concepts fondamentaux d'un SMA situés sur les quadrants

Agent

C'est au niveau de la structure interne d'un agent que l'on conçoit tout ce qui lui est nécessaire pour prendre des décisions et déterminer son comportement. Par définition la structure interne d'un agent contient tous ses états internes. C'est ici que s'exprime le vécu d'un agent. Dans le cas d'un agent cognitif, sa structure interne contient ses croyances, ses intentions, ses buts, etc. Pour un agent réactif, il n'y a pas de traitement symbolique, mais le comportement est aussi déterminé au niveau de sa structure interne. Dans un SMA, la structure interne n'est (ou ne devrait être) accessible que pour l'agent lui-même.

La structure interne d'un agent correspond donc évidemment au quadrant I-I.

Environnement

Dans un SMA, l'environnement établit et maintient une réalité commune pour l'ensemble des agents et qui est objective de leur point de vue. L'environnement fournit le support même de l'activité des agents : c'est dans l'environnement que les agents peuvent manifester leur comportement. L'environnement est aussi une structure qui met en relation l'ensemble des objets et des agents du système.

L'environnement correspond donc aux deux quadrants externes (E-I et E-C).

Organisation

Les organisations ont été proposées dans le cadre des SMA pour permettre une structuration du système au travers des notions de *groupe* et de *rôle*, avec la motivation d'éliminer les états mentaux de cette structuration. Elles établissent des relations entre rôles qui sont de nature objective, et donc externe. Par exemple, dans le modèle AGR, la structure de groupe et de rôle détermine qui peut communiquer avec qui.

Les organisations se situent donc au niveau du quadrant E-C.

Norme, engagement social, institution

Tous les aspects normatifs, qu'ils s'agissent de norme, d'engagement social ou d'institution, ont été étudiés dans le cadre des SMA pour permettre un certain équilibre entre le besoin de contrôle pour le système ou une sous partie du système, et le besoin de préserver l'autonomie individuelle des agents. Ces notions ont aussi été étudiées pour faciliter la coordination entre des agents hétérogènes, notamment dans un système ouvert. Toutes ces motivations peuvent se résumer à une seule : établir une pression sociale de la part d'un groupe ou d'une société sur ses membres tout en préservant leur autonomie. Il s'agit bien ici de définir des interprétations subjectives de l'activité dans le système qui sont partagées par un groupe d'agents.

Les aspects normatifs d'un SMA se situent donc dans le quadrant I-C.

Interaction

Pour qu'il y ait une interaction, sous quelque forme que ce soit, il faut tout d'abord qu'il y ait des actions individuelles. Et pour qu'un agent agisse, il faut qu'il en ait l'intention. De plus la possibilité d'interaction dépend des structures qui sont définies dans le système, notamment par les organisations. Et enfin, les interactions peuvent avoir un sens qui leur est attribué dans une certaine culture. La culture peut même constituer une motivation pour produire certaines interactions spécifiques.

Pour toutes ces raisons, l'analyse de l'interaction dans un SMA concerne l'ensemble des quadrants : elle dépend aussi bien de la structure interne des agents, que des propriétés de l'environnement, que des structures organisationnelles ainsi que des aspects normatifs.

La figure 3.3 récapitule la position dans les quadrants de chacun des concepts évoqués pour les SMA, elle représente la « carte des SMA ».

Distinguer système social et culture

Nous avons présenté à la section 2.4 les approches sociales pour les SMA. En adoptant un point de vue collectif, ces approches ont donc abordés des problématiques relatives aux deux quadrants collectifs (I-C et E-C). A notre connaissance, il n'y a pas dans les approches existantes que l'on peut qualifier de « sociales » de séparation explicite entre la perspective interne et la perspective externe d'un collectif. Par exemple, dans l'approche VOYELLES [Demazeau, 2001] la dimension « organisation » comprends des éléments internes comme la régulation du système et des éléments externes comme la structuration du système.

Les quadrants nous permettent de mettre en évidence cette confusion qui est faite dans le traitement des problématiques sociales dans les SMA. C'est la distinction nette que fait Wilber entre un système social et une culture.

Un système social, même si il est considéré comme décrété, est un réseau externe ; il établit une structuration et des relations entre les éléments du système qui ont une manifestation dans l'environnement. Un système de messagerie instantanée est un exemple de système social ; dans un tel système, les utilisateurs sont mis en relation via la notion de *contact* ou de *carnet d'adresse*, un utilisateur donné ne pouvant communiquer qu'avec ses contacts. Cette relation est donc objective et s'impose aux utilisateurs du système.

Une culture quant à elle établit un ensemble de valeurs, de conventions et d'interprétations ; autant d'éléments qui sont de nature subjective et partagés par un groupe d'agents. Alors qu'un système social s'impose aux agents, une culture peut être acceptée ou refusée. De plus, une culture n'a pas d'impact direct sur l'environnement, mais elle peut en avoir sur les états internes des agents. Dans le cas d'un système de client de messagerie instantanée, une règle telle que « il est préférable de ne pas communiquer avec

un utilisateur qui affiche le statut occupé » est de nature culturelle car elle ne modifie pas les capacités objectives des utilisateurs.

Le concept de rôle est central dans les travaux sur les organisations. Le rôle est une abstraction qui est employé pour décrire à la fois l'aspect structurel d'une organisation (externe) et le comportement attendu d'un agent qui joue un rôle particulier (interne).

Pour illustrer notre propos, considérons le modèle AGR et son implémentation avec la plateforme MadKit. Dans MadKit/AGR la structure de groupe et de rôle contraint la communication entre les agents : pour pouvoir communiquer, deux agents doivent jouer un rôle chacun dans un groupe commun. Les rôles sont aussi utilisé comme abstraction pour la communication ou plus généralement l'interaction : un agent peut s'adresser à un autre agent parce que celui-ci joue un rôle spécifique. Ces deux caractéristiques des rôles sont bien objectives : elles permettent ou non à un agent de communiquer avec un autre, et elles permettent à un agent d'entrer en contact avec un autre agent dont il ne connaît pas nécessairement l'identité.

D'autre part, les rôles dans MadKit/AGR sont aussi employés pour supporter la définition de comportements attendus par le biais de normes attachées à un rôle [Tranier *et al.*, 2003]. Les normes attachées à un rôle s'appliquent aux agents qui jouent ce rôle. Ces normes ne sont qu'une description du comportement qui est attendu et par essence elles peuvent être transgressées, elles ne contraignent donc pas de manière absolue l'activité des agents. Il s'agit bien d'un aspect subjectif du rôle.

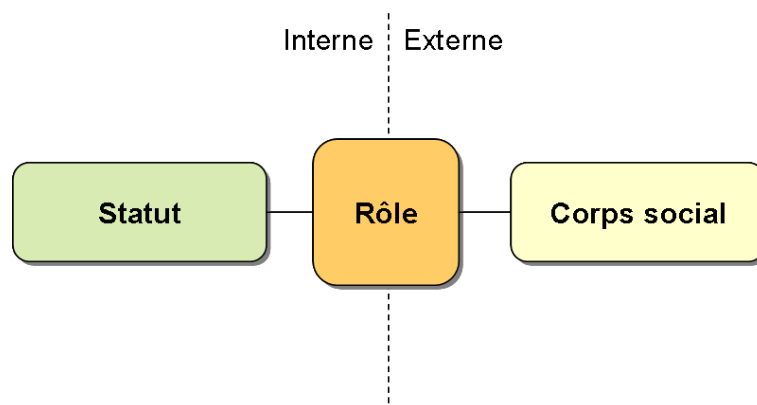


FIG. 3.4 – Un rôle est l'association entre un corps social et un statut

Bien évidemment il y a un lien très étroit entre le quadrant E-C et le quadrant I-C. La structure sociale sert de support à la définition de comportements attendus décrits par des normes. Pour lever l'ambiguïté qui est attaché à la notion de rôle qui peut être employée dans les deux quadrants collectifs, nous utilisons la notion de *corps social* pour le quadrant E-C et celle de *statut* pour le quadrant I-C. Le fait qu'un agent possède un

certain corps social est un fait qui est matérialisé dans l'environnement et qui peut donc être perçu ou exploité dans l'interaction, alors que le fait qu'un agent possède un certain statut doit être accepté ou reconnu, cela ne modifie pas directement l'environnement mais peut changer l'attitude des agents.

Le rôle tel qu'il a été proposé dans la littérature correspond de notre point de vue à une association entre un corps social et un statut (figure 3.4), cette association pouvant être capturée par l'opérateur *counts-as* de Searle : la possession d'un corps social est alors un fait brut qui « compte pour » l'attribution d'un statut. Le corps social confère à son détenteur différentes capacités sociales (comme celle de communiquer par exemple) et désigne un élément dans une structure sociale, alors que le statut correspond à une position qui est reconnue dans une culture et qui est assortie d'un ensemble de normes.

Unification de l'activité physique et de l'activité sociale

Nous situons donc la notion d'organisation en tant que structure dans le quadrant E-C et donc dans l'environnement. Une organisation constitue un moyen objectif d'établir un cadre pour l'interaction dans l'environnement. Les modèles AGRE [Ferber *et al.*, 2004] puis AGREEN [Báez-Barranco *et al.*, 2006] ont déjà fourni une étude de la relation entre une organisation et l'environnement dans laquelle les auteurs proposent d'unifier les notions de corps physique et de corps social. Le comportement d'un agent est une manifestation externe, c'est-à-dire perceptible dans l'environnement, qu'il s'agisse de déplacer un objet ou bien d'émettre un message.

Situer des modèles sur les quadrants

De la même manière que nous avons situé les concepts fondamentaux d'un SMA sur les quadrants, il est possible d'y situer tout modèle qui s'applique aux SMA.

Par exemple nous pouvons situer l'architecture BDI au niveau du quadrant I-I car elle définit une manière d'établir et de maintenir la dimension interne d'un agent. Par contre cette architecture ne dit pas comment construire un environnement, une organisation ou une culture.

Pour prendre un autre exemple situé dans un quadrant différent, le modèle AGR permet de construire des organisations en termes de groupes et de rôles, il se situe donc dans le quadrant E-C. Les capacités individuelles qui sont prévues par AGR sont réduites, elles ne concernent que l'envoi de message, la création de groupe et l'acquisition de rôle (quadrant E-I) ; AGR considère l'existence des agents mais il ne présuppose aucun modèle d'agent spécifique et cette partie doit donc être couverte par d'autres modèles (quadrant I-I) ; et de même le modèle AGR considère la possibilité d'attacher des normes aux rôles

pour décrire des comportements attendus, mais ici aussi rien n'est dit sur l'expression et le fonctionnement des normes, un autre modèle traitant des normes doit donc être employé (quadrant I-C).

Enfin considérons un troisième exemple, ISLANDER (cf. section 2.4.6), qui nous permet d'illustrer l'intérêt de cette carte pour mettre en évidence les aspects traités par un modèle. Avec ISLANDER, une institution est décrite à partir de quatre éléments de base : *cadre dialogique*, *scene*, *structure performative* et *norme*. Considérons ces notions par rapport aux quadrants. Les cadres dialogiques, les scènes et les structures performatives décrivent les schémas de conversation qui sont possibles pour les agents. Cet aspect s'impose aux agents, ils sont objectifs de leur point de vue, et il convient donc de les placer dans les quadrants externes. Les normes quant à elles sont décrites comme un moyen d'imposer des obligations à un agent (quadrant I-C) et de modifier ses possibilités dans la structure performative (quadrant E-I).

Le sens donné à l'institution pour ISLANDER est différent de celui qui est donné par Searle (qui se trouve dans le quadrant I-C). L'intérêt de la carte que nous proposons est alors de dépasser la terminologie pour réduire l'ambiguïté quand il s'agit de comparer des approches.

Aussi bien l'architecture BDI, que le modèle AGR, et que l'outil ISLANDER apportent des solutions pour construire un SMA, mais aucun de ces modèles ne permet de construire complètement un SMA, il est nécessaire de les relier à d'autres modèles qui couvrent les aspects qu'ils ne couvrent pas. L'intérêt de la carte intégrale est de pouvoir situer les différents modèles afin de pouvoir plus facilement les mettre en relation, et finalement de pouvoir couvrir tous les quadrants lors de la conception d'un SMA. Comme nous venons de le voir sur les exemples utilisés, un modèle n'est pas nécessairement situé précisément dans un quadrant, il peut couvrir de manière plus ou moins étendue plusieurs quadrants.

3.4 Analyser au travers des 4 quadrants

Une approche intégrale, dans le sens donné par Wilber, consiste à considérer un problème suivant les quatre quadrants. Dans cette section nous allons illustrer comment une approche intégrale peut guider l'analyse de problèmes avec une approche multi-agent en prenant comme exemple le cas de la communication et celui du contrôle social pour les SMA.

3.4.1 La communication

La communication a été initialement traitée comme un échange direct de messages entre les agents, et la sémantique des langages de communication tel que KQML [Finin *et al.*, 1994] ou FIFA-ACL [FIPA, 2001] est définie en termes d'états mentaux des agents. Une telle approche de la communication est une réduction à un seul quadrant : le quadrant I-I.

Plus récemment cette approche de la communication a été critiquée, et certains travaux, dont [Fornara *et al.*, 2007], ont proposé de donner une sémantique institutionnelle aux langages de communication. Cette nouvelle considération de la communication intègre le quadrant de la culture (quadrant I-C). Une culture permet de définir une ontologie qui est partagée par tous les membres d'une société ainsi que des règles institutionnelles qui engendrent la sémantique du langage. La sémantique d'une affirmation, par exemple, n'est plus de modifier directement les croyances de l'interlocuteur, mais d'établir un engagement social pour celui qui fait l'affirmation : affirmer qu'un énoncé est vrai, c'est s'engager sur la véracité de cet énoncé. Cette approche permet de donner une sémantique culturelle au langage qui est nécessaire pour qu'une compréhension mutuelle soit possible. Elle permet aussi de préserver l'hétérogénéité des agents en ne posant pas de contrainte sur leur architecture interne.

Les deux quadrants internes concernent la signification des messages échangés et les effets produits sur les interlocuteurs par ces messages. Il faut aussi considérer l'aspect externe de la communication, c'est-à-dire les moyens de communication qui sont offerts aux agents par l'environnement. A ce niveau on considère la forme des messages qui sont supportés par le système (un message est au moins une suite de bits) et les mécanismes de transfert des messages. Indépendamment de la sémantique, l'environnement peut établir toutes les contraintes liées à la communication dans le système, comme : qui peut communiquer avec qui ? Quand ? Sous quel mode ? Dans quelles conditions ? Etc.

3.4.2 Le contrôle social

Le contrôle social désigne l'ensemble des moyens et des ressources par lesquels une société parvient à faire respecter ses normes. Le but d'un contrôle social est de permettre un fonctionnement correct de la société en réduisant les conflits.

Ce principe retient beaucoup d'attention dans la communauté multi-agent pour fournir des réponses au problème de la dualité entre l'autonomie et le contrôle. Le contrôle social permet à une société d'agents d'exercer une pression sociale sur chacun d'eux. Ils sont alors influencés dans leur comportement, par le biais de récompenses ou de sanctions, mais leur autonomie est préservée car ils conservent la liberté de choisir quelles actions effectuer même si celles-ci enfreignent les contraintes sociales.

C'est pour rendre possible des formes de contrôle social dans un SMA qu'ont été étudiées des notions comme les organisations, les normes ou encore les institutions. Bien évidemment ces notions sont toutes nécessaires pour qu'un contrôle social soit possible, mais ce que nous apporte la vision intégrale appliquée à ce problème, c'est que les autres quadrants ne doivent pas être négligés.

Dans cette thèse, nous cherchons non seulement à analyser les aspects normatifs d'un SMA en eux-mêmes, mais aussi les relations qu'ils entretiennent avec les autres aspects du système. Il s'agit aussi d'établir ce qui relève des aspects normatifs dans le système et ce qui au contraire doit être traité par d'autres moyens.

Si l'on considère le problème du contrôle social du point de vue du quadrant I-I on peut dégager différents besoins. Tout d'abord le contrôle social sous-entend une forme de contrôle effectuée par les agents eux-mêmes. Il est donc nécessaire de placer des agents dans le système qui seront en charge d'effectuer ce contrôle, que cette tâche incombe à certains agents désignés ou bien à l'ensemble des agents. Dans tous les cas, les agents en charge du contrôle doivent être en mesure de déterminer les cas d'infraction au vu des faits et des normes en vigueur dans le système. D'autre part, si l'on veut que les agents soient influencés par les contraintes sociales, il faut qu'ils disposent des capacités cognitives leur permettant d'intégrer ces contraintes sociales dans leur raisonnement.

Nous pouvons citer comme travaux qui se sont intéressés au contrôle social du point de vue du quadrant I-I l'extension de l'architecture BDI pour prendre en compte les normes (cf. section 2.4.3), et tous les travaux sur les logiques déontiques qui proposent des mécanismes de raisonnement sur les normes (cf. section 2.4.2).

Le point de vue du quadrant E-I est moins évident à première vue dans le cas du contrôle social, mais il est tout de même important. Avant tout, pour qu'un contrôle so-

cial soit possible, il est nécessaire que le comportement des agents soit observable. C'est l'environnement qui permet de donner une perspective publique aux actions réalisées par les agents en préservant les traces du comportement des agents. On peut citer les travaux de [Kaminka *et al.*, 2002] qui ont proposé le principe d'*écoute flottante* pour surveiller de manière non intrusive l'activité d'agents autonomes. Ensuite, les moyens « physiques » d'observation qui sont donnés aux agents contrôleurs ont nécessairement un impact sur le contrôle social. Enfin, pour qu'un contrôle social soit réellement coercitif, il est nécessaire de pouvoir appliquer en dernier recours des sanctions physiques. Ici encore les capacités matérielles qui sont données aux agents pour appliquer des sanctions influencent directement la pression qui est mise sur les membres d'une société, et donc le résultat effectif du contrôle social mis en place.

3.5 Approche de conception pour les systèmes multi-agents

Nous avons vu l'intérêt de la vision intégrale en ce qui concerne l'analyse et la compréhension d'un SMA, nous allons maintenant la considérer dans le cadre de la conception d'un SMA.

La théorie de la vision intégrale a été développée par Wilber non pas pour déterminer ce que serait la réalité de notre monde, mais pour en donner une meilleure compréhension, ce que Wilber exprime par la métaphore suivante : la vision intégrale est une carte qu'il ne faut pas confondre avec le territoire. Si l'on considère maintenant un SMA artificiel, les choses sont différentes. En poursuivant la métaphore de Wilber, la conception d'un SMA peut être assimilée à la création d'un territoire. Ce territoire constitue alors la « réalité » pour les agents du système.

Notre proposition est d'adopter une approche de conception pour les SMA qui reflète en elle-même les quatre quadrants. Cette approche se veut très générale pour permettre l'intégration de différents modèles spécifiques. Elle repose sur une conception explicite et séparée de chacune des quatre dimensions qui correspondent aux quadrants.

3.5.1 Séparer les quatre dimensions

Pour concevoir séparément les quatre dimensions d'un SMA, il faut extraire et traiter explicitement les relations qu'elles entretiennent entre elles. Nous allons donc maintenant considérer les distinctions qui sont nécessaires à la séparation de ces quatre dimensions.

Distinguer la volonté d'agir de l'action

Si l'on considère la carte intégrale, la volonté d'agir et l'exécution d'une action se situe dans deux quadrants différents. La volonté d'agir trouve son origine à l'intérieur d'un agent. Pour un agent cognitif, l'intention de réaliser une action résulte de ses états mentaux comme ses croyances et ses buts, pour un agent purement réactif elle est le résultat d'un processus déterministe qui ne prend en compte que les percepts. Ainsi quel que soit le type d'agent, c'est bien au niveau de l'architecture interne de l'agent qu'est produite une volonté d'agir (quadrant I-I).

L'exécution d'une action, quant à elle, correspond à une transformation de l'environnement. L'ensemble des actions qu'exécute un agent est la manifestation de son comportement. Les actions qu'un agent peut réaliser et les conséquences de ces actions sont des aspects qui relèvent de l'environnement, et plus particulièrement des lois qui régissent l'environnement (quadrants externes).

Cette distinction a déjà été étudiée dans la littérature sous la dénomination séparation *esprit/corps* (cf. section 2.2.5). Cette séparation est rendue possible par l'application du principe *influence/réaction* (cf. section 2.2.4). Elle permet de préserver à la fois l'intégrité des agents et celle de l'environnement. Il s'agit bien ici de distinguer deux dimensions, l'une étant gouvernée par l'autonomie de l'agent, l'autre part les lois de l'environnement.

Distinguer les structures individuelles des structures collectives

Faire la distinction entre les structures individuelles et les structures collectives revient à décrire séparément le niveau micro et le niveau macro. Cela permet aussi de définir séparément les comportements individuels possibles des interactions qui peuvent résulter de ces comportements.

Un SMA est un système dont les éléments sont des agents et des objets. Il est tout aussi nécessaire de décrire individuellement chacune des entités que de décrire les relations (physiques et sociales) qu'elles entretiennent.

Les travaux sur les organisations (voir section 2.4.1) ont proposé des solutions pour décrire des structures collectives, et cela indépendamment des états internes des agents qui vont intégrer ces structures.

La notion d'environnement dynamique proposée par [Helleboogh *et al.*, 2007] (cf. section 2.2.6) permet de décrire explicitement la dynamique individuelle de chacune des entités de l'environnement avec la notion d'activité ainsi que les interférences qui sont produites par la mise en commun de ces activités individuelles.

Distinguer l'interprétation individuelle de l'interprétation collective

La subjectivité (le quadrant I-I) est présente dans tout SMA au travers des agents : chaque agent peut se faire sa propre interprétation de l'activité dans l'environnement. Ce qui par contre n'est pas toujours traité dans les SMA, c'est l'intersubjectivité. Les travaux qui se sont intéressés aux aspects normatifs pour les SMA (quadrant I-C) ont motivé le besoin d'introduire une dimension institutionnelle, culturelle ou intersubjective pour renforcer la cohésion sociale et la compréhension mutuelle entre les agents, et ainsi favoriser la coordination. Le concept d'institution tel qu'il a été défini par Searle permet de décrire des réalités institutionnelles, qui sont de nature collective, c'est-à-dire indépendantes des états internes des individus.

En pratique le quadrant I-C a souvent été réduit au quadrant I-I. Les informations relatives à une culture, comme une ontologie par exemple, sont alors encodées dans la structure interne des agents. L'inconvénient est que la culture est alors imposée aux agents. Cela réduit l'hétérogénéité des agents et va à l'encontre de la conception de systèmes ouverts.

Distinguer réalité brute et réalité culturelle

Nous reprenons ici la terminologie de Searle : la réalité brute désignant l'ensemble des éléments externes (quadrants E-I et E-C) et une réalité culturelle¹ désignant l'interprétation de la réalité brute qui peut être construite par une culture.

Il est primordial de distinguer la réalité brute des différentes réalités culturelles qui peuvent exister car leurs impacts sur le système sont de nature très différente. La réalité brute s'impose aux agents alors qu'une réalité culturelle peut être rejetée. Searle illustre cette distinction en comparant une palissade avec une frontière arbitraire. Une palissade peut, par ses propriétés physiques empêcher le passage. Un agent situé d'un côté de la palissade n'a pas la possibilité physique de se rendre de l'autre côté. Dans une telle situation l'agent n'a pas le choix de passer ou non de l'autre côté, la réalité brute le contraint à rester du côté où il se trouve. Maintenant si l'on considère une frontière arbitraire, telles celles qui sont établies entre les pays, la frontière en elle-même n'empêche pas le passage. Cependant celui-ci peut être interdit. La frontière peut remplir la même fonction qu'une palissade, à la condition que les agents acceptent² collectivement l'interdiction qui est associée à la frontière. Dans cette situation, un agent peut d'après

¹Searle ne parle pas de réalité culturelle mais de réalité institutionnelle. Nous considérons des réalités institutionnelles comme des éléments parmi d'autres d'une réalité culturelle.

²Il faut faire attention ici au sens donné à « accepter ». Il se peut que les agents ne respectent l'interdiction de passage que par crainte des sanctions encourues. Mais dans ce cas, le résultat est bien là, ils reconnaissent la fonction de la frontière.

sa rationalité, décider de respecter ou de transgresser l'interdiction.

L'environnement naturel dans lequel nous vivons s'impose à nous. Nous pouvons en changer l'état, mais nous ne pouvons transgresser les lois de la physique qui le régissent. Toutes les constructions humaines qui ne relèvent pas des lois de la physique sont donc nécessairement institutionnelles ou culturelles. Par contre, un SMA artificiel est lui, entièrement construit. Le choix de modéliser telle ou telle chose comme partie de la réalité brute ou de la réalité culturelle est un choix arbitraire qui est fait par le ou les concepteurs du système. Rien n'empêche de modéliser dans un système des éléments de notre culture comme étant partie intégrante de la réalité brute du système.

Le choix de modéliser tel ou tel aspect au niveau de la réalité brute ou de la réalité culturelle dépend des besoins du système considéré. La réalité brute permet d'établir certaines garanties quant à ce qui pourra ou ne pourra pas se passer dans le système. La réalité brute est donc essentielle pour maintenir la sécurité du système, elle offre des possibilités pour protéger tous les aspects sensibles d'un système. Mais il y a aussi un inconvénient à la modélisation au niveau de la réalité brute, c'est que les règles qui sont ici décrites sont figées. La réalité brute permet d'apporter des solutions dans les situations où l'on est en mesure de prévoir à l'avance le résultat attendu, mais pour les autres cas, ce n'est pas à ce niveau que l'on permet au système de s'adapter. La capacité d'adaptation d'un système provient de l'autonomie de ses constituants. Si pour un problème donné on souhaite obtenir plus de flexibilité et d'adaptation, c'est au niveau de la culture qu'il faudra alors travailler. La culture permet de préconiser des choix, mais fait toujours intervenir en dernier recours le libre arbitre de l'agent. Et c'est la rationalité de l'agent appliquée à son activité locale qui permettra au système de s'adapter.

3.5.2 Les relations entre les quatre dimensions

Nous venons de dire que les quatre dimensions sont toutes distinctes mais interdépendantes. Nous présentons sur la figure 3.5 les relations qui existent entre elles du point de vue des SMA.

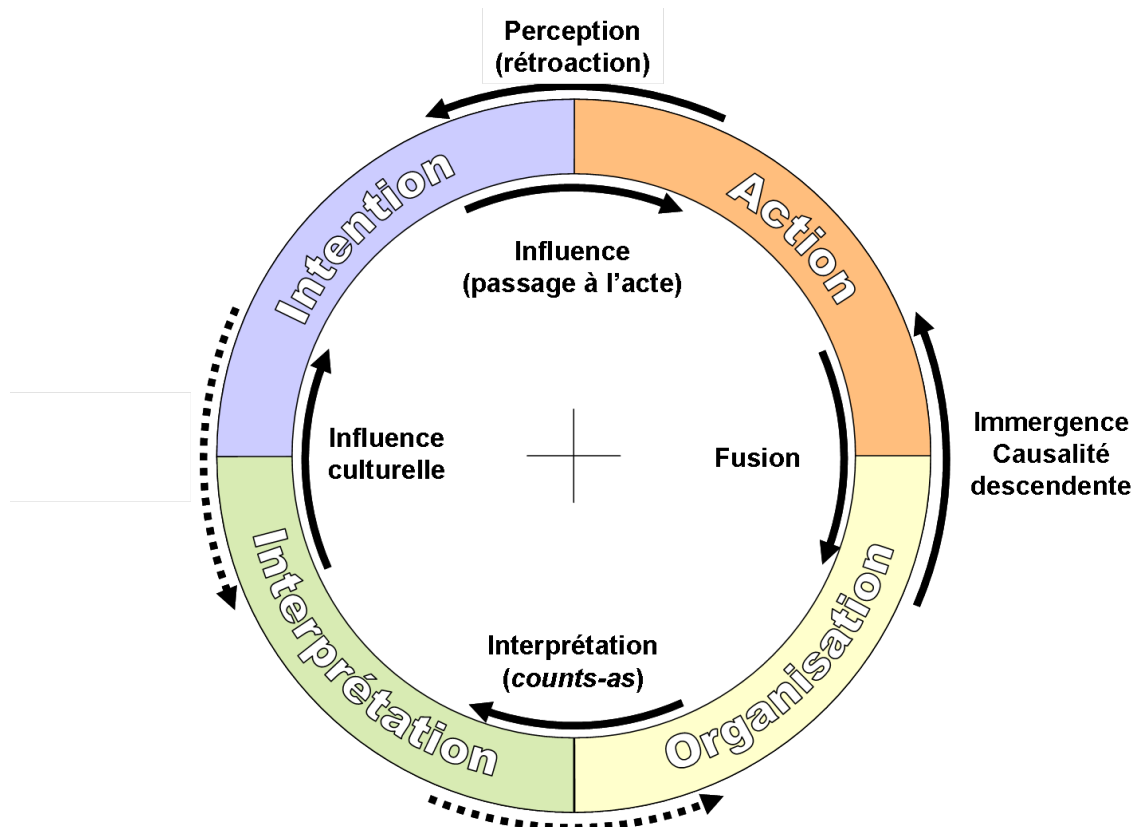


FIG. 3.5 – Relations entre les quadrants

Les relations entre le quadrant I-I et le quadrant E-I expriment le rapport entre la structure interne des agents et l'environnement. L'activité interne d'un agent engendre ses intentions qui le conduisent à agir dans l'environnement. Le passage de la dimension I-I à la E-I pour un agent correspond au *passage à l'acte*. Suivant le principe *influence/réaction*, cela correspond à l'émission d'influences de l'agent vers l'environnement. Toujours d'après ce principe, un agent ne peut pas déterminer lui-même les conséquences de ses actes et il ne peut pas avoir accès à l'état de l'environnement, il ne peut donc que constater au travers de sa perception le résultat de ses actes et l'évolution de l'environnement. L'environnement émet en retour des percepts à l'agent, qu'il va pouvoir intégrer dans sa structure interne, ce qui constitue alors une *boucle de rétroaction*.

Les relations entre les quadrants E-I et E-C expriment le rapport entre le niveau micro et le niveau macro. D'après le quadrant E-I, on considère le comportement individuel de chaque entité du système, et d'après le quadrant E-C, on considère les structures

collectives dans lesquelles les comportements individuels ont lieu. Le passage du niveau micro au niveau macro correspond à la *fusion* de l'individu avec la structure englobante, c'est-à-dire comment l'individu est relié au reste du système. La relation réciproque est la contrainte appliquée par les structures collectives sur les éléments qui les constituent, c'est l'*immersion* ou la *causalité descendante* (*downward causation*).

La relation entre le quadrant E-C et le quadrant I-C est à sens unique : la culture attribue une interprétation collective à la réalité brute mais la culture n'a pas d'influence directe sur la réalité brute. Du point de vue externe, tout est factuel et objectif, les choses sont ou ne sont pas ; du point de vue interne, la culture attribue des jugements de valeur, des statuts ou des états déontiques aux phénomènes. Les faits bruts engendrent les interprétations collectives, mais ces interprétations ne changent pas les faits. Cependant, il y a bien une relation entre la culture et la réalité brute, mais cette relation est indirecte (elle est représentée en pointillé sur la figure). La culture influence les agents dans leur interprétation individuelle, elle peut donc les conduire à agir différemment, et donc transformer différemment l'environnement. L'impact de la culture sur l'environnement dépend donc de l'impact de la culture sur les individus.

La relation entre le quadrant I-C et le quadrant I-I dans ce sens correspond à l'*influence culturelle*, c'est-à-dire comment une culture influence un individu dans ses choix, et donc l'amène à agir différemment. La relation réciproque correspond à l'évolution de la culture. Nous considérons que cette relation est indirecte, car si les agents sont responsables de l'évolution d'une culture, cette évolution ne peut être accomplie par la simple pensée. La communication ou du moins l'interaction entre les membres d'une société est nécessaire pour faire évoluer une culture. L'évolution d'une culture est un problème très complexe qui ne sera pas abordé dans cette thèse et qui constitue une des perspectives ouvertes par celle-ci.

3.5.3 Quadrants – Comportements – Capacités

Nous venons de voir comment les quatre dimensions d'un SMA pouvaient être conçues séparément, et quelles étaient les relations entre elles. Maintenant nous allons analyser le rapport entre chacune de ces dimensions et les notions de *comportement* et de *capacité*. Nous employons ici le terme capacité pour désigner une aptitude d'un agent à modifier une dimension du système.

La figure 3.7 montre la relation entre les quadrants et la notion de capacité.

Pour le quadrant I-I, nous parlerons des capacités cognitives d'un agent qui lui permettent de modifier son état interne. Ces capacités peuvent être plus ou moins sophistiquées pour aller de l'agent réactif le plus simple jusqu'à un agent cognitif complexe.

Pour les quadrants externes (E-I + E-C) nous parlons de capacités matérielles qui

expriment l'aptitude à transformer la réalité brute, c'est-à-dire l'environnement.

Enfin pour le quadrant I-C nous parlons de pouvoirs (institutionnels). Un pouvoir étant l'aptitude d'un agent à transformer une réalité culturelle. Ces pouvoirs sont déterminés par les règles constitutives des institutions.

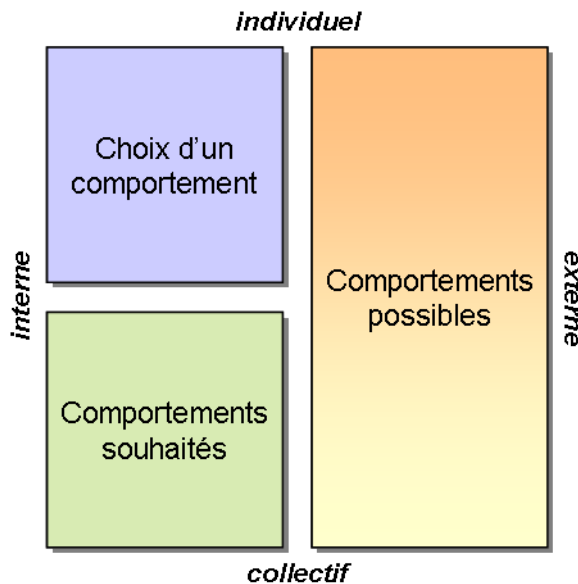


FIG. 3.6 – Quadrants et comportements

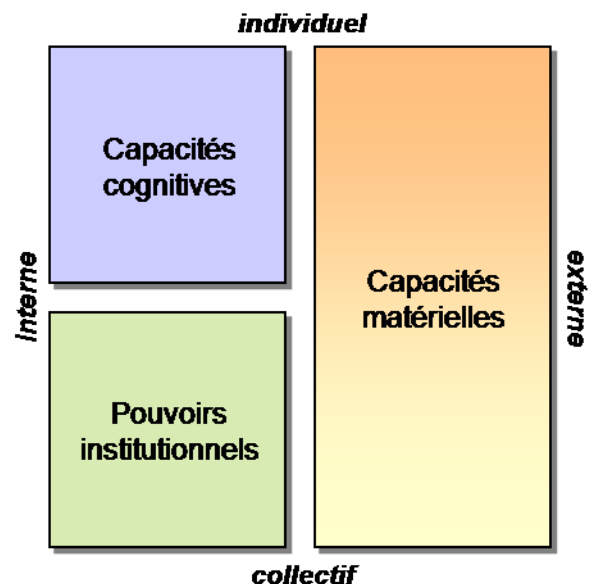


FIG. 3.7 – Quadrants et capacités

La figure 3.6 illustre le rôle de chaque quadrant relativement aux comportements dans un SMA. Les deux dimensions externes (E-I + E-C) induisent tous les comportements qui sont possibles dans le système. Ces comportements possibles dépendent à la fois des capacités individuelles et des structures d'interaction. Cet aspect permet de délimiter l'autonomie des agents dans le système.

La dimension du quadrant I-I, détermine le choix d'un comportement qui est fait par un agent parmi tous les comportements possibles. C'est à ce niveau que se situe l'autonomie d'un agent, il doit être le seul à pouvoir déterminer quel comportement il adopte parmi toutes les possibilités qui lui sont offertes.

Enfin la dimension I-C décrit les comportements qui sont attendus ou au contraire non désirés par un groupe d'agents.

3.5.4 Intérêts de cette approche de conception

La conception séparée des dimensions intentionnelle, comportementale, sociale et culturelle d'un SMA présente de nombreux avantages.

Tout d'abord ces quatre dimensions sont orthogonales et sont toutes nécessaires pour

la conception d'un SMA. Le rôle de chacune de ces dimensions et leurs effets sur le système sont différents. Le fait de les concevoir séparément permet donc d'obtenir un meilleur contrôle sur chacune d'elles.

D'autre part, nous avons vu comment il était possible d'exploiter la vision intégrale comme cadre d'analyse. En adoptant cette approche de conception, il y a un correspondance directe entre le processus d'analyse et celui de conception, ce qui permet bien évidemment une meilleure compréhension de la conception.

Intrinsèquement, cette approche de conception permet de construire des SMA plus modulaires. En reliant les différents éléments du système par un couplage faible, il est plus facile de modifier ou de substituer un élément du système sans pour autant devoir modifier l'intégralité du système. Ainsi l'environnement peut être modifié sans intervenir sur les agents ou au contraire il est possible de modifier les agents sans intervenir sur l'environnement. Il en va de même pour les différentes cultures qui peuvent être conçues indépendamment du reste du système.

Cette approche est très générale, ce qui lui permet d'exploiter différents modèles spécifiques en pratique. Elle constitue un cadre intégrateur dans lequel les modèles d'agents, d'environnement, d'organisation, de normes ou d'institutions peuvent plus facilement trouver leur place, et être rassemblés pour concevoir un système dans sa globalité.

Enfin, et c'est le point le plus important, cette approche est particulièrement intéressante pour concevoir des systèmes ouverts. La difficulté que pose la conception des systèmes ouverts est justement que les agents sont conçus séparément du reste du système. Avec cette approche, les relations entre les agents et le reste du système sont beaucoup moins ténues, il est alors beaucoup plus aisé de les concevoir séparément. De plus la prise en compte de la culture permet de réduire considérablement les contraintes qu'un système doit poser sur l'architecture interne de ses agents pour atteindre une compréhension mutuelle entre les agents.

Chapitre 4

MASQ : un méta-modèle fondé sur les 4 quadrants

DANS ce chapitre nous présentons le méta-modèle MASQ (pour *Multi-Agent System based on Quadrants*). MASQ est fondé sur les quatre quadrants de la vision intégrale. Il permet de décrire de manière formelle un système multi-agent suivant chaque quadrant. Il repose sur quatre concepts primitifs : l'*esprit*, l'*objet*, l'*espace brut* et la *culture* ; chacun de ces concepts étant destiné à capturer une des quatre dimensions correspondant aux quadrants. La formalisation de ces concepts a été établie dans un souci de généralisation afin de pouvoir intégrer et mettre en relation des modèles spécifiques d'agent, d'environnement, d'organisation, de norme ou encore d'institution. De plus, MASQ permet de spécifier séparément les agents, l'environnement et les aspects culturels. Cette caractéristique rend ce modèle intéressant pour la modélisation de systèmes ouverts.

Contents

4.1	Introduction	92
4.2	Esprit	95
4.3	Objet	97
4.4	Corps : incarnation d'un esprit dans un objet	100
4.5	Espace brut	105
4.6	Culture	123
4.7	Synthèse	136

4.1 Introduction

L'objectif principal de cette thèse est de définir un cadre général pour les SMA qui permette d'intégrer différents modèles proposés dans des contextes spécifiques (cf. chapitre 2) en les mettant en relation à l'aide du cadre d'analyse proposé par Wilber (cf. chapitre 3). C'est dans cette perspective que nous proposons le méta-modèle MASQ (pour *Multi-Agent System based on Quadrants*) qui fait l'objet de ce chapitre.

MASQ est un méta-modèle formel générique pour la description de SMA. Ce méta-modèle se veut suffisamment générique pour permettre l'exploitation de modèles spécifiques selon les besoins du système considéré.

MASQ s'inscrit dans une continuité initiée par les modèles AGR [Ferber et Gutknecht, 1998], AGRE [Ferber *et al.*, 2004] puis AGREEN [Báez-Barranco *et al.*, 2006]. Ce méta-modèle est à la fois le résultat d'une intégration de différents modèles et principes, et a pour objectif d'établir un cadre intégrateur.

Le méta-modèle MASQ permet de décrire un SMA à partir de :

- quatre concepts primitifs : *esprit*, *objet*, *espace brut* et *culture* ;
- un ensemble de relations entre ces concepts ;
- un ensemble de lois qui permettent de décrire la dynamique du système.

4.1.1 Concepts primitifs

Le méta-modèle MASQ repose sur quatre concepts primitifs : *esprit*, *objet*, *espace brut* et *culture* ; chacun de ces concepts étant destiné à capturer une des quatre dimensions correspondant aux quadrants. Le figure 4.1 présente ces concepts primitifs situés dans les quadrants.

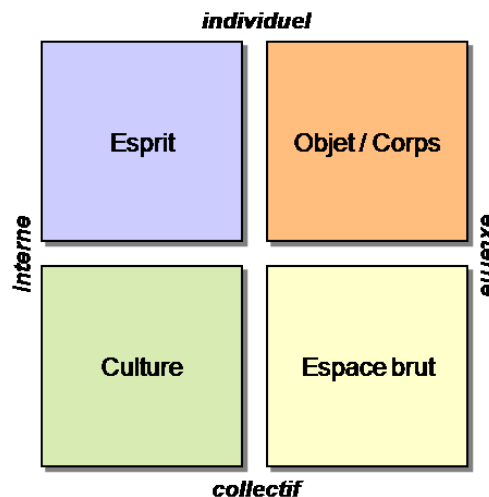


FIG. 4.1 – Les concepts primitifs : *esprit*, *objet*, *espace brut* et *culture*

Avant de présenter en détail chacun de ces concepts, nous en donnons ici une vue d'ensemble :

Esprit : La description d'un esprit est celle de l'architecture interne d'un agent. Le concept d'esprit capture la dimension interne individuelle.

Objet : Un objet est une entité de l'environnement, ou de la réalité brute au sens de Searle. Le concept d'objet permet de décrire de manière individuelle l'état et le comportement que manifeste une entité dans l'environnement. Certains objets sont connectés à un esprit, on parle alors de corps. Le concept d'objet capture la dimension externe individuelle.

Espace brut : Les espaces bruts mettent en relation les objets de l'environnement. Ils définissent les structures nécessaires à l'interaction. Le concept d'espace brut capture la dimension externe collective.

Culture : Une culture contient les éléments subjectifs qui sont partagés par un ensemble d'esprits tels que les interprétations collectives, les normes sociales, les ontologies ou de manière plus générale les connaissances partagées.

Comme nous allons le voir ces concepts sont tous reliés entre eux ; si l'on se place du point de vue d'un agent, un agent raisonne avec son esprit, agit dans un monde brut au travers de son corps, ce monde brut est constitué de différents objets avec lesquels il va pouvoir entrer en interaction, et son interprétation du monde a pour support le contexte culturel dans lequel il est plongé.

4.1.2 Principes

L'objectif du méta-modèle MASQ est de permettre pour un SMA de décrire séparément les quatre dimensions induites par les quadrants. Pour cela nous nous appuyons sur un ensemble de principes qui ont été proposés pour les SMA et qui ont chacun été présenté précédemment. Nous rappelons ici chacun d'eux.

Séparation esprit/corps

Ce principe (cf. section 2.2.5) a pour but de séparer la partie interne d'un agent (son esprit) de sa partie externe (ses corps). Un esprit correspond à la structure interne d'un agent qui comporte tout ce qui est mis en œuvre dans l'agent pour décider de son comportement. Les corps de l'agent déterminent son existence dans l'environnement : ils lui donnent ses capacités d'action et de perception sur l'environnement et ils représentent sa manifestation dans l'environnement qui le rend perceptible pour les autres agents. Les corps d'un agent sont soumis aux contraintes environnementales, ainsi ils déterminent un

cadre à l'autonomie de l'agent dans l'environnement : l'autonomie de l'agent consiste à choisir les actions à effectuer parmi celles que ses corps lui offrent.

Principe d'intégrité de l'agent

L'esprit d'un agent (sa structure interne) n'est pas publique. Aucun élément extérieur à l'agent ne peut y avoir accès, que ce soit l'environnement ou tout autre agent. De l'extérieur, seul le comportement manifesté dans l'environnement permet de raisonner sur un agent. Ce principe permet de préserver l'autonomie et l'hétérogénéité des agents.

Principe influence/réaction

Le principe *influence/réaction* [Ferber et Müller, 1996] (cf. section 2.2.4) permet de distinguer la volonté d'agir d'un agent des effets de ses actions sur l'environnement. Un agent décide lui-même quelles actions il produit dans son environnement, mais c'est l'environnement qui détermine les conséquences de ces actions. Un agent n'a donc pas la possibilité de modifier lui-même son environnement, il peut seulement l'influencer. L'environnement quant à lui réagit aux influences produites par tous les agents pour déterminer son évolution. En appliquant ce principe, on peut décrire les lois qui caractérisent l'évolution de l'environnement en prenant en compte les activités concurrentes des agents ; ces lois s'imposent alors aux agents qui ne peuvent que les subir. Ainsi, tout ce qui n'est pas prévu par l'environnement n'est simplement pas possible pour les agents.

L'association ce principe avec celui d'intégrité de l'agent interdit toute interaction directe entre deux esprits, ce qui signifie qu'il ne peut y avoir de « télépathie » entre deux esprits. L'activité d'un agent n'est possible qu'au travers de l'environnement, en conséquence de quoi tout comportement est observable dans l'environnement.

Distinction entre activité individuelle et interaction

Ce principe propose de représenter explicitement les différentes entités qui composent l'environnement – les corps des agents ainsi que les simples objets – et les activités individuelles que manifestent ces entités, comme par exemple un ballon qui roule ou bien un robot qui se déplace. L'interaction résulte alors de la mise en commun de ces activités individuelles qui peut produire des interférences, comme par exemple la trajectoire d'un ballon qui est déviée suite à une collision avec un obstacle. Ce principe permet de considérer séparément le comportement individuel d'une entité et le résultat de l'interaction de l'ensemble des entités.

Unification de l'activité physique et de l'activité sociale

Que l'activité d'un agent soit physique ou sociale, elle consiste à manifester un comportement dans l'environnement au travers de ses corps. Ainsi une activité sociale telle que l'envoi d'un message sera traitée de la même manière qu'une activité physique telle que le déplacement d'un robot.

Distinction entre réalité brute et culture

Ce principe découle des travaux de Searle sur la construction de la réalité sociale (cf. section 2.4.7). Il consiste à faire une distinction explicite entre ce qui constitue la réalité brute, c'est-à-dire ce qui est objectif du point de vue d'un agent, et les différentes interprétations (collectives) qui peuvent être faites par une société d'agents (la culture). Dans un SMA la réalité brute est décrite par l'environnement, elle fournit le contexte pour les activités des agents, et cette réalité s'impose aux agents. Les différentes interprétations de la réalité brute qui sont données par les cultures peuvent influencer les agents dans leurs décisions, notamment en fonction de la pression sociale exercée par ces sociétés, mais ces interprétations n'ont pas d'incidence directe sur la réalité brute, elles n'imposent aucune contrainte matérielle inviolable pour les agents.

4.2 Esprit

L'esprit correspond à la structure interne d'un agent, c'est-à-dire sa partie décisionnelle. Un esprit est un processus indépendant de l'environnement mais qui est en interaction avec celui-ci : il a son propre cycle de vie et il entretient une communication asynchrone permanente avec l'environnement. Cette communication est un échange de percepts délivrés par l'environnement et d'influences émises par l'esprit. L'esprit produit son « raisonnement » en prenant en compte les percepts délivrés par l'environnement ; le résultat de cette phase de délibération détermine les intentions de l'agent qui se manifestent par l'envoi d'influences à l'environnement.

Un esprit est un *système dynamique ouvert* [Michel, 2004], on peut donc le caractériser par :

- son *état interne* qui est généralement représenté par un ensemble de *variables d'état*.
- son *mécanisme de changement d'état* qui détermine comment les variables d'état évoluent au cours du temps en prenant en compte les percepts qu'il reçoit en entrée. Il faut noter que l'évolution d'un esprit n'est pas synchronisée avec son environnement qui lui délivre ses percepts, ainsi pour un cycle l'entrée peut être constituées de plusieurs percepts produits à des instants différents ou au contraire l'entrée peut ne

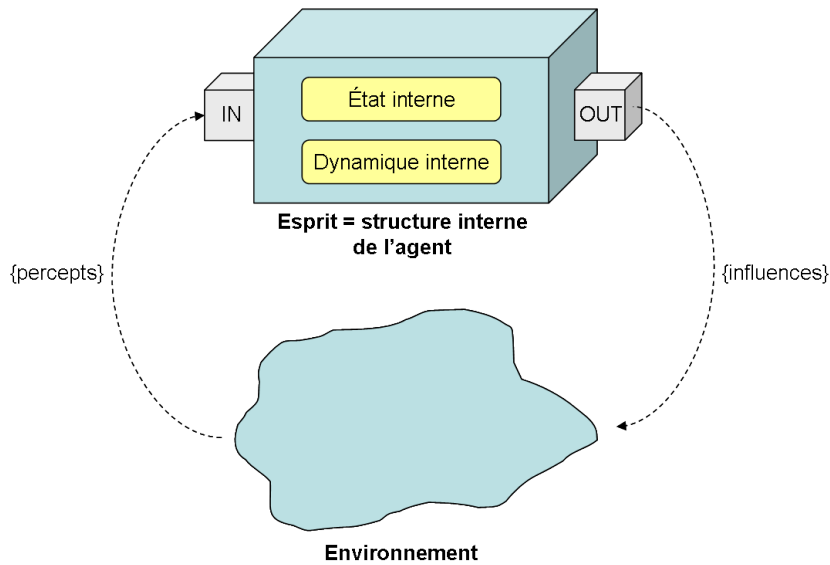


FIG. 4.2 – La relation entre un esprit et son environnement

contenir aucun percept. Ce mécanisme de changement d'état peut être modélisé par une *fonction de transition d'état*.

- son *mécanisme de production* qui détermine les influences produites par l'esprit en fonction de son état interne. Cet aspect peut être modélisé par une *fonction de production*

Les deux derniers points sont généralement regroupés sous le terme de *dynamique interne*.

L'état interne d'un agent modélise sa réalité interne individuelle au sens des quadrants, et sa dynamique interne exprime ses capacités cognitives, c'est-à-dire comment son état interne peut évoluer.

En notant Σ_m l'ensemble des états internes possibles pour un esprit m , Π l'ensemble des percepts possibles, Γ l'ensemble des influences possibles, et $\mathbb{P}(E)$ l'ensemble des parties de l'ensemble E , la dynamique interne d'un esprit est une fonction D_m définie par :

$$D_m : \Sigma_m \times \mathbb{P}(\Pi) \rightarrow \Sigma_m \times \mathbb{P}(\Gamma)$$

$$(\sigma_m, P) \mapsto (\sigma'_m, I)$$

Pour un état interne σ_m et un ensemble de percepts P , la fonction D_m retourne un nouvel état interne σ'_m et un ensemble d'influences I .

Définition 4.1 (Esprit) *Un esprit est un tuple $m = \langle id_m, \sigma_m, D_m \rangle$ avec id_m un identifiant unique, $\sigma_m \in \Sigma_m$ l'état interne de m et D_m sa dynamique interne.*

Définition générique d'un esprit

La définition d'un esprit que nous venons de donner est volontairement très générique, elle permet ainsi de capturer des modèles d'agents très différents, et donc de faire co-exister dans un SMA des agents hétérogènes. La seule contrainte que nous imposons sur un esprit est d'être capable de recevoir des percepts de son environnement et d'émettre des influences vers son environnement.

Nous pouvons aussi remarquer que cette définition ne pose aucune contrainte sur le langage utilisé pour implémenter un esprit et sur le « lieu » d'exécution d'un esprit. Un esprit pourra aussi bien être exécuté sur la même machine que son environnement, que sur une machine distante du moment que l'infrastructure sous-jacente permet la communication entre l'esprit et son environnement.

4.3 Objet

Dans MASQ, le concept d'objet permet de décrire individuellement les entités qui composent l'environnement. Contrairement aux esprits, les objets ne sont ni autonomes ni pro-actifs, leur évolution est entièrement déterminée par les lois de l'environnement et les différentes activités qui apparaissent dans l'environnement. Dans le modèle MASQ, les objets sont des entités passives car leur évolution est contrôlée par l'environnement, cependant ils ne sont pas inertes car ils manifestent des activités comme par exemple un ballon qui roule ou bien un feu de circulation qui change périodiquement de couleur.

Chaque objet est caractérisé à tout instant par son état, mais il faut faire attention au sens particulier que prend la notion d'état comme le montre le paradoxe suivant :

Paradoxe de Zénon

Zénon demandait : « Soit une flèche en vol. A un instant, est-ce qu'elle est au repos ou en mouvement ? ». Si on répondait qu'elle est en mouvement, il disait « Mais être en mouvement, c'est changer de position. A un instant, la flèche a une position, elle n'en change pas. Elle n'est donc pas en mouvement. » Si on répondait qu'elle est au repos, il disait « Mais si elle est au repos à cet instant, elle est aussi au repos à tous les autres instants, elle est donc toujours au repos. Elle n'est jamais en mouvement. Mais alors comment peut-elle passer d'une position à une autre ? »

Pour fournir une réponse à ce paradoxe, nous utilisons la notion d'*état dynamique*. Un état dynamique permet de décrire pour un instant t à la fois l'état d'un objet (ce qu'il est)

par des variables d'état et son activité à cet instant (ce qu'il est entrain de faire) par des variables de dynamique. Sur l'exemple de la flèche de Zénon, la position de la flèche peut être décrite par des variables d'états et son déplacement peut être décrit à un instant t par des variables de dynamique qui représentent sa vitesse de déplacement instantané.

Définition 4.2 (Etat dynamique) *Un état dynamique pour un objet o à un instant t est un tuple $\delta_o(t) = \langle v_1, v_2, \dots, v_n \rangle \in \Delta_o$ avec :*

- Δ_o l'ensemble des états dynamiques possibles pour l'objet o , et
- chaque v_i une variable d'état ou une variable de dynamique.

Dans l'environnement naturel, l'évolution d'un objet peut être déterminée si l'on a une connaissance suffisante des lois de la physique. Dans le cas d'un environnement virtuel, les lois d'évolutions qui le régissent doivent être décrites explicitement. Pour cela, nous introduisons la notion de *loi d'évolution instantanée* qui permet de décrire l'évolution de l'état dynamique d'un objet en l'absence de perturbation extérieure.

Définition 4.3 (Loi d'évolution instantanée) *Une loi d'évolution instantanée Φ est une fonction qui, à un état dynamique δ , associe un nouvel état dynamique δ' :*

$$\begin{aligned} \Phi : \Delta &\rightarrow \Delta \\ \delta &\mapsto \delta' \end{aligned}$$

Bien que notre formalisation soit différente, les notions d'état dynamique et de fonction d'évolution instantanée pour les objets que nous présentons ici peuvent être rapprochées de la notion d'*activité* présentée dans [Helleboogh *et al.*, 2007]. Nous emploierons également ce terme pour désigner ce qu'un objet est entrain de faire.

Il ne faut pas confondre la dynamique interne que nous avons présentée précédemment pour les esprits et la fonction d'évolution instantanée pour les objets. La fonction d'évolution instantanée décrit pour un objet l'évolution de son état en l'absence de perturbation extérieure, c'est-à-dire l'évolution de cet objet si il était la seule entité de l'environnement. Mais comme un objet n'est pas seul dans l'environnement il n'est pas possible de prévoir son état futur uniquement à partir de sa fonction d'évolution instantanée. Par exemple, la flèche peut rencontrer un obstacle durant son vol qui stoppera son mouvement. Ainsi l'évolution d'un objet ne peut pas être déterminée en ne considérant que l'objet lui-même, nous verrons par la suite le rôle que joue l'environnement dans l'évolution des objets qui le composent.

Pour décrire les objets nous utilisons la notion de *type d'objet*, tout objet étant instance d'un certain type d'objet. Un type d'objet est décrit par les états dynamiques possibles pour les objets de ce type, c'est-à-dire l'ensemble des variables nécessaires à la modélisation de l'état dynamique et leur domaine de définition respectif, ainsi que la loi d'évolution instantanée qui s'applique aux objets de ce type.

Définition 4.4 (Type d'objet) *Un type d'objet est un tuple $T = \langle \Delta_T, \Phi_T \rangle$ avec :*

- Δ_T l'ensemble des états dynamiques pour les objets de type T , et
- $\Phi_T : \Delta_T \rightarrow \Delta_T$ la loi d'évolution instantanée pour les objets de type T .

Pour un objet o de type T nous noterons $o \in T$.

Pour permettre une plus grande abstraction dans la description des objets, nous utiliserons les notions de *type abstrait* et d'héritage de type. Un type abstrait est décrit par un ensemble de variables et ne peut avoir d'instance directe. Un sous-type possède les mêmes variables que le ou les types dont il hérite plus des variables supplémentaires.

Par exemple un type abstrait peut être défini pour regrouper tous les objets mobiles de l'environnement qui sont au moins caractérisé par une position et une vitesse de déplacement.

Finalement nous donnons la définition suivante d'un objet :

Définition 4.5 (Objet) *Un objet o est un tuple $o = \langle id_o, T_o, \delta_o \rangle$ avec :*

- id_o un identifiant unique,
- $T_o = \langle \Delta_T, \Phi_T \rangle$ le type de l'objet, et
- $\delta_o(t) \in \Delta_T$ l'état dynamique de o à l'instant t .

Exemple 4.1 – La flèche de Zénon.

Pour illustrer la modélisation des objets de l'environnement nous considérons le cas d'une flèche en mouvement, et pour simplifier les expressions, nous nous limitons à un espace à une dimension.

A tout instant une flèche peut être caractérisée par sa position dans l'espace $x(t) \in \mathbb{R}$ et par sa vitesse de déplacement instantanée $v(t) \in \mathbb{R}$.

Pour modéliser ce problème, nous définissons un type d'objet flèche permettant de décrire les propriétés de toute flèche :

$$T_{arrow} = \langle \mathbb{R}^2, \Phi_{arrow} \rangle$$

La fonction d'évolution instantanée d'une flèche peut s'exprimer par :

$$\begin{aligned}\Phi_{arrow} : \mathbb{R}^2 &\rightarrow \mathbb{R}^2 \\ \langle x, v \rangle &\mapsto \langle x + v, decelerate(v) \rangle\end{aligned}$$

où *decelerate* est une fonction de décélération.

Finalement une flèche instance du type T_{arrow} s'exprime par le tuple :

$$\langle arrow_1, T_{arrow}, \langle x(t), v(t) \rangle \rangle$$

4.4 Corps : incarnation d'un esprit dans un objet

Certains objets ont la particularité d'être relié à un esprit, on parle alors de *corps*. Un corps joue le rôle de médiateur entre un esprit et son environnement : il permet à l'esprit d'agir sur son environnement, de percevoir son environnement et d'être perçu par les autres esprits. Un corps est la manifestation d'un agent dans son environnement, il permet son existence même dans celui-ci.

Nous avons dit à la section 4.2 qu'un esprit pouvait influencer et percevoir son environnement. Nous pouvons maintenant préciser qu'il le fait au travers d'un corps : c'est un corps qui émet des percepts vers son esprit et qui reçoit des influences de son esprit. Le corps va permettre de transformer les influences d'un esprit en activité dans l'environnement.

Nous notons *HoldBody* la relation entre un corps et un esprit. Nous imposons qu'un corps ne soit relié qu'à un unique esprit, par contre nous ne posons pas par défaut la contrainte qu'un esprit ne soit relié qu'à un unique corps. Cette contrainte pourra être établie dans des cas spécifiques.

Définition 4.6 (Relation *HoldBody*) La relation *HoldBody* est définie entre l'ensemble des esprits M et l'ensemble des objets O de la manière suivante :

$$\forall m \in M, \quad \forall o \in O, \quad \text{HoldBody}(m, o) \Leftrightarrow o \text{ est un corps de } m$$

Un esprit peut être en relation avec plusieurs corps mais un corps n'est en relation qu'avec un seul esprit :

$$\forall o \in O, \forall m, m' \in M \quad \text{HoldBody}(m, o) \wedge \text{HoldBody}(m', o) \Leftrightarrow m = m'$$

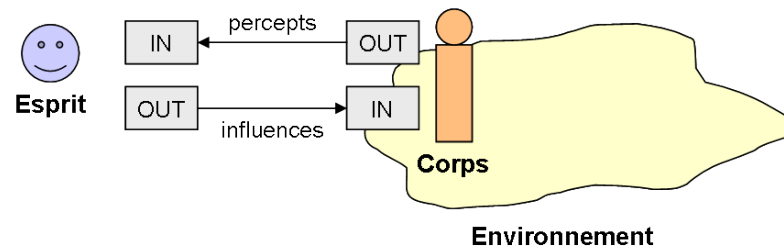


FIG. 4.3 – Lien esprit – corps

La figure 4.3 illustre la relation entre un esprit et un corps. On peut voir sur cette figure que nous plaçons un esprit à l'extérieur de l'environnement, alors qu'un corps est situé dans l'environnement.

Définir les capacités matérielles d'un agent sur son environnement

Un agent va pouvoir agir et percevoir son environnement au travers de son corps, cependant un agent n'a pas la possibilité de contrôler son corps de manière absolue, il peut seulement l'influencer. Un corps est un objet de l'environnement, l'évolution de son état est donc soumise aux lois de l'environnement ; mais cette évolution prend aussi en compte les influences qui sont émises par son esprit.

Ainsi le concept de corps permet de décrire explicitement ce que peut faire et ne peut pas faire un esprit dans l'environnement. Il permet donc de définir les *capacités matérielles* qui permettent aux agents de modifier la réalité brute. Un corps délimite l'autonomie de son esprit : cette autonomie correspond alors au fait de pouvoir exploiter librement les capacités offertes par son corps dans l'environnement. Cet aspect est essentiel pour garantir la sécurité du système, notamment dans un système ouvert où il

n'est pas possible de faire des présupposés sur les esprits.

Pour décrire comment un esprit peut intervenir sur l'évolution de l'état dynamique d'un de ses corps nous utilisons la notion d'influence qui permet de décrire comment l'esprit souhaite que l'état dynamique de son corps évolue, et la notion de *loi de réaction* qui décrit comment un ensemble d'influences peut modifier concrètement l'état dynamique d'un corps. Le concept de loi de réaction est ici similaire à celui employé dans [Helleboogh *et al.*, 2007], nous en adaptons simplement la formalisation pour prendre en compte l'état dynamique d'un objet.

Définition 4.7 (Influence) *Une influence est un tuple*

$f = \langle m, o, name, \langle p_1, p_2, \dots, p_k \rangle \rangle$ avec :

- m est l'esprit qui a émis l'influence ;
- o est l'objet à qui l'influence est destinée ; la relation $HoldBody(m, o)$ doit être vérifiée pour que l'influence soit valide ;
- $name$ est le nom de l'influence ;
- $\langle p_1, p_2, \dots, p_k \rangle$ est une liste d'arguments.

Définition 4.8 (Loi de réaction) *Une loi de réaction pour un type d'objet T est une fonction définie par :*

$$RLaw : \Delta_T \times \mathbb{P}(\Gamma) \rightarrow \Delta_T$$

$$(\delta, \{\gamma_1, \gamma_2, \dots, \gamma_i\}) \mapsto \delta'$$

Une loi de réaction est définie pour un type d'objet. C'est une fonction qui permet de calculer le nouvel état dynamique d'un corps à partir de son état dynamique actuel et des influences qui n'ont pas encore été prises en compte.

Pour un état dynamique et un ensemble d'influences donnés, une loi de réaction peut retourner l'état dynamique initial, ce qui signifie que les influences émises par un esprit n'engendrent pas nécessairement une transformation de son corps. Un esprit peut potentiellement envoyer tout type d'influences à son corps, mais seuls certains types d'influences auront un effet réel sur le corps. Nous parlerons alors d'*influences valides*.

C'est donc la définition de la loi de réaction pour un type de corps qui détermine comment un corps peut être influencé par un esprit.

Nous étendons pour la définition 4.4 d'un type d'objet en lui ajoutant une loi de réaction car deux types de corps peuvent avoir les mêmes propriétés mais réagir différemment aux influences.

Définition 4.9 (Type d'objet) Un type d'objet T est un tuple $\langle \Delta_T, \Phi_T, RLaw_T \rangle$ où Δ_T est un ensemble d'états dynamiques possibles, Φ_T est une fonction d'évolution instantanée et $RLaw_T$ est une loi de réaction.

Exemple 4.2 – Robot mobile.

Considérons un robot doté d'un moteur lui permettant de se déplacer et d'un frein. Le robot constitue ici un corps qui va pouvoir être relié à un esprit. Les possibilités qui sont données l'esprit relié à ce corps sont 1) régler la puissance du moteur et 2) régler la puissance du frein. Afin de simplifier l'exemple nous ne considérons qu'une seule dimension pour le déplacement du robot.

Pour modéliser l'état du robot nous utilisons les variables suivantes :

- $x \in \mathbb{R}$: sa position.
- $v \in \mathbb{R}$: sa vitesse de déplacement instantanée.
- $e \in \mathbb{N}$: son niveau d'énergie.
- $p \in [0, 1]$: le réglage en pourcentage de la puissance moteur.
- $b \in [0, 1]$: le réglage en pourcentage de la puissance du frein.

La fonction d'évolution d'état instantanée du robot est définie par :

$$\Phi_{robot}(x, v, e, p, b) = (x', v', e', p, b)$$

$$\text{avec } \begin{cases} x' = x + v \\ v' = Power(v, e, p, b) \\ e' = Consumption(e, p) \end{cases}$$

Les influences valides pour le robot sont de la forme :

- $\langle m, o, setPower, \langle val \rangle \rangle$, avec $val \in [0, 1]$, ou
- $\langle m, o, setBrake, \langle val \rangle \rangle$, avec $val \in [0, 1]$.

La loi de réaction¹ du robot est définie par :

$$RLaw_{robot}(\langle x, v, e, p, b \rangle, \langle m, o, name, val \rangle) = \langle x, v, e, p', b' \rangle$$

¹Une loi de réaction prend en paramètre un ensemble d'influences. Pour simplifier les notations sur cet exemple, nous ne l'avons définie que pour une unique influence. On peut considérer pour cet exemple que seule la dernière influence reçue est prise en compte.

$$\text{avec } p' = \begin{cases} val \text{ si } name = setPower \text{ et } val \in [0, 1] \\ p \text{ sinon} \end{cases}$$

$$\text{et } b' = \begin{cases} val \text{ si } name = setBrake \text{ et } val \in [0, 1] \\ b \text{ sinon} \end{cases}$$

Calcul de la réaction à des influences d'un corps

Les corps ne sont pas synchronisés avec leur esprit, l'émission d'influences par un esprit à un corps n'entraîne pas une transformation immédiate de son état dynamique. Les influences émises sont simplement stockées (ce qui est représenté par la boîte `in` du corps sur la figure 4.3). C'est l'environnement qui gère la dynamique des objets qui le composent, c'est donc lui qui déterminera le moment de la réaction d'un corps. La phase de réaction d'un corps correspond à l'application de la loi de réaction de son type sur les influences qui sont stockées.

Nature des corps

La notion de corps que nous définissons dans le modèle MASQ est très générale. Cette notion suggère généralement une composante physique, comme dans l'exemple du robot mobile, mais nous considérons ici cette notion comme tout moyen de percevoir et d'agir dans un environnement, quelque soit la nature de celui-ci. Cette notion de corps est identique à la notion de *mode* présentée dans [Báez-Barranco *et al.*, 2006].

Pour illustrer la nature différente que peut avoir un corps nous donnons les exemples suivants :

- si l'on considère l'Internet comme un environnement, un navigateur peut être considéré comme un corps dans cet environnement ;
- un client chat est un autre type de corps qui permet d'interagir sur l'Internet ;
- enfin nous pouvons citer la notion d'avatar qui est communément utilisée dans les jeux et qui illustre parfaitement la notion de corps : un avatar est la représentation d'un joueur dans un jeu, c'est aussi le moyen de définir les actions que le joueur peut réaliser dans le jeu.

Dynamique de l'incarnation d'un esprit dans un corps

Dans le méta-modèle MASQ, la relation entre un corps et un esprit est dynamique. Nous avons déjà dit qu'un esprit pouvait posséder plusieurs corps, de plus un esprit peut au cours du temps acquérir de nouveaux corps ou bien perdre des corps qu'il possède. Les capacités matérielles dont dispose un agent sur l'environnement ne sont donc pas figées, elles peuvent évoluer suivant les corps que qui sont pris ou libérés par l'esprit. Nous reviendrons sur ce point dans la prochaine section.

4.5 Espace brut

Nous introduisons le concept d'espace brut pour décrire l'environnement d'un SMA. D'un point de vue conceptuel, l'environnement est composé d'objets et son rôle principal est de mettre en relation ces objets et d'établir les lois qui vont déterminer leurs interactions.

Dans le méta-modèle MASQ nous adoptons une vision systémique de l'environnement : nous ne le considérons pas comme une entité monolithique centrale mais comme un système d'espaces bruts. Nous définissons donc l'environnement comme un ensemble d'espaces bruts qui sont en relation.

L'ensemble des espaces bruts définit une structure sur l'ensemble des objets de l'environnement : chaque objet appartient à un unique espace brut. Pour exprimer cette relation entre l'ensemble des objets O et l'ensemble des espaces bruts BS nous définissons la relation *BelongTo*.

Définition 4.10 (Relation *BelongTo*) *La relation *BelongTo* est définie entre l'ensemble des objets O et l'ensemble des espaces bruts BS de la manière suivante :*

$$\forall o \in O, \quad \forall bs \in BS \quad \text{BelongTo}(o, bs) \Leftrightarrow \text{l'objet } o \text{ appartient à l'espace brut } bs$$

Tout objet appartient nécessairement à un unique espace brut, ce qui se traduit par :

$$\forall o \in O \quad \exists! bs \in BS \quad \text{tel que } \text{BelongTo}(o, bs)$$

La figure 4.4 illustre cette structuration de l'environnement par un ensemble d'espaces bruts.

Dans un premier temps, nous allons considérer la dimension interne d'un espace brut. Ensuite nous aborderons les relations qui peuvent être établies entre les espaces bruts pour constituer l'environnement du système.

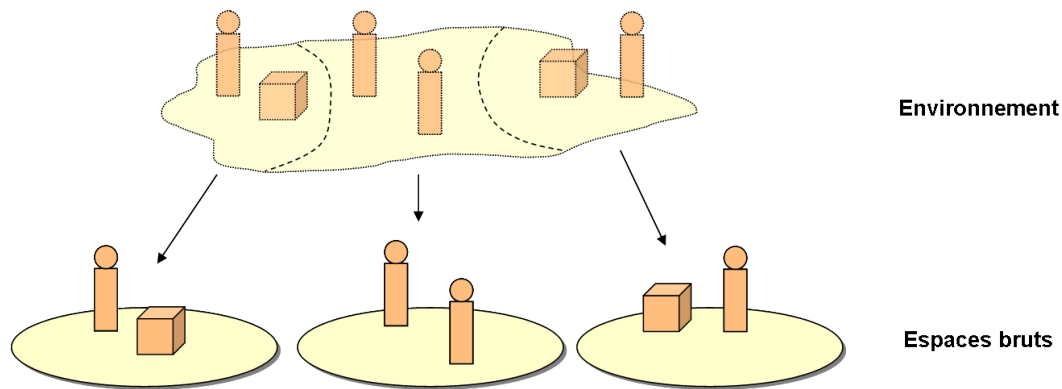


FIG. 4.4 – Structuration de l’environnement par des espaces bruts

4.5.1 Dimension interne d’un espace brut

Fragment de la réalité brute

L’environnement d’un SMA constitue la réalité brute du système (les quadrants E-I et E-C). Il maintient un état des choses qui est objectif pour les agents et qui s’impose à eux. De plus cette réalité brute est le seul lieu possible pour leur activités.

Un espace brut est un fragment de l’environnement et donc un fragment de la réalité brute du système. Ce fragment est décrit par les objets qui sont dans l’espace brut et par les relations qu’ils entretiennent au sein de l’espace. Un espace brut n’est pas une simple collection d’objets, mais c’est un véritable système au sens de la théorie des systèmes :

- Les objets sont les parties constituantes d’un espace brut ;
- L’espace brut constitue une frontière qui sépare ses objets du reste de l’environnement ;
- Les objets sont inter-reliés dans l’espace brut. Les relations entre objets peuvent être de toutes sortes. Deux exemples significatifs de relations sont les transports (c’est-à-dire comment les objets peuvent se déplacer dans l’espace) et les communications (comment de l’information peut être échangée entre les objets).

Espace physique – Espace social

La nature des espaces bruts peut être très variée, mais deux grandes classes peuvent être distinguées : les espaces physiques et les espaces sociaux.

Un espace physique est la modélisation d’une partie du monde physique (un terrain de football par exemple). Un tel espace peut être doté d’une topologie particulière qui permet de situer les objets dans l’espace, et d’établir par exemple des relations de distance.

Un espace social définit ses propres structures d’interaction. Pour un espace de

nature sociale, on parle d'objet et de corps sociaux. Ce type d'espace brut peut par exemple servir de support d'interaction pour des agents communicants. En fournissant les structures d'interaction nécessaires à la communication, l'espace détermine les capacités matérielles dont disposent les agents pour communiquer, les contraintes sur qui peut communiquer avec qui, et les mécanismes nécessaires à l'acheminement des messages. Il est aussi possible de parler de topologie pour les espaces sociaux [Zambonelli *et al.*, 2003]. Par exemple une organisation qui utilise des rôles de « maître » et des rôles « d'esclave » définit une topologie hiérarchique.

Cependant dans le modèle MASQ la distinction physique / social n'est que conceptuelle. En pratique il ne s'agit que d'objets et de corps disposant de capacités pour agir et percevoir dans un espace brut et de lois qui permettent de déterminer le résultat des interactions entre tous ces objets, qu'il s'agisse de déplacer un ballon ou d'engager une conversation.

Relation aux esprits

Les esprits sont reliés aux objets par la relation *HoldBody* et les objets sont reliés aux espaces bruts par la relation *BelongTo*, il y a donc une relation indirecte entre les esprits et les espaces bruts que nous exprimons par une relation d'*incarnation* : un esprit est incarné dans un corps qui est situé dans un espace brut. Nous notons cette relation *Embody*.

Définition 4.11 (Relation *Embody*) *La relation *Embody* est une relation ternaire entre l'ensemble des esprits M , l'ensemble des objets O et l'ensembles des espaces bruts BS définie par :*

$$\forall m \in M, \forall o \in O, \forall bs \in BS \quad \text{Embody}(m, o, bs) \Leftrightarrow \text{HoldBody}(m, o) \wedge \text{BelongTo}(o, bs)$$

Un esprit peut posséder plusieurs corps dans différents espaces bruts. De manière générale, nous n'imposons pas la contrainte d'un corps unique pour un esprit dans un espace brut, mais cette contrainte peut être établie spécifiquement pour certains espaces.

Pour les esprits les espaces bruts sont des boites noires : un espace brut n'est perceptible qu'au travers des percepts délivrés par un corps situé à l'intérieur de cet espace. Un espace brut délimite la portée de la perception et des actions possibles pour un esprit qui possède un corps dans cet espace. La perception dans un espace brut reste cependant locale, sauf cas particulier, un corps ne permet pas de percevoir l'intégralité d'un espace.

Contexte d'interaction

Un espace brut définit un contexte d'interaction pour les objets qui le composent. Chaque objet manifeste individuellement une activité qui est exprimée à chaque instant par son état dynamique. Ces différentes activités qui sont réalisées au sein d'un espace brut peuvent entrer en interférence les unes avec les autres comme par exemple lorsque deux objets en mouvement entrent en collision. C'est au niveau de l'espace brut que l'on décrit sous quelles conditions une interférence peut apparaître et quels en sont les effets. Nous reprenons ici le concept de *loi d'interférence* proposé dans [Helleboogh *et al.*, 2007] pour décrire cet aspect.

Comme dans [Helleboogh *et al.*, 2007] nous considérons qu'une interférence peut résulter en une transformation d'activités pour les objets qui sont impliqués. Par exemple, lorsque que deux objets entrent en collision, leur vitesse et leur direction de déplacement peuvent être modifiées. Une telle transformation d'activités s'exprime dans MASQ par un changement de l'état dynamique de ces objets.

Mais nous considérons aussi deux autres types d'effets possibles pour une interférence :

- Une interférence peut conduire à la création ou à la destruction d'objets dans l'espace brut. Un exemple est le cas d'un robot disposant de la capacité de tirer des missiles. Le tir d'un missile peut être modélisé comme la création d'un objet missile à l'instant du tir qui évoluera ensuite selon sa propre dynamique. Par la suite, la collision du missile avec un obstacle donnera lieu à une nouvelle interférence dont un des effets est de détruire le missile.
- L'attribution d'un corps ou au contraire la libération d'un corps peut aussi être la conséquence d'une interférence. Ceci permet de définir toutes les règles qui régissent l'attribution et la libération d'un corps. Ainsi l'acquisition d'un nouveau corps pour un esprit dans un espace brut nécessite l'apparition d'une interférence particulière dans cet espace. De manière générale, une telle interférence est le résultat de certaines activités qui se produisent dans l'espace brut. Cette approche permet de décrire différents modes possibles d'attribution d'un corps. Un esprit peut s'attribuer lui-même un nouveau corps s'il dispose d'ores et déjà de corps capables de produire l'interférence nécessaire, ou bien cette capacité peut être détenue uniquement par un autre agent qui est alors le seul en mesure d'attribuer le nouveau corps. Enfin l'attribution d'un corps peut aussi nécessiter l'interaction de différents agents. D'autre part, une interférence peut avoir pour conséquence la perte d'un corps pour un esprit.

Finalement, une loi d'interférence comprend :

1. *Une condition d'application* qui s'exprime en termes de prédicats sur les états dynamiques des objets et sur les relations qui sont établies dans l'espace brut entre les objets.
2. *Un ensemble d'effets*, chaque effet pouvant être :
 - une transformation de l'état dynamique d'un objet ;
 - la création ou la suppression d'un objet ;
 - l'établissement ou la suppression du lien entre un esprit et un corps.

Nous donnons la définition suivante d'une transformation d'états dynamiques :

Définition 4.12 (Transformation d'états dynamiques) *Une transformation d'états dynamiques est une fonction trans définie par :*

$$\begin{aligned} trans : \Delta_1 \times \Delta_2 \times \dots \Delta_n &\rightarrow \Delta_1 \times \Delta_2 \times \dots \Delta_n \\ \delta_1 \times \delta_2 \times \dots \delta_n &\mapsto \delta'_1 \times \delta'_2 \times \dots \delta'_n \end{aligned}$$

A partir des états dynamiques d'un ensemble d'objets, une transformation d'états dynamiques produit un nouvel état dynamique pour chacun de ces objets.

Finalement, nous définissons une loi d'interférence par :

Définition 4.13 (Loi d'interférence) *Une loi d'interférence est un tuple :*

$$ILaw = \langle Cond, Trans, O^+, O^-, Embod^+, Embod^- \rangle$$

avec :

- *Cond* la condition d'application de la loi,
- *Trans* un ensemble de transformations d'états dynamiques,
- O^+ un ensemble d'objets qui seront créés dans l'espace brut,
- O^- un ensemble d'objets qui seront détruits dans l'espace brut,
- $Embod^+$ un ensemble de couples $\langle m, o \rangle$ où m est un esprit et o un objet, chaque couple dénotant l'établissement d'un lien entre l'esprit m et l'objet o ,
- $Embod^-$ un ensemble de couples $\langle m, o \rangle$ où m est un esprit et o un objet, chaque couple dénotant la suppression du lien existant entre l'esprit m et l'objet o .

L'utilisation des lois d'interférences est illustrée sur l'exemple 4.3.

Exemple 4.3 – Collision de deux billes.

Considérons le cas de deux billes en mouvement qui entrent en collision (figure 4.5). Chacune des deux billes possède à tout instant un état dynamique qui caractérise sa position et sa vitesse de déplacement.

Nous allons exprimer par une loi d’interférence dans quelle circonstance une collision peut avoir lieu entre deux billes et quelle sera la transformation à appliquer à leur état dynamique respectif.

Pour simplifier nous considérons un espace à une dimension et une collision sans frottement.

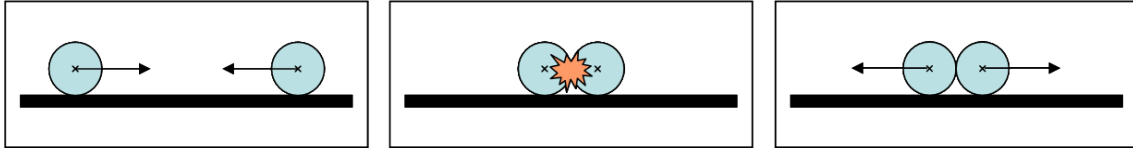


FIG. 4.5 – Collision de deux billes

Pour modéliser cet exemple nous définissons le type d’objet T_{Ball} avec les variables d’états suivantes :

- $x \in \mathbb{R}$: la position de l’objet ;
- $v \in \mathbb{R}$: la vitesse de l’objet.

La loi d’interférence pour la collision entre objets mobiles est définie par :

$$ICollide = \langle Cond, \{Bounce\}, \emptyset, \emptyset, \emptyset, \emptyset \rangle$$

avec :

$$Cond : \exists o_1 \in T_{Ball}, \exists o_2 \in T_{Ball} \text{ tel que } dist(o_1, o_2) = \epsilon$$

$$\begin{aligned} Bounce : \Delta_{Ball} \times \Delta_{Ball} &\rightarrow \Delta_{Ball} \times \Delta_{Ball} \\ (x_1, v_1), (x_2, v_2) &\mapsto (x_1, -v_1), (x_2, -v_2) \end{aligned}$$

La condition exprime que l’interférence $ICollide$ s’applique lorsque la distance entre deux objets du type T_{Ball} est égale à la constante ϵ . La fonction distance est donnée par $dist(o_1, o_2) = |o_1.x - o_2.x|$.

La transformation $Bounce$ qui est appliquée aux objets exprime leur rebond, elle consiste à changer le signe de la vitesse des deux objets.

Pour cette interférence il n'y a pas d'objet créé ou supprimé et les liens corps – esprit ne sont pas affectés.

L'exemple 4.4 donne une autre illustration des lois d'interférence, dans le contexte d'un espace social cette fois-ci.

Exemple 4.4 – Attribution d'un corps social.

Considérons l'exemple d'un forum pour lequel il existe un système de modération. Nous distinguons deux types de compte pour les utilisateurs de ce forum : les comptes de type modérateur et les comptes de type simple utilisateur. Pour les besoins de cet exemple, nous ne considérons qu'un aspect du problème : le fait que seul un modérateur possède la capacité de promouvoir un simple utilisateur au niveau de modérateur.

Une modélisation possible de cet exemple est la suivante. L'ensemble des comptes utilisateurs sont administrés au niveau d'un espace social $bs_{account}$. Cet espace contient des corps sociaux de deux types différents : le type $T_{moderator}$ pour les modérateurs et le type T_{user} pour les utilisateurs simples.

Enfin le fait de promouvoir un simple utilisateur au niveau d'un administrateur peut être modélisée par l'attribution d'un corps de type $T_{moderator}$ à un esprit qui possède déjà un corps de type T_{user} . Et pour rendre possible cette attribution de corps, avec la contrainte que, seul un corps de type $T_{moderator}$ peut la produire, nous définissons la loi d'interférence suivante :

$$IGrantModerator = \langle Cond, \emptyset, \{new_moderator\}, \emptyset, \{new_embody\}, \emptyset \rangle$$

avec :

$$Cond : \exists moderator \in T_{moderator}, \exists user \in T_{user}$$

$$\text{tel que } moderator.grantModerator = id_{user}$$

$$new_moderator \text{ un nouveau corps de type } T_{moderator}$$

$$new_embody = \langle m, new_moderator \rangle \text{ avec } m \in M \text{ tel que } Embody(m, user)$$

Finalement nous donnons la définition suivante pour le concept d'espace brut :

Définition 4.14 (Espace brut) *Un espace brut est un tuple*

$bs = \langle id, O_{bs}, NRel, ILaws \rangle$ avec :

- id l'identifiant de bs ,
- $O_{bs} \subseteq O$ l'ensemble des objets de bs ,
- $NRel$ un réseau de relations² entre les objets de bs ,
- $ILaws$ l'ensemble de lois d'interférence de bs .

Cycle d'évolution d'un espace brut

La dynamique interne d'un espace brut est déterminée par ses propres lois.

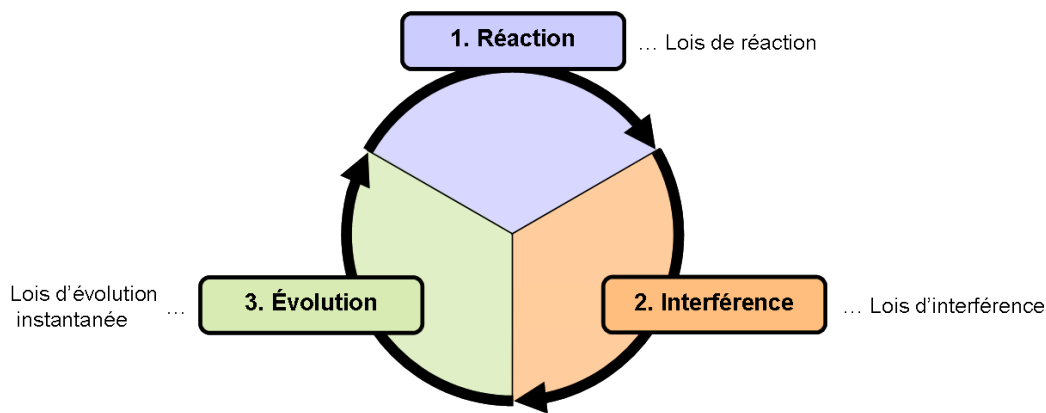


FIG. 4.6 – Le cycle à 3 phases d'un espace brut

Comme l'illustre la figure 4.6, l'évolution d'un espace brut est un cycle perpétuel qui se décompose en trois phases :

1. **Réaction** : Cette première phase correspond à la prise en compte des influences des esprits. C'est au cours de cette phase que les lois de réaction sont appliquées pour changer l'état dynamique des corps qui ont reçu des influences. Dans le cas général, un espace brut n'est pas synchronisé avec les esprits, mais ceci peut être le cas notamment pour faire de la simulation, la phase de réaction est alors bloquante et le cycle est interrompu le temps que tous les esprits émettent leurs influences.
2. **Interférence** : Cette deuxième phase correspond à la détection des interférences éventuelles et à leur résolution, c'est-à-dire l'application des transformations d'états dynamiques nécessaires, les créations et les suppressions d'objets, ainsi que les attributions et les libérations de corps.
3. **Évolution** : La dernière phase correspond à la transformation de l'espace brut pour passer d'un instant t à un instant $t + 1$. Cette transformation est effectuée en

²L'expression concrète de ce réseau de relations dépend du domaine qui est modélisé. Il comporte toutes les relations qui peuvent être établies entre les objets de l'espace brut. La relation de distance est un exemple de relation.

appliquant à chaque objet sa fonction d'évolution instantanée.

D'un point de vue opérationnel, le cycle d'exécution d'un espace brut peut être associé à son propre *thread* ou bien être activé par un évènement extérieur. Dans le premier cas, le cycle s'exécute continuellement et la dimension temporelle est gérée au niveau de l'espace. Le second cas permet de synchroniser différents espaces.

Contrainte d'intégrité interne pour un espace brut

Chaque espace brut est constitué de ses propres objets et structures d'interaction, et il est régi par ses propres lois (lois de réaction, lois d'évolution et lois d'interférence). Nous imposons une contrainte d'intégrité interne pour les espaces bruts : aucun élément extérieur ne peut accéder ou modifier directement l'état interne d'un espace brut. Cette contrainte permet d'assurer à tout moment que l'évolution d'un espace brut est bien conforme à ses lois. Ainsi chaque espace brut gère sa propre évolution de manière indépendante du reste de l'environnement.

Un espace brut n'est pas pour autant isolé du reste de l'environnement. Il est en interrelation avec le reste du système, et des facteurs extérieurs, tels que les influences des esprits, peuvent avoir des conséquences sur son évolution. Cependant, ces conséquences sont nécessairement le résultat de l'application des lois de l'espace.

Dans MASQ, ce sont les lois de réaction qui permettent d'intégrer dans un espace brut des évènements extérieurs. Nous avons présenté leur fonctionnement en ce qui concerne les influences des esprits, et nous aborderons dans la prochaine section leur rôle dans les échanges qui peuvent se produire entre plusieurs espaces bruts. Ce qu'il est important de noter, et qui garantit l'intégrité des espaces bruts, c'est que tout évènement extérieur à un espace brut qui n'est pas capturé par une loi de réaction n'aura aucune incidence sur l'évolution de cet espace.

Cette conception de l'environnement en espaces bruts présente différents avantages :

- Le fait de garantir l'intégrité des espaces bruts permet d'assurer une cohérence locale de l'environnement. Chaque espace brut, au travers de ses lois, détermine les activités qui sont possibles en son sein, et le résultat de ces activités. Cette « réalité » que constitue un espace brut s'impose alors aux esprits qui peuvent la percevoir, l'exploiter pour atteindre leurs objectifs, mais en aucun cas en transgresser les règles. Cet aspect est essentiel pour la sécurité du système.
- Cette approche est une application du principe *diviser pour mieux régner*. L'environnement n'est pas conçu comme une entité monolithique pour laquelle il faudrait

décrire toutes les lois qui régissent le système global. Au contraire, l'environnement est décomposé en plusieurs sous-parties, les espaces bruts, chacune définissant son propre contexte d'activité qui est régi par ses propres règles. Cette distribution de l'environnement sur un ensemble d'espaces bruts permet de réduire la complexité de la conception de l'environnement.

- Chaque type d'espace brut peut être conçu et mis à jour séparément, de part l'indépendance qu'il existe entre l'évolution de chaque espace brut. Il est ainsi plus facile de faire évoluer l'environnement d'un SMA, notamment en ajoutant de nouveaux types d'espace brut pour permettre de nouvelles formes d'interaction dans le système.

4.5.2 Relations entre espaces bruts

Nous avons considéré jusqu'à présent l'activité interne d'un espace brut, nous allons maintenant envisager comment les espaces bruts s'agencent entre eux pour constituer l'intégralité de l'environnement et quelles sont les interactions qui peuvent avoir lieu entre eux.

Deux types principaux de relation sont nécessaires entre des espaces bruts : les transports et les communications.

La relation de transport induit la mobilité des agents dans l'environnement, ou comment un esprit peut accéder à de nouveaux espaces bruts. Au cours du temps, les esprits vont pouvoir entrer et sortir de différents espaces bruts, ce qui revient à acquérir et libérer des corps dans ces espaces bruts. Une relation d'accessibilité doit donc être établie entre les espaces bruts. Cette relation peut bien évidemment dépendre de la nature du corps détenu dans l'espace de départ ou bien de la nature du corps demandé dans l'espace d'arrivée.

La relation de communication induit la circulation d'information entre espaces bruts. Un espace brut peut acquérir de l'information qui est produite par les différentes activités qui se produisent dans un autre espace brut.

Nous proposons donc dans MASQ un mécanisme qui permette aux différents espaces bruts d'interagir, et nous analyserons différents usages qui peuvent être fait de ces interactions.

4.5.2.1 Mécanisme d'interaction entre espaces bruts

Les relations et les interactions qui sont nécessaires entre des espaces bruts pour modéliser l'intégralité de l'environnement peuvent être de nature très différente. Dans

MASQ nous proposons un mécanisme générique sur la base duquel il est possible de construire différentes relations spécifiques entre espaces bruts. Nous allons tout d'abord présenter ce mécanisme puis ensuite nous présenterons quelques patrons de conception pour modéliser des relations spécifiques en exploitant ce mécanisme.

Ce mécanisme repose sur deux notions : les *références externes* et les *influences entre espaces bruts*.

Référence externe

Chaque espace brut possède un identifiant unique qui permet de le désigner, et chaque objet possède lui même un identifiant unique au sein d'un espace brut. Le couple (id espace brut – id objet) permet de désigner un objet dans un espace depuis l'extérieur. Ainsi une référence externe est la désignation d'un objet dans un espace brut depuis l'extérieur.

Il est important de préciser qu'aucun élément extérieur à un espace brut ne peut avoir accès aux objets qui se trouvent à l'intérieur. Il est donc simplement possible d'en avoir une référence. Et comme il n'existe pas de contrainte d'intégrité entre les espaces bruts, la référence à un objet ne garantit pas l'existence effective de l'objet dans l'espace désigné.

Influence entre espaces bruts

Pour que des échanges entre espaces bruts soient possibles tout en préservant l'intégrité de chacun des espaces bruts, nous étendons la notion d'influence que nous avons présenté pour les esprits en introduisant la possibilité qu'un objet dans un espace brut soit influencé par un autre espace brut.

Ainsi, une influence pourra aussi être produite par un espace brut et destinée à un objet situé dans un autre espace brut.

La production d'une influence pourra être un des effets d'une interférence qui se produit dans un espace brut. L'émission d'une influence vers un objet externe est rendue possible par les références externes. Enfin toute influence émise par un espace brut pourra être traitée de la même manière que les influences émises par les esprits, c'est-à-dire en appliquant la loi de réaction de l'objet qui reçoit l'influence.

Avec ce mécanisme d'influences et de réactions, un espace brut n'a pas la possibilité de modifier directement un autre espace brut, mais simplement d'influencer son évolution. L'émission d'une influence depuis un espace brut E_1 vers un espace brut E_2 dépend des lois de l'espace E_1 . Par contre l'effet qui peut être produit dans l'espace E_2 ne dépend que des lois de celui-ci.

Notons que tout objet peut ainsi recevoir des influences d'un autre espace brut, que ce soit un simple objet ou bien un corps associé à un esprit. A un instant donné, un objet

peut donc être influencé par différentes sources.

4.5.2.2 Patrons de conception

En génie logiciel, un patron de conception (*design pattern* en anglais) est un concept destiné à résoudre des problèmes récurrents de conception. A la différence d'un algorithme qui s'attache à décrire d'une manière formelle comment résoudre un problème particulier, les patrons de conception décrivent des procédés de conception généraux.

Dans MASQ, toute interaction entre espaces bruts est traitée en exploitant les références externes et les influences entre espaces bruts. Ce mécanisme très générique permet de concevoir des solutions pour des relations et des interactions de nature très différente entre espace brut. Cependant pour apporter plus de précision quant aux mises en œuvre possibles de ce mécanisme, nous proposons un certain nombre de patrons de conception qui répondent à des problèmes qui nous paraissent essentiels et récurrents dans la conception d'un environnement. Il faut bien noter que ces patrons de conceptions qui seront présentés par la suite ne font pas partie intégrante du méta-modèle MASQ, ce ne sont que des mises en œuvre possibles du méta-modèle.

Relation d'identité

Dans le méta-modèle MASQ l'identité d'un esprit est masquée par le ou les corps qu'il possède dans un espace brut. En d'autres termes, il est possible pour un esprit de percevoir l'identité d'un corps dans un espace brut, mais pas l'identité de l'esprit qui se trouve derrière le corps. Ainsi, deux esprits m_1 et m_2 peuvent interagir dans différents espaces bruts sans que l'un et l'autre ne sache qu'ils interagissent avec le même esprit car les corps dans chaque espace brut sont différents.

Cependant, dans certaines situations particulières, il est souhaitable que la relation d'identité puisse être établie entre deux corps situés dans des espaces différents. L'exemple 4.5 illustre une de ces situations.

Exemple 4.5 – Relation d'identité entre espaces brut.

Considérons le contexte d'un match de football pour lequel on modélise à la fois une composante physique et une composante sociale. La composante physique est donnée par l'espace brut « terrain de football » dans lequel se trouve les corps physiques des joueurs qui leur donne la capacité de jouer au football. La composante sociale correspond à la structure sociale d'une équipe, elle est modélisée par l'espace brut « équipe de football ». Dans cet espace

brut les esprits possèdent des corps sociaux dont les capacités dépendent de la nature de l'équipe.

La figure 4.7 illustre ce scénario.

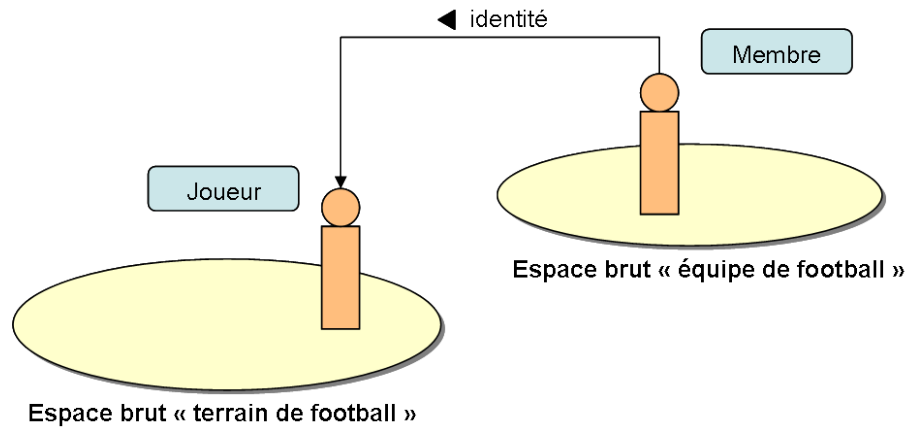


FIG. 4.7 – Relation d'identité entre objets de différents espaces bruts

Sur cet exemple, la relation d'identité est nécessaire pour que les esprits d'une équipe puissent faire la correspondance entre les membres sociaux de leur équipe et les joueurs sur le terrain. Par contre on peut remarquer que si cette correspondance doit être faite pour les membres de l'équipe, elle n'a pas à l'être pour des agents qui peuvent percevoir sur le terrain mais qui n'appartiennent pas à l'équipe. Ces derniers ne connaissent pas la structure sociale de l'équipe en question, et donc n'ont pas à faire cette association.

Pour modéliser cette relation d'identité avec MASQ, nous proposons de définir pour les objets une variable d'état spécifique qui contiendra les références externes de tous les objets pour lesquels la relation d'identité doit être établie. En définissant dans l'espace brut les capacités nécessaires à la perception de cette variable, les esprits pourront « connaître » l'identifiant des objets qui sont associés à un objet spécifique. Bien évidemment la relation d'identité entre deux objets n'est pertinente que pour un esprit qui se situe dans les deux espaces bruts, sans quoi la référence externe n'est pas une information utile. De plus, on peut remarquer que cette relation d'identité n'est pas symétrique. Elle peut être perçue dans un des deux espaces bruts et pas dans l'autre.

Publication d'information

Un type d'échange entre espace brut qui est particulièrement important est la publication d'information produite par un espace brut dans un autre espace brut. La publication d'information peut avoir de nombreux usages comme en particulier la mise en place d'un service de pages jaunes utile pour référencer les services qui sont offerts

dans d'autres espaces bruts.

Pour modéliser ce type d'échange, nous proposons le patron de conception suivant. Soient bs_{dest} l'espace brut dans lequel l'information doit être publiée et bs_{source} l'espace brut qui produit l'information. Dans bs_{dest} , un objet o_{info} sert de support à l'information qui doit être publiée. Cette information est exprimée par l'état de o_{info} , ce qui la rend perceptible par les agents qui sont présents dans bs_{dest} . bs_{source} quant à lui doit posséder une référence externe vers o_{info} . La production ou la mise à jour d'information se traduit alors par l'émission d'influences par bs_{source} vers o_{info} . Le traitement de ces influences auront pour effet de transformer l'état de o_{info} en conséquence et ainsi de mettre à jour l'information disponible dans bs_{dest} .

Il est important de noter que si la cohérence est garantie au sein d'un espace brut, elle ne l'est pas entre plusieurs espaces bruts. Ainsi l'information qui est publiée à distance peut s'avérer être inexacte à un instant donné, notamment parce que des influences de mises à jour n'ont pas encore été prises en compte.

Entrée dans un espace brut

Pour un esprit, entrer dans un espace brut signifie se voir attribuer un corps dans cet espace. Ce processus relève d'une interaction externe avec l'espace brut de destination puisque tant qu'un esprit ne dispose pas de corps dans l'espace, il n'a pas accès à la partie interne de cet espace.

L'accès à un espace brut peut être restreint de différentes manières : un mot de passe peut être nécessaire, des conditions peuvent être posées sur les propriétés du candidat, ou encore des contraintes peuvent être relative à l'état de l'espace brut de destination comme par exemple un nombre limite au delà duquel toute entrée est rejetée.

Le processus d'entrée dans un espace brut se déroule de la manière suivante :

1. Émission de la requête d'entrée.
2. Traitement de la requête (celle-ci peut être ignorée ou rejetée) suivant les lois de l'espace de destination.
3. Création d'un nouveau corps ou sélection d'un corps existant.
4. Attribution du corps au demandeur (ce qui consiste à établir un canal de communication entre l'esprit et le corps pour l'échange d'influences et de percepts).

Pour qu'une requête puisse être reçue par l'espace de destination, il est nécessaire qu'il soit en mesure de recevoir des influences depuis l'extérieur. De la même manière que nous

avons proposé d'utiliser un objet comme support pour la publication d'information, nous proposons ici d'utiliser un objet $o_{entrance}$ pour traiter les requêtes d'entrée. C'est la loi de réaction de l'objet $o_{entrance}$ qui permettra de prendre en compte ou d'ignorer les requêtes ; une requête sera prise en compte si l'influence correspondante a produit une certaine activité de l'objet $o_{entrance}$. C'est cette activité qui pourra alors causer une interférence dans l'espace de destination et ainsi créer le nouveau corps et l'attribuer au demandeur.

En ce qui concerne l'émission de la requête d'entrée, différentes situations sont possibles. Tout d'abord un esprit peut émettre lui-même sa propre requête à un espace brut. Dans ce cas il doit posséder une référence externe sur cet espace brut, que cette référence ait été encodée dans son architecture interne ou bien qu'il l'ait découvert durant ses activités précédentes. Un espace brut qui traite les requêtes directement effectuées par les esprits sera considéré comme *public*.

D'autres espaces bruts au contraire ne traiteront pas les requêtes qui proviennent directement des esprits afin de restreindre l'accès en fonction de la provenance du demandeur. Un esprit pourra alors exploiter un corps qu'il possède déjà dans un autre espace brut et qui dispose de la capacité nécessaire pour émettre la requête. Dans ce cas, l'esprit n'a pas besoin de posséder de référence externe sur l'espace brut de destination. C'est son espace d'origine qui doit la posséder, et celle-ci n'est pas forcément publique pour les esprits.

Enfin le dernier type de situations est celui où l'esprit ne possède d'aucun corps disposant de la capacité nécessaire. La seule option qui lui reste est alors de solliciter un autre agent qui lui a la capacité de provoquer l'interférence nécessaire à l'émission de la requête. Cette approche permet de donner uniquement à certains agents la capacité de faire entrer d'autres agents dans un espace brut.

Communication inter-espace brut

Une autre forme d'échange d'information entre espaces bruts peut être réalisée par l'envoi de messages par un corps dans un espace brut vers un corps dans un autre espace brut.

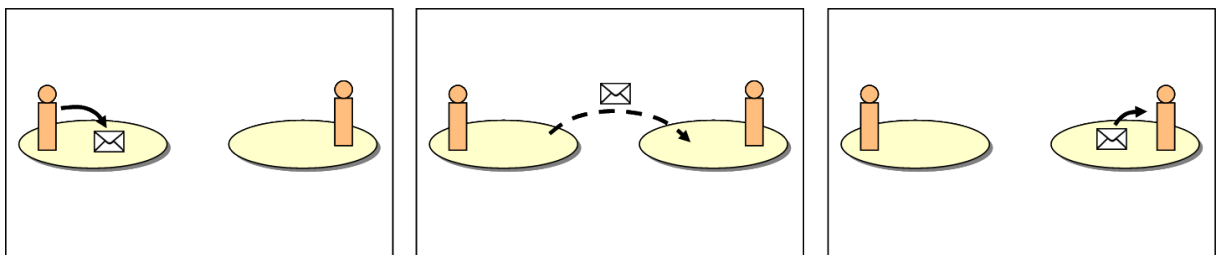


FIG. 4.8 – Communication par passage de messages entre différents espaces

Comme l'illustre la figure 4.8, l'envoi de message depuis un espace brut vers un autre

espace brut nécessite trois étapes. Tout d'abord le message doit être produit dans l'espace initial. Ensuite ce message doit être transporté vers l'espace brut de destination par l'intermédiaire de l'envoi d'une influence. Enfin le message doit être créé (un nouvel objet) dans l'espace d'arrivée pour qu'il puisse être ensuite perçu par son destinataire.

La possibilité d'une telle forme de communication entre deux espaces bruts dépend des lois des deux espaces. Les lois de l'espace initial doivent permettre la création d'un message ainsi que sa transmission (par influence) vers l'espace brut distant. Les lois de l'espace de destination doivent traiter ce type d'influence afin que la réaction engendre la création d'un message en local.

Chaque espace brut peut posséder ses propres règles en ce qui concerne l'émission à distance et la réception depuis l'extérieur.

Modélisation holonique des objets

Il est possible d'adopter une conception holonique de l'environnement en établissant des relations entre un objet et un espace brut.

Nous rappelons qu'un holon est à la fois un tout composé de sous-parties, et une sous-partie d'un tout de plus haut niveau. Rapporté à l'environnement, cela revient à considérer qu'un objet dans un espace brut est lui-même un système de composants en interaction à un autre niveau. Cette approche permet de ne pas réduire l'évolution interne d'un objet à une fonction (la loi d'évolution), mais de considérer que cette évolution dépend des interactions entre des entités d'un niveau inférieur (des objets dans un autre espace brut).

Dans le méta-modèle MASQ, cette approche peut être modélisée en définissant les échanges d'influences possibles ainsi que leurs effets entre l'objet principal et les objets qui le composent (cf. figure 4.9).

Dans une approche holonique, les espaces bruts à chaque niveau permettent de décrire les interactions possibles à ce niveau. La finesse de la décomposition dépend des besoins de modélisation du système considéré.

Par exemple, un robot mobile modélisé comme un corps à un certain niveau, peut être défini comme un ensemble d'éléments mécaniques en interaction à un niveau inférieur. Au niveau du robot, l'influence « avancer » peut se traduire par l'envoi d'influences vers les composants du robot. Dans ce cas, le changement d'état du robot ne sera pas déterminé par une unique loi d'évolution, mais ce sont les interférences entre les composants mécaniques du robot qui produiront des influences destinées à faire changer son état d'une certaine

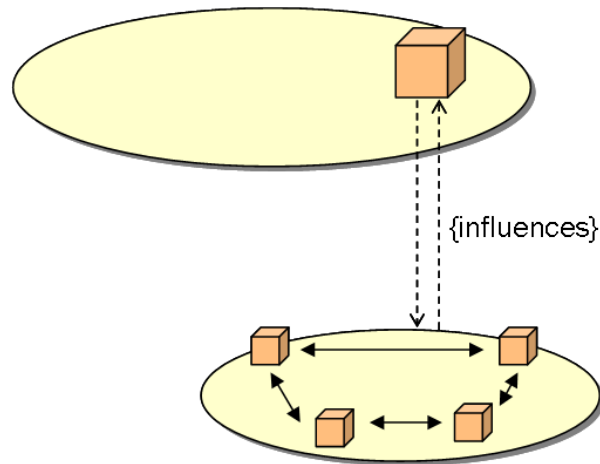


FIG. 4.9 – Modélisation holonique

manière, pour finalement le faire avancer effectivement.

4.5.3 Structuration de l'environnement

Nous venons de voir que les espaces bruts et les relations qu'ils entretiennent définissent une structuration de l'environnement. Cette structuration permet d'obtenir une synchronisation régionale du système et ainsi de trouver un compromis entre une synchronisation centralisée, qui ne permet pas de passer à l'échelle et qui n'est pas robuste, et une synchronisation totalement distribuée qui rend très difficile voire impossible le maintien de la cohérence. Avec le méta-modèle MASQ, il est possible faire varier le niveau de synchronisation en jouant sur la granularité des espaces bruts. Un modèle centralisé correspond à un unique espace brut pour représenter l'environnement alors qu'un modèle totalement distribué peut être obtenu en réduisant au maximum l'expression des espaces bruts, un espace brut étant alors une cellule dans le sens donné par Ferber [Ferber, 1999].

On peut faire une analogie entre la structuration de l'environnement et des architectures classiques de réseaux qui sont représentées sur les figures 4.10, 4.11, 4.12 :

- Un environnement monolithique centralisé correspond au modèle client–serveur (figure 4.10) : l'environnement joue le rôle de serveur pour tous les esprits qui sont des clients. Ce contrôle central permet de garantir une cohérence absolue des règles de l'environnement mais au prix d'un goulot d'étranglement et d'une dégradation de la robustesse du système qui repose entièrement sur la fiabilité du serveur.
- Un système sans environnement (autre que celui du réseau sous-jacent) dans lequel les esprits (les agents) interagissent directement et librement peut être associé à un modèle pair-à-pair (P2P) pur. Dans ce type d'approche, il n'est pas possible de

définir des règles environnementales et donc les agents doivent prendre en charge eux-mêmes tous les aspects qui incombent à l'environnement. Par exemple, un agent peut potentiellement communiquer avec n'importe quel autre agent, chacun doit donc gérer lui-même avec qui il accepte ou refuse de communiquer. La décentralisation étant absolue, il n'est pas possible de garantir des propriétés globales.

- Le modèle de structuration en espace brut que nous proposons peut être rapproché du modèle *super-peer* [Yang et Garcia-Molina, 2003] (figure 4.12). Dans ce modèle, certains nœuds du réseau, les *super-peers*, jouent le rôle de serveur pour un ensemble de simples clients (les *peers*). Les *super-peers* sont interconnectés entre eux à la manière d'un réseau P2P. L'intérêt de cette approche est de permettre un compromis entre centralisation et décentralisation. Un espace brut peut être assimilé à un *super-peer* qui centralise l'activité de plusieurs esprits qui en sont les clients. La cohérence des interactions entre ces esprits est garantie par l'espace brut. D'autre part, les espaces bruts sont interconnectés entre eux d'une manière pair-à-pair.

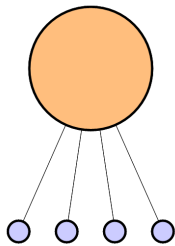


FIG. 4.10 – Architecture client-serveur

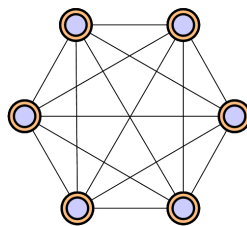


FIG. 4.11 – Architecture pair-à-pair (P2P)

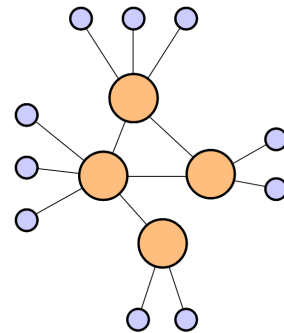


FIG. 4.12 – Architecture *super-peer*

Dynamique de la structure de l'environnement

La structuration de l'environnement par les espaces bruts est dynamique : au cours du temps des espaces bruts peuvent être créés et d'autres peuvent être détruits.

La création d'un nouvel espace brut est le résultat d'une interférence dans un espace brut.

Seul un espace brut peut lui-même se détruire, ce qui est aussi le résultat d'une interférence produite dans cet espace. La destruction d'un espace entraîne automatiquement la destruction de tous ses objets.

4.6 Culture

Nous introduisons le concept de culture pour capturer la dimension interne collective. Il s'agit à ce niveau de décrire les éléments subjectifs qui sont partagés par un ensemble d'esprits tels que les interprétations collectives, les normes sociales, les ontologies ou de manière plus générale les connaissances partagées.

Dans le contexte spécifique des SMA, l'intérêt de la culture est de fournir à un groupe d'agents un contexte qui leur permet d'atteindre une compréhension mutuelle. Pour un agent, la culture lui permet d'une part d'anticiper le comportement des autres agents, et d'autre part de déterminer dans certaines situations le comportement que l'on attend de lui. Du point de vue d'une société, la culture permet d'influencer les esprits dans leurs décisions tout en préservant leur intégrité et leur hétérogénéité. Ainsi, la culture permet la mise en place des formes de pression sociale pour obtenir une meilleure cohésion entre les membres d'une société et réduire l'apparition de conflits.

Dans le méta-modèle MASQ, un esprit peut être plongé dans une ou plusieurs cultures, ce qui lui donne accès aux connaissances partagées qui y sont définies. Pour exprimer cette relation entre l'ensemble des esprits M et l'ensemble des cultures C nous définissons la relation *Immersed*.

Définition 4.15 (Relation *Immersed*) *La relation *Immersed* est définie entre l'ensemble des esprits M et l'ensemble des cultures C de la manière suivante :*

$$\forall m \in M, \forall c \in C \quad \textit{Immersed}(m, c) \Leftrightarrow \textit{l'esprit } m \textit{ est plongé dans la culture } c$$

Pour une culture, nous distinguons trois types de connaissances partagées :

1. Les données partagées : des données exprimables sous la forme de variables qui appartiennent non pas aux connaissances individuelles mais à une culture.
2. Les procédures comportementales partagées : des modèles de comportement qui sont appliqués par tous les individus d'une même culture dans certaines situations données.
3. Les interprétation collectives : des mécanismes d'interprétations des phénomènes qui se produisent dans la réalité brute qui ne sont pas propres à un unique individu mais qui sont collectivement acceptés dans une culture.

Pour illustrer l'utilisation des données partagées et des procédures comportementales partagées nous considérons deux exemples de modélisation possibles de comportement

instinctif chez les animaux. Il ne s'agit bien ici que d'un choix de modélisation possible, et non d'un choix épistémologique sur l'origine des instincts.

Exemple 4.6 – Exemple de donnée partagée.

Chez certaines espèces d'oiseaux migrateurs, tous les individus effectuent à une période donnée une migration vers une destination commune. Un choix de modélisation possible serait donc de coder cette destination au niveau de la culture de ces individus.

Exemple 4.7 – Exemple de procédure comportementale partagée.

Par ailleurs, certaines espèces d'animaux ont un rite de reproduction bien défini. Ce rite peut être modélisé par une procédure comportementale qui est définie au niveau de la culture.

Pour décrire des formes d'interprétation collective, nous utilisons la notion d'*institution* telle qu'elle a été définie par Searle [Searle, 1997] (cf. section 2.4.7). Une culture peut contenir plusieurs institutions, chaque institution étant définie par un ensemble de règles constitutives et de règles régulatrices. Nous donnerons la formalisation de ces règles dans le méta-modèle MASQ dans la prochaine section.

A ce stade nous pouvons donner la définition formelle d'une culture dans le méta-modèle MASQ.

Définition 4.16 (Culture) Une culture est un tuple

$c = \langle SData_s, SProcs, CLaws, RLaws \rangle$ avec :

- $SData_s$: un ensemble de données partagées ;
- $SProcs$: un ensemble de procédures partagées ;
- $CLaws$: un ensemble de règles constitutives ;
- $RLaws$: un ensemble de règles régulatrices.

4.6.1 Interprétation collective

Selon Searle, une institution permet de construire une réalité institutionnelle à partir de la réalité brute et/ou d'autres réalités institutionnelles. Pour une institution, la construction de sa réalité institutionnelle est définie par des *règles constitutives* et des *règles régulatrices*. Les règles constitutives permettent de donner du sens à un fait brut ou un fait provenant d'une autre réalité institutionnelle : « qu'est-ce qui compte pour

quoi ». Les règles régulatrices permettent de donner une caractérisation normative aux faits institutionnels : « comment les choses devraient être ». La distinction entre ces deux types de règles est essentielle. Ce sont les règles constitutives qui permettent la création d'une réalité institutionnelle, c'est-à-dire une description institutionnelle des phénomènes³. Les règles régulatrices quant à elles permettent de décrire ce qui est interdit, obligatoire ou permis relativement à cette description institutionnelle.

Ontologie et description de la réalité brute

Nous avons vu que les règles constitutives permettent de donner une description institutionnelle à des phénomènes de la réalité brute. Pour cela il est nécessaire de pouvoir décrire ces phénomènes, et donc de disposer d'ontologies de la réalité brute. Ainsi à chaque espace brut il doit correspondre une ou plusieurs ontologies qui permettent de décrire le fragment de réalité correspondant. Une même ontologie peut être partagée par différents espaces bruts. Les percepts qui sont émis par un corps vers un esprit sont exprimés suivant ces ontologies.

Pour décrire un phénomène de la réalité brute nous utilisons la notation suivante :

$$bs_1.P(x_1, x_2, \dots, x_i)$$

où $bs_1 \in BS$ est un espace brut et $P(x_1, x_2, \dots, x_i)$ est un prédicat qui est vérifié sur bs_1 .

Ontologie et description d'une réalité institutionnelle

Un système de règles constitutives permet d'établir une ontologie pour décrire une réalité institutionnelle. Si l'on considère un jeu tel que le football, les règles du jeu sont les règles constitutives qui permettent de construire la réalité relative au jeu. Ainsi, les règles du jeu définissent ce qu'est un but, un hors-jeu, une touche, etc. à partir de phénomènes physiques tels que la trajectoire du ballon ou la positions des joueurs. Sur la base des règles du jeu, il est possible d'interpréter la réalité d'un match de football. C'est le fait que cette interprétation soit collectivement acceptée qui permet de jouer au football.

Les règles constitutives permettent aussi d'établir une réalité institutionnelle sur la base d'autres réalités institutionnelles comme l'illustrent les énoncés suivants :

Tels et tels sons comptent pour des mots en français.

³Nous utilisons le terme *phénomène* pour désigner aussi bien des entités, des actions, des événements ou des états des choses.

Tels et tels mots comptent pour une promesse.

Telle promesse dans un certain contexte compte pour un mariage.

Chaque énoncé est l'expression d'une règle constitutive. Le premier énoncé se fonde sur un phénomène brut. Les deux derniers énoncés expriment la constitution de nouveaux faits institutionnels à partir de faits institutionnels existants; ils expriment comment construire une nouvelle réalité institutionnelle à partir d'une réalité institutionnelle existante.

Pour décrire un fait institutionnel relatif à une institution i_1 nous utilisons une notation similaire à celle pour décrire un phénomène de la réalité brute :

$$i_1.P(x_1, x_2, \dots, x_i)$$

Règle constitutive

En utilisant ces notations, il est possible de décrire des règles constitutives comme cela est illustré sur les deux exemples suivants.

Exemple 4.8 – Règle constitutive de la monnaie.

Soit une institution *money* qui comporte la règle constitutive « tel bout de papier *compte pour* un billet de banque de 5 euros ». En utilisant la notation que nous venons de présenter, nous exprimons cette règle constitutive par :

$$bs.DescBill5E(obj) \Rightarrow money.Bill5E(obj)$$

Le prédicat *DescBill5E* est la description physique d'un billet de banque. Le prédicat *Bill5E* caractérise le statut de billet de banque dans la réalité engendrée par l'institution *money*.

Exemple 4.9 – Règles constitutives d'une vente aux enchères.

Considérons le cas d'une vente aux enchères. Usuellement, lorsqu'une enchère est en cours, le fait de lever la main permet de manifester sa volonté de renchérir par rapport au prix actuel. Cette convention est une règle constitu-

tive qui s'exprime par :

$$\left. \begin{array}{l} iauction.AuctionPending \wedge \\ auctionroom.Human(x) \wedge \\ auctionroom.RaisedHand(x) \end{array} \right\} \Rightarrow iauction.MakeBid(x)$$

avec :

- *iauction* l'institution des enchères ;
- *auctionroom* l'espace brut dans lequel se déroule l'enchère ;
- *Auctionpending* le fait que l'enchère est en cours ;
- *Human(x)* le fait que x soit un objet de type humain ;
- *Raisedhand(x)* le fait que x ait la main levé ; et
- *MakeBid(x)* le fait que x surenchérit.

Sur le même exemple, nous pouvons donner une illustration d'un pouvoir institutionnel. Considérons qu'une personne possède le statut de commissaire-priseur (dénnoté par le prédicat *iauction.Auctioneer(x)*) et que l'enchère est considérée comme ouverte lorsque le commissaire la déclare l'enchère ouverte. Ce pouvoir institutionnel est donné par la règle constitutive suivante :

$$\left. \begin{array}{l} iauction.Auctioneer(x) \wedge \\ auctionroom.Tell(x, "declare open") \end{array} \right\} \Rightarrow iauction.AuctionPending$$

Définition 4.17 (Règle constitutive) Une règle constitutive est une expression de la forme $X \Rightarrow Y$, où

- X : La description de faits bruts et/ou de faits institutionnels.
- Y : La description d'un fait institutionnel.

Règle régulative

Le deuxième aspect d'une interprétation collective est l'interprétation déontique. Sur la base d'une description institutionnelle, les règles régulatrices permettent de décrire ce qui est considéré comme obligatoire ou interdit pour une institution.

Exemple 4.10 – Exemple de règle régulative d'une vente aux enchères.

Considérons l'extension suivante pour l'exemple de la vente aux enchères. Lorsqu'une enchère est ouverte, il est interdit aux acheteurs de communiquer.

Nous exprimons cette contrainte par la règle régulative suivante :

$$\begin{aligned} & iauction.AuctionPending \wedge iauction.Buyer(x) \\ & \rightarrow \mathbf{I}(iauction.Communicate(x)) \end{aligned}$$

avec \mathbf{I} qui dénote l'opérateur d'interdiction.

Une règle constitutive, qui n'est pas donnée ici, devra définir ce qui compte pour une communication pour l'institution de l'enchère.

Définition 4.18 (Règle régulative) *Une règle régulative est une expression qui attribue un caractère déontique à une description institutionnelle.*

Cette définition est volontairement très générale car nous ne souhaitons pas présupposer un modèle spécifique de représentation des règles régulatrices. Différents formalismes déontiques ont d'ores et déjà été proposés pour formaliser des normes qui sont des règles régulatrices. Notre objectif, ici, est simplement d'exprimer dans le méta-modèle MASQ la possibilité de décrire des règles régulatrices et de mettre en évidence comment ces règles sont reliées à l'activité dans le système. En pratique, il sera bien évidemment nécessaire d'exploiter un formalisme précis de description de normes.

4.6.2 Perception d'une réalité institutionnelle

Dans cette section nous considérons les moyens dont peuvent disposer les esprits pour accéder à une interprétation institutionnelle.

Le travail de Searle [Searle, 1997] est fondamental dans notre approche car il permet de comprendre comment une réalité institutionnelle peut être établie à partir de la réalité brute et des règles constitutives. Cependant Searle n'a pas traité dans ce travail du rapport des individus à une réalité institutionnelle, c'est-à-dire comment les individus peuvent prendre connaissance des faits institutionnels qui sont établis.

Le modèle AGREEN

Le modèle AGREEN [Báez-Barranco *et al.*, 2006] a fourni une première solution à ce problème. Avant de présenter l'approche qui est proposée dans ce modèle, il est important de souligner la différence entre AGREEN et MASQ car si elle peut paraître subtile, elle n'en est pas moins fondamentale. Dans AGREEN il n'y a pas une distinction

entre réalité brute et culture, mais il y a une distinction entre espace physique et espace social. Dans la terminologie AGREEN les espaces physiques contiennent des objets et des corps physiques et les espaces sociaux contiennent des objets sociaux et des rôles. Ces deux types d'espaces étant des contextes pour l'action et la perception des agents (au travers des corps et des rôles). Le modèle AGREEN prévoit la possibilité de définir des règles constitutives de type *counts-as* entre un espace physique et un espace social, certaines activités dans un espace physique pouvant alors produire des effets de manière causale dans différents espaces sociaux.

Dans cette approche les moyens qui sont donnés aux agents pour percevoir une réalité institutionnelle sont définis par la notion de rôle. Une réalité institutionnelle étant décrite dans un espace social, il suffit qu'un agent possède un rôle dans cet espace pour pouvoir en avoir une perception.

Nous pensons que cette approche soulève plusieurs problèmes. Tout d'abord, une réalité institutionnelle est entièrement déterminée par l'application des règles constitutives sur la réalité brute. Ainsi le fait de matérialiser une réalité institutionnelle constitue une redondance d'information, ce qui engendre inévitablement des coûts de maintenance pour préserver la cohérence entre la réalité brute et cette réalité institutionnelle.

D'autre part, et c'est de notre point de vue le problème le plus important, la perception que peut avoir un agent au travers de ses rôles est découplée de son activité physique. En pratique, un agent peut percevoir un fait institutionnel dans un espace social par l'intermédiaire d'un de ses rôles, sans que son corps lui ne soit capable de percevoir le phénomène physique qui en est à l'origine. Ce mécanisme viole le principe de localité de la perception. Considérons un exemple concret pour illustrer ce problème.

Exemple 4.11 – Le problème de la perception d'une réalité institutionnelle.

En France, le résultat d'une élection présidentielle a pour effet d'attribuer le statut institutionnel de président de la république. Ainsi du moment que le décompte des voix est effectif, cette attribution de statut est aussi effective dans la réalité institutionnelle.

Considérons maintenant un citoyen français qui se trouve isolé au moment du résultat de l'élection présidentielle dans un lieu où il ne dispose d'aucun moyen de communication. Ce citoyen ne peut pas avoir connaissance du résultat de l'élection présidentielle.

Ce que nous voulons illustrer sur cet exemple, c'est qu'un agent ne peut pas avoir une

connaissance intégrale d'une réalité institutionnelle, et que cette connaissance qu'il peut en avoir est contrainte par son activité physique.

Une approche fondée sur la perception de la réalité brute

Nous adoptons donc une démarche différente de celle proposée dans le modèle AGREEN en considérant que la connaissance qu'un agent peut avoir d'une réalité institutionnelle est directement conditionnée par sa perception du monde brut. Ainsi de même que les agents ne peuvent avoir qu'une représentation partielle du monde brut dans lequel ils évoluent, ils ne peuvent avoir qu'une représentation partielle d'une réalité institutionnelle.

Dans MASQ, une culture constitue une connaissance partagée par tous les esprits qui appartiennent à cette culture. Pour respecter le principe de localité de la perception dans le monde brut, nous proposons de ne pas placer les faits institutionnels au niveau d'une culture mais uniquement les règles d'une institution, c'est-à-dire les règles constitutives et les règles régulatrices. Rapporté à l'exemple de l'élection de la république, cela revient à dire que la règle constitutive « x remporte l'élection présidentielle compte pour x est président de la république » est une connaissance partagée ; par contre le fait que *untel* soit président de la république n'est pas une connaissance partagée, elle n'est pas située dans la culture.

Pour un agent, la connaissance d'un fait institutionnel nécessite une expérience dans la réalité brute qui permet de déduire ce fait institutionnel d'après les règles constitutives.

Pour prendre un autre exemple, la culture nous décrit ce qu'est un billet de banque, mais elle ne décrit pas qu'un individu tient un billet de banque dans la main. Ceci relève de l'interprétation d'une situation d'après la culture.

Pour résumer ce qui vient d'être dit, nous considérons que les mécanismes d'interprétations fournis par une institution relève de la connaissance collective, mais qu'une interprétation donnée de la réalité brute relève de la connaissance individuelle.

On peut faire un parallèle entre la perception qu'un agent a du monde brut et celle qu'il a d'une réalité culturelle. Les connaissances d'un agent sur son environnement sont généralement exprimées en termes de croyances ; en notant B_i l'opérateur modal associé aux croyances d'un agent i , on exprime par $B_i(x)$ le fait que l'agent i croit qu'une propriété x est vérifiée sur l'environnement. Bien évidemment on a pas l'implication $B_i(x) \rightarrow x$, c'est ce qui traduit le fait qu'un agent peut avoir des croyances erronées sur

son environnement. De la même manière nous considérons qu'un agent ne peut avoir que des croyances sur une réalité institutionnelle. En utilisant l'opérateur D_s formalisé par Jones et Sergot [Jones et Sergot, 1996] pour décrire une réalité institutionnelle, on note alors $B_i(D_s(y))$ le fait qu'un agent i croit que la propriété y est vérifiée dans une réalité institutionnelle s . Et de la même manière que pour la réalité brute, on a pas l'implication $B_i(D_s(F)) \rightarrow D_s(F)$, les croyances d'un agent sur une réalité institutionnelle peuvent donc être erronées.

Il est évident que cette approche permet de préserver le principe de localité de la perception pour les agents car la perception est restreinte au monde brut. La « perception » d'une réalité institutionnelle n'est que le résultat d'une interprétation de percepts sur la réalité brute.

Un autre intérêt de cette approche est de permettre l'explication du phénomène d'une reconnaissance de statut illégitime. Les faux billets sont un exemple représentatif de ce phénomène. Le statut de billet de banque est conféré par une règle constitutive du type « tel morceau de papier compte pour un billet de banque ». Un faux billet ne correspond pas exactement à la description de « tel morceau de papier » de la règle précédente sans quoi il ne serait pas faux ; un faux billet n'a donc pas le statut de billet de banque. Pourtant nous pouvons constater que dans la réalité les faux billets sont parfois reconnus comme des billets de banque, ce qui leur permet de remplir la même fonction. Nous pouvons expliquer ce phénomène dans le méta-modèle MASQ par l'imprécision des percepts d'un agent. La perception qu'a un agent peut être imprécise au point de ne pas pouvoir distinguer un faux billet d'un vrai ; l'agent reconnaîtra par erreur le statut de billet de banque au faux billet.

Nous allons maintenant analyser comment un esprit peut se constituer une représentation des différentes réalités institutionnelles qui le concernent.

4.6.3 Internalisation des règles institutionnelles

Règle formelle et règle informelle

On distingue généralement les règles formelles et les règles informelles. De manière générale une règle formelle est associée à un écrit comme par exemple dans le cas des codes de règles (code pénal, code civil, code de la route, etc.). Les règles formelles ont donc une manifestation dans la réalité brute. Au contraire une règle informelle correspond à une connaissance partagée ou une coutume (comme par exemple la manière de se saluer) qui est admise par les membres d'une culture mais qui n'est pas décrite de manière formelle.

Cette distinction entre règle formelle et règle informelle est exprimable dans le méta-modèle MASQ de la manière suivante.

Une règle formelle est une règle qui est réifiée dans un espace brut, c'est-à-dire que cette règle sera exprimée dans un certain langage et sera supportée par un objet particulier qui modélise un code de règles. Ces règles peuvent donc être perçues directement dans un espace brut, les esprits ont la possibilité de les apprendre.

Au contraire, les règles informelles n'ont pas de contrepartie dans le monde brut, elles n'existent qu'au niveau d'une culture. Pour avoir connaissance de règles informelles, un esprit doit appartenir à la culture dans laquelle elles sont établies. Pour un esprit étranger à une culture donnée, un processus d'apprentissage est nécessaire pour intégrer les règles de cette culture. Cet apprentissage peut être réalisé de différentes manières : par imitation des autres, en observant et en généralisant le comportement des autres, en étant informé des pratiques par des membres de cette culture, ou encore en fonction des récompenses et des pénalités reçues.

Les processus d'acquisition sont donc différents pour les règles formelles et les règles informelles. Dans le cas des règles formelles il suffit de consulter les « documents officiels » alors que le cas de règles informelles il est nécessaire de les découvrir par soi-même.

Processus d'internalisation des règles

Que les règles soient formelles ou informelles, il est nécessaire qu'un esprit les internalise pour être en mesure de se faire une représentation d'une réalité institutionnelle. Ce processus d'internalisation consiste à l'intégration des règles dans les croyances mêmes d'un esprit. Une règle constitutive de la forme $X \Rightarrow_s Y$ est internalisée par un agent i sous la forme $B_i(X) \rightarrow B_i(D_s(Y))$. L'internalisation exprime que si un agent croit la prémisse X d'une règle constitutive, alors il croit que Y est admis pour l'institution s .

Reconsidérons l'exemple de l'élection présidentielle. Les citoyens français ont internalisé la règle constitutive qui attribue le statut de président de la république au vainqueur de l'élection présidentielle. Ainsi pour qu'un citoyen reconnaisse ce statut à une personne, il est nécessaire qu'il croie que celui-ci a remporté l'élection. Et c'est la crédibilité qui est donnée par les citoyens aux médias de communication qui relaient cette information qui permet l'acceptation de ce statut.

Représentation individuelle d'une réalité institutionnelle

Dans notre approche, il n'y a pas de représentation explicite d'une réalité institutionnelle qui soit externe aux esprits. Au contraire, chaque esprit peut s'en faire sa propre représentation d'après les règles qu'il a internalisées et son expérience dans la réalité brute. En conséquence de quoi chaque esprit peut avoir une représentation partielle et erronée d'une réalité institutionnelle, ce qui est le prix à payer pour préserver le principe de localité de la perception. De plus la construction d'une représentation du contexte institutionnel est confiée aux esprits eux-mêmes, et comme les esprits ne peuvent pas être contrôlés depuis l'extérieur, il est possible que certains ne se constituent pas de telles représentations. Ce choix relève de l'autonomie de chaque esprit, sachant que le fait d'avoir une représentation d'une réalité institutionnelle permet pour un esprit de favoriser la coordination avec les autres et de réduire le risque de sanction.

4.6.4 Distinction entre culture et espace brut

Dans le méta-modèle MASQ, la distinction entre un espace brut et une culture est essentielle. Un espace brut permet la réalisation d'une activité alors qu'une culture permet l'interprétation collective d'une activité. Dans l'environnement, c'est-à-dire dans les espaces bruts, les choses sont telles quelles sont. L'environnement ne fournit aucun jugement de valeur sur les éléments qu'il contient. La culture donne du sens aux phénomènes qui se produisent dans l'environnement.

Les règles qui sont définies pour un espace brut et celles définies pour une culture ont des conséquences de nature très différente.

Les règles d'un espace brut conditionnent son évolution, ce qu'un agent peut faire, ce qu'il ne peut pas faire, ainsi que les conséquences de ce qu'il fait. Ces règles s'imposent aux agents.

Les règles d'une culture sont de nature descriptive, elles n'ont pas d'impact direct sur la réalité brute. Elles ne modifient pas les possibilités d'action d'un agent dans un espace brut. Ainsi qu'un acte soit considéré comme « bon » ou « mauvais » ne change pas la possibilité de l'effectuer, pas plus que cela ne change les conséquences de l'acte dans l'environnement. Par contre si la culture ne modifie pas directement la réalité brute, elle influence les esprits dans leurs décisions, ce qui peut les conduire à adopter des comportements différents, et finalement faire évoluer la réalité brute d'une manière différente. Il y a donc un lien indirect entre cultures et espaces bruts.

Pour prendre un exemple commun, l'obligation de rouler à droite en France est bien une réalité, et cette obligation nous influence tous dans notre comportement sur la route, cependant rien ne nous empêche physiquement de rouler sur la gauche.

Brut ou culturel ? un choix de modélisation

Durant la phase de modélisation d'un système il est nécessaire de déterminer ce qui doit être modélisé au niveau de la réalité brute et ce qui doit l'être au niveau de la culture. Rien n'impose que des caractéristiques culturelles ou institutionnelles dans les sociétés humaines ne soient modélisées au niveau de la réalité brute d'un SMA. En ce qui concerne l'environnement naturel, les lois physiques sont données, la création d'institution est donc le seul moyen d'établir de nouvelles activités qui ne dépendent pas des lois physiques. Ce n'est pas le cas pour un environnement artificiel car dans ce cas les lois sont établies par son concepteur. Par exemple, pour mettre en place un système de vote, les humains doivent mettre en place une institution (par exemple : lever la main compte pour un vote) car l'action de voter n'est pas naturelle ; par contre un environnement artificiel peut établir des capacités qui permettent aux agents de voter « de manière brute ».

Finalement le choix de modéliser un aspect du système au niveau de la réalité brute ou au niveau d'une culture dépend des besoins du système relativement à cet aspect. Le fait d'exploiter la réalité brute permet d'établir certaines garanties, pour favoriser la sécurité par exemple. Par contre la modélisation au niveau de la réalité brute présente un inconvénient : elle réduit les possibilités d'adaptation du système. La modélisation à ce niveau ne permet pas aux agents de contourner les règles ; ce choix est donc intéressant uniquement si il est possible de prévoir *a priori* le résultat que l'on attend du système relativement à l'aspect considéré. Si dans certaines circonstances les lois qui ont été établies se révèlent être inadaptées, il ne sera pas possible pour les agents de les contourner. Le choix de modéliser un aspect du système au niveau d'une culture est une alternative qui permet de favoriser l'adaptation au prix d'une perte du déterminisme. Les agents peuvent ne pas respecter les lois d'une institution, ils peuvent donc adopter un comportement différent de celui qui est attendu dans une situation donnée. C'est cette possibilité qui permet au système global de s'adapter à des situations qui n'ont pas été prévues à la conception. Cette solution est donc intéressante lorsque les comportements à adopter ne sont prédéterminables pour toutes les situations et que l'on souhaite faire intervenir la rationalité des agents « en situation ».

4.6.5 Discussion

La dimension culturelle est sans aucun doute un des aspects les plus complexes à mettre en œuvre dans un SMA. De plus l'intérêt pour cette dimension est relativement récent par rapport aux autres dimensions. La formalisation de la culture que nous proposons avec le méta-modèle MASQ est un premier pas vers une exploitation de cette dimension.

Nous allons conclure cette section par deux points sur lequel nous ne donnons pas de réponse tranchée.

Le changement culturel

Nous n'avons pas traité dans ce travail du problème de la création dynamique et de l'évolution d'une culture. Notre méta-modèle ne donne pas de réponse aux questions suivantes : comment un individu intègre-t-il une nouvelle culture ? Comment une connaissance partagée peut-elle évoluer au cours du temps ? Le problème de l'évolution d'une culture est certainement très complexe et l'analyse spécifique de cette problématique est une des perspectives qui est ouverte par cette thèse.

Holisme *vs.* individualisme méthodologique

Il existe en sociologie deux courants qui s'opposent radicalement en ce qui concerne l'explication des phénomènes sociaux et culturels : l'*individualisme méthodologique* et le *holisme méthodologique*.

L'individualisme méthodologique considère que le niveau social est uniquement le résultat des actions individuelles de ses membres. Il rejette l'existence de structures collectives comme « l'inconscient collectif » de Jung ou la « conscience collective » de Durkheim. Il considère que les réalités institutionnelles ou culturelles n'« existent » pas : ce ne sont que des constructions conceptuelles. S'il y a consensus sur quoi que ce soit au sein d'un groupe, c'est uniquement parce que des individus ont forcé ou convaincu les autres à croire à la même chose. L'élément essentiel dans les comportements sociaux est donc la rationalité de l'individu.

Le holisme méthodologique est essentiellement le contraire. Il considère que les groupes sont constitués de « tendances collectives » ayant une énergie propre qui déterminent les actions des individus. Le social ou la culture sont considérés comme une « mécanique » qui explique le comportement des individus par une causalité.

Le débat entre ces deux conceptions est loin d'être terminé. Ceci étant, quelques approches ont été proposées pour tenter de les réconcilier en prenant une position intermédiaire. Quoi qu'il en soit, en ce qui concerne les SMA, nous ne souhaitons pas effectuer un choix entre ces deux approches, mais au contraire donner la possibilité d'utiliser l'une ou l'autre en fonction des besoins. Ainsi avoir une conception individualiste ou une conception holiste peut être un choix de modélisation selon les caractéristiques mêmes du problème qui est traité.

Dans le méta-modèle MASQ, une conception purement individualiste consiste à considérer la uniquement comme une commodité d'analyse. Au niveau de l'implémentation, les cultures disparaissent : elles sont réduites aux connaissances individuelles. Un esprit n'a donc pas accès à d'autres connaissances que ses propres connaissances.

Une conception holiste pure correspond au contraire à exprimer les trames de comportement au niveau de la culture, les esprits étant réduits à des moteurs de règles qui déterminent leur comportement en appliquant les règles définies au niveau collectif.

Enfin il est possible de combiner des éléments des deux approches. L'esprit dispose d'une certaine autonomie, mais l'esprit est aussi plongé dans une culture qui permet d'établir des connaissances partagées entre plusieurs esprits.

4.7 Synthèse

Dans ce chapitre nous avons présenté le méta-modèle MASQ qui est une formalisation du cadre conceptuel qui a été développé au chapitre 3. Ce méta-modèle permet de décrire séparément les quatre dimensions qui correspondent aux quadrants, les lois qui déterminent leur évolution, ainsi que les relations qui les unissent. Les concepts primitifs sont synthétisés sur la figure 4.13, les lois sur la figure 4.14 et les relations sur la figure 4.15.

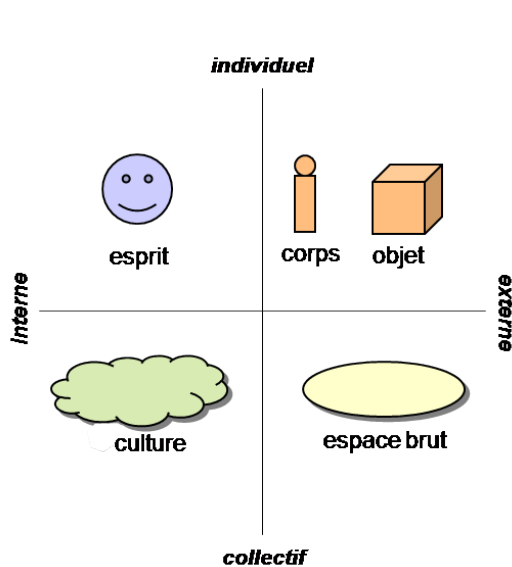


FIG. 4.13 – MASQ : les concepts

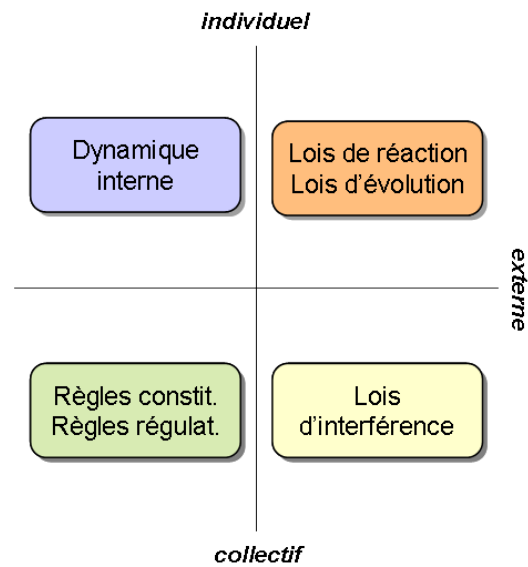


FIG. 4.14 – MASQ : les lois

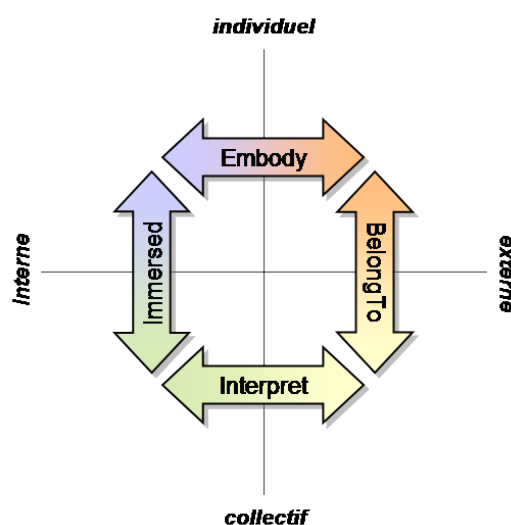


FIG. 4.15 – MASQ : les relations

Le méta-modèle MASQ s'inscrit totalement dans l'approche de conception que nous avons exposé à la section 3.5, il bénéficie donc des avantages de cette approche, à savoir :

- il traite des quatre quadrants ;
- il permet de concevoir séparément les différents éléments ;
- il permet une conception modulaire ;
- il unifie l'activité physique et l'activité sociale ;
- la définition générique des concepts permet d'intégrer des modèles spécifiques ;
- il est adapté à la spécification de système ouvert.

De plus, au travers de MASQ, nous proposons une mise en œuvre de la notion d'institution telle que Searle l'a définie. L'approche que nous avons proposée permet aux agents de se construire une représentation des différentes réalités institutionnelles, ce qui va favoriser leur compréhension mutuelle et donc leur coordination, et cela tout en préservant le principe de localité de la perception.

La figure 4.16 présente le diagramme UML du méta-modèle MASQ. On peut constater sur ce diagramme qu'il n'y a pas d'association entre les classes **Culture** et **Espace brut**. Les espaces bruts et les cultures sont bien en relation dans MASQ, mais par l'intermédiaire des esprits. Ceci est une conséquence des choix que nous avons effectués pour la représentation des réalités culturelles. La culture ne fournit aux agents que des règles d'interprétation qu'ils pourront alors exploiter au cours de leur expérience dans les espaces bruts.

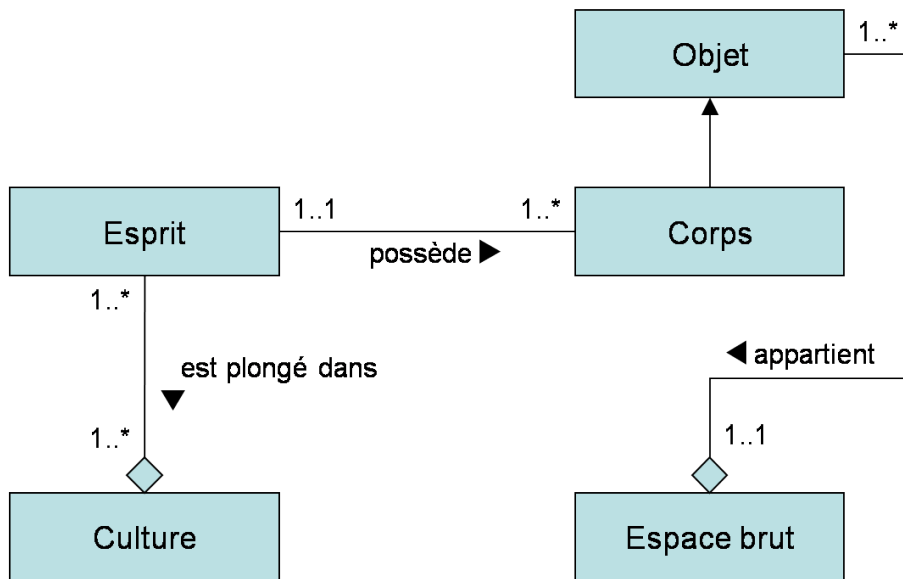


FIG. 4.16 – Diagramme UML du méta-modèle MASQ

Finalement nous pouvons conclure sur le méta-modèle MASQ en indiquant que son intérêt majeur est de donner une spécification précise des concepts primitifs qui respecte les principes que nous avons identifiés comme essentiels d’après l’état de l’art.

Chapitre 5

Modélisation et conception avec MASQ

Ce chapitre est consacré à la modélisation et à la conception de SMA avec le méta-modèle MASQ. Pour illustrer la mise en œuvre de MASQ, nous considérerons deux systèmes différents. Le premier, un système de type wiki, qui supporte la collaboration entre agents humains et qui est de nature « sociale ». Le deuxième, *Warbot* (la guerre des robots), fait intervenir des agents logiciels et comporte un aspect « physique » dans lequel les agents sont situés.

5.1 Modélisation d'un wiki

Un *wiki*¹ est un système de gestion de contenu de site Web qui rend les pages Web librement et également modifiables par tous les visiteurs autorisés. On utilise les wikis pour faciliter l'écriture collaborative de documents avec un minimum de contraintes.

Créé en 2001, Wikipédia² est le représentant le plus populaire de cette classe de site web. Wikipédia est un projet d'encyclopédie collaborative sur Internet. Ce projet a permis de montrer l'intérêt à grande échelle de ce type de système, notamment en ce qui concerne l'autorégulation du système.

Nous avons choisi l'exemple d'un wiki pour illustrer le méta-modèle MASQ car ce type de système possède les caractéristiques d'un SMA ouvert, à savoir :

- les participants ne sont pas connus à l'avance, et de nouveaux participants peuvent joindre le système à tout moment ;

¹Le wiki a été inventé en 1995 par Ward Cunningham, pour une section d'un site sur la programmation informatique qu'il a appelée WikiWikiWeb. Le mot « wiki » vient du redoublement hawaïen wiki wiki, qui signifie « rapide ».

²<http://www.wikipedia.org/>

- le contenu rédactionnel est le résultat de la collaboration entre les différents participants ;
- généralement, l’ensemble du contenu rédactionnel dépasse les compétences individuelles de chaque participant ;
- un wiki est généralement un système très dynamique du fait du nombre et de la diversité des participants (par exemple, des informations d’actualités apparaissent très rapidement dans l’encyclopédie Wikipédia) ;
- la modération du contenu rédactionnel est le résultat d’une autorégulation entre les participants, il s’agit d’une forme de contrôle social.

Les utilisateurs de wikis sont encore très majoritairement des utilisateurs humains, aussi nous ne considérerons dans cet exemple que des agents humains. Cependant nous pouvons noter que des *bots* sont d’ores et déjà utilisés par Wikipédia pour des tâches telles que la correction orthographique, la suppression de vandalisme grossier ou encore l’importation automatique d’information depuis des sources fiables.

5.1.1 Description du wiki

Nous allons maintenant décrire le wiki que modéliserons par la suite avec le méta-modèle MASQ. Le diagramme UML de ce wiki est présenté sur la figure 5.1.

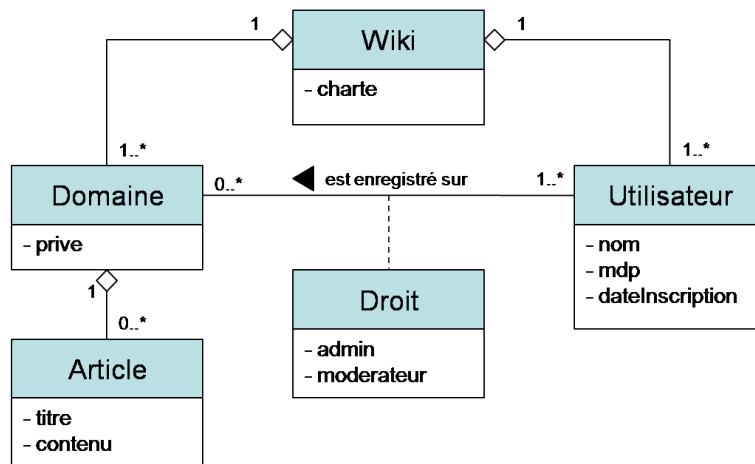


FIG. 5.1 – Diagramme UML du wiki

Domaine

Le contenu rédaction du wiki sera structuré en domaine. Un domaine peut être relatif à une thématique particulière (par exemple *Informatique*) ou dédié à un groupe d’utilisateurs spécifique (par exemple l’équipe SMA du LIRMM). De plus, un domaine peut être public, c’est-à-dire accessible à des utilisateurs non authentifiés ; ou privé, dans ce

cas seuls les utilisateurs enregistrés sur ce domaine ont la possibilité d'y accéder. Dans chaque domaine, le contenu rédactionnel est composé d'un ensemble d'articles.

Gestion des utilisateurs

Le wiki comporte un ensemble de comptes utilisateurs. Chaque utilisateur peut s'authentifier depuis la page d'accueil (ou portail) du wiki à l'aide de son nom et de son mot de passe. Les utilisateurs qui ne sont pas authentifiés n'ont accès qu'aux domaines publics.

Pour chaque domaine, nous considérons trois types de droit :

- *droit d'accès* : permet à l'utilisateur d'entrer sur le domaine ;
- *droit de modération* : permet à l'utilisateur de modérer le contenu rédactionnel de ce domaine ;
- *droit d'administration* : permet à l'utilisateur de modifier les droits des autres utilisateurs pour ce domaine.

Fonctionnalités de gestion du contenu rédactionnel

Le contenu rédactionnel est propre à chaque domaine, mais les fonctionnalités offertes sont identiques pour tous les domaines, à savoir :

- la recherche d'article ;
- la consultation d'article ;
- l'édition d'article (cette fonctionnalité peut être contrainte pour un article par les modérateurs).

Fonctionnalités de modération

Les fonctionnalités supplémentaires qui sont données aux modérateurs sont les suivantes :

- le verrouillage d'article (seuls les modérateurs peuvent éditer un article verrouillé) ;
- le déverrouillage d'article.

Charte

Enfin le wiki dispose d'une charte d'utilisation dans laquelle sont consignées les règles d'utilisation et de modération. Ce document étant destiné à des utilisateurs humains il est rédigé dans le langage naturel. Il comporte un ensemble de règles telles que « il est interdit d'utiliser des propos injurieux » ou encore « il est obligatoire de placer un minimum de références bibliographiques dans un article ».

Cas d'utilisation

La figure 5.2 représente les cas d'utilisations pour ce wiki. Les différentes fonctionnalités ont été regroupées en quatre catégories : les fonctionnalités offertes depuis le portail ou la page d'accueil, les fonctionnalités relatives à la gestion des comptes utilisateurs, les fonctionnalités offertes au sein d'un domaine et les fonctionnalités spécifiques à l'édition d'article.

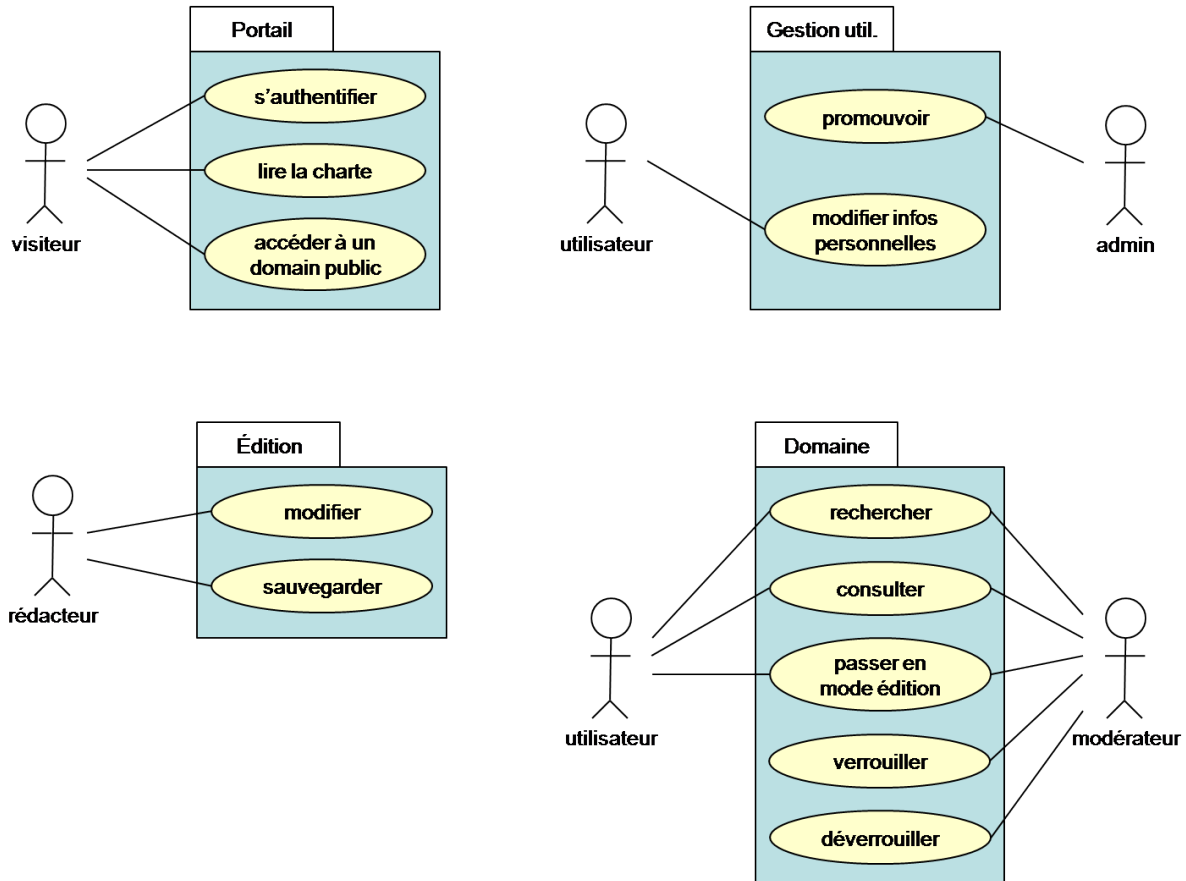


FIG. 5.2 – Les cas d'utilisations du wiki

5.1.2 Modélisation avec MASQ

Considérons tout d'abord le wiki du point de vue des quatre quadrants :

Interne – Individuel : dans ce scénario, les esprits sont des utilisateurs humains ; l'interface web doit donc leur offrir la possibilité d'une part d'émettre des influences, et d'autre part d'avoir une représentation des perceptions qui sont délivrées par le système ;

Externe – Individuel : nous considérons depuis cette perspective les capacités individuelles qui sont données aux utilisateurs ; celles-ci dépendent pour un domaine des droits de l'utilisateur (accès, modération, administration) et du contexte d'interaction dans lequel se trouve l'utilisateur (page d'accueil, gestion des utilisateurs, consultation, édition) ; nous considérons aussi depuis cette perspective les propriétés qui caractérisent les objets d'interactions du système (charte, compte utilisateur, article) ;

Externe – Collectif : de ce point de vue, nous considérons la structuration du système ; dans ce scénario, la structuration a trois fonctions : 1) distinguer les différents modes d'interaction offerts par le système (page d'accueil, gestion des utilisateurs, consultation et modération dans un domaine, édition), 2) structurer le contenu rédactionnel en domaine (il faut notamment garantir la sécurité des domaines privés), 3) assurer la synchronisation nécessaire à l'édition concurrente d'un même article ;

Interne – Collectif : nous pouvons distinguer deux origines aux aspects culturels qui jouent un rôle dans le cas d'un wiki : 1) la culture propre au système : cette culture est décrite par la charte d'utilisation du système ; 2) les cultures qui sont propres aux utilisateurs comme par exemple le langage utilisé ou encore la culture scientifique ; ces cultures permettent aux utilisateurs de donner un sens au texte contenu dans les articles.

Pour ce scénario, les utilisateurs étant humains, nous ne donnerons pas de concrétisation des aspects culturels hormis la réification de la culture du wiki dans la charte d'utilisation.

Représentation MASQ du wiki

La figure 5.3 illustre une représentation du wiki avec les concepts de MASQ. Sur cette figure, on peut voir trois utilisateurs (X , Y et Z). Ces utilisateurs sont plongés dans différentes cultures (la culture du wiki et la culture scientifique sur cet exemple). Sur la partie droite de la figure sont représentés les objets, les corps et les espaces bruts qui sont mis en jeu pour modéliser le wiki. Les hexagones contiennent les types d'objet et de

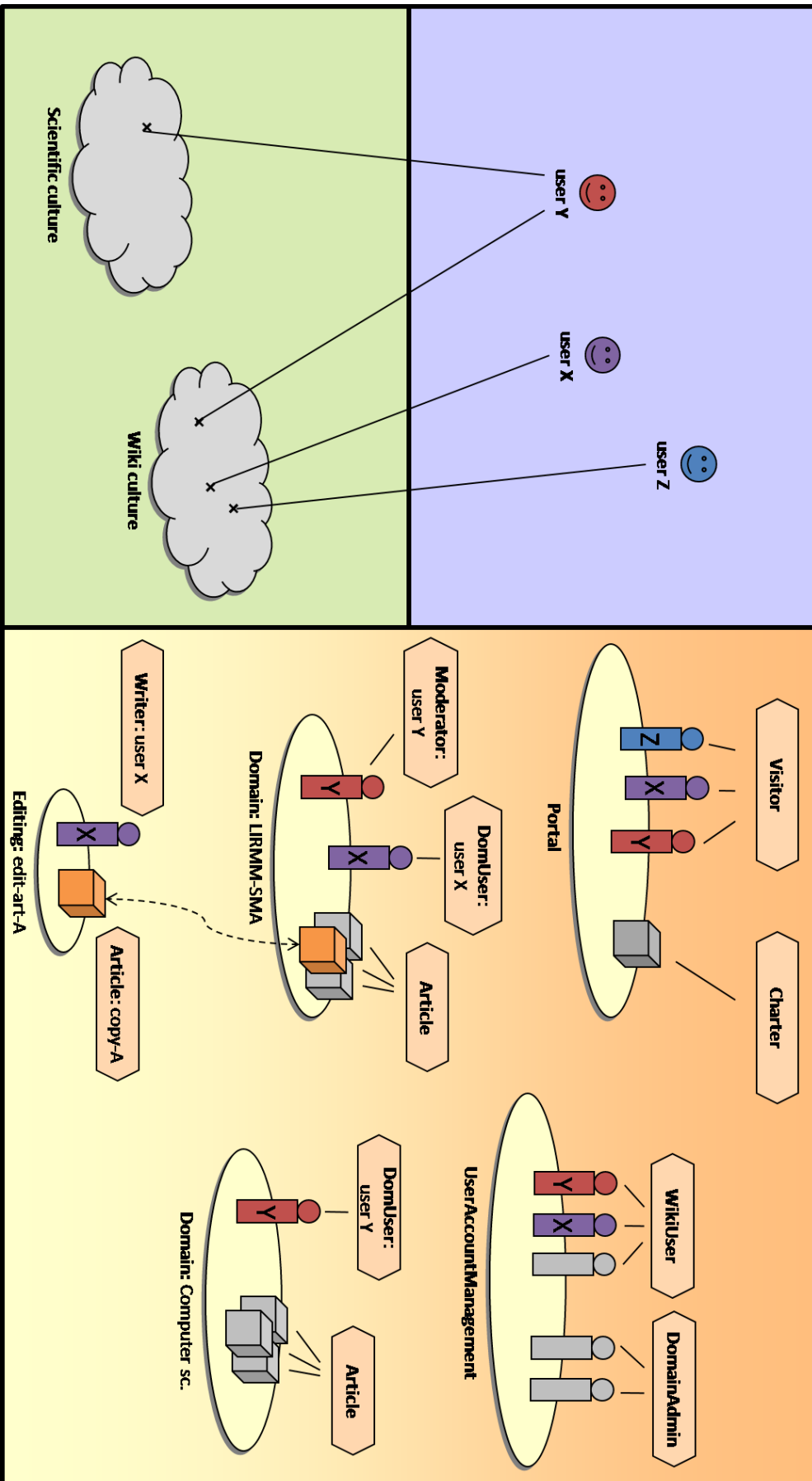


FIG. 5.3 – Représentation MASQ du wiki

corps, complété quand cela est nécessaire par l'identifiant de l'instance. Pour exprimer la relation d'incarnation entre un esprit et son corps nous avons utilisé un code de couleur : l'esprit est incarné dans tous les corps de même couleur que lui. Pour rendre la figure plus lisible, nous avons fusionné la représentation des deux quadrants externes, il ne s'agit toutefois que d'une commodité de représentation : le quadrant E-I comporte tous les objets et les corps, et le quadrant E-C comporte tous les espaces bruts. D'autre part, ce scénario mettant en jeu des utilisateurs humains, seul les éléments des quadrants externes seront implantés.

Nous donnons tout d'abord une description informelle des types d'espaces bruts qui sont représentés sur la figure avant de définir plus précisément chacun de ces types.

- **Portal** : ce type d'espace brut correspond aux interactions possibles depuis le portail ou la page d'accueil (authentification, lecture de la charte et entrée sur un domaine public) ; ce type d'espace brut n'a qu'une seule instance qui contient un objet pour matérialiser la charte et autant de corps de type **Visitor** qu'il y a d'utilisateurs connectés au système ;
- **UserAccountManagement** : ce type offre les fonctionnalités relatives à la gestion des comptes utilisateurs ; deux types de corps sont distingués : **WikiUser** et **DomainAdmin** ; dans MASQ les corps sont des objets, la modification des droits d'un compte utilisateur peut donc être modélisée comme le résultat d'une interaction entre un corps de type **DomainAdmin** et un corps de type **WikiUser** considéré comme un objet ;
- **Domain** : ce type correspond aux interactions possibles dans un domaine (recherche, consultation, modération et passage en mode d'édition) ; chaque domaine du wiki sera modélisé par un espace brut de type **Domain** ;
- **Editing** : ce type offre un contexte d'interaction pour l'édition d'un article ; nous faisons ici le choix de modéliser par des d'espaces bruts différents la consultation et l'édition car l'édition pouvant être concurrente, elle requiert des mécanismes de synchronisation entre les différents rédacteurs ; chaque édition d'article sera effectué dans un espace brut dédié de type **Editing**, sur une copie de l'article original ; l'article original ne sera mis à jour (par une influence entre les deux espaces bruts) qu'au moment de la sauvegarde des modifications.

Définition des types d'objet, de corps et d'espace brut

La figure 5.4 donne une représentation des différents types d'espaces bruts, d'objets et de corps en utilisant les stéréotypes UML. Trois types de stéréotypes sont utilisés :

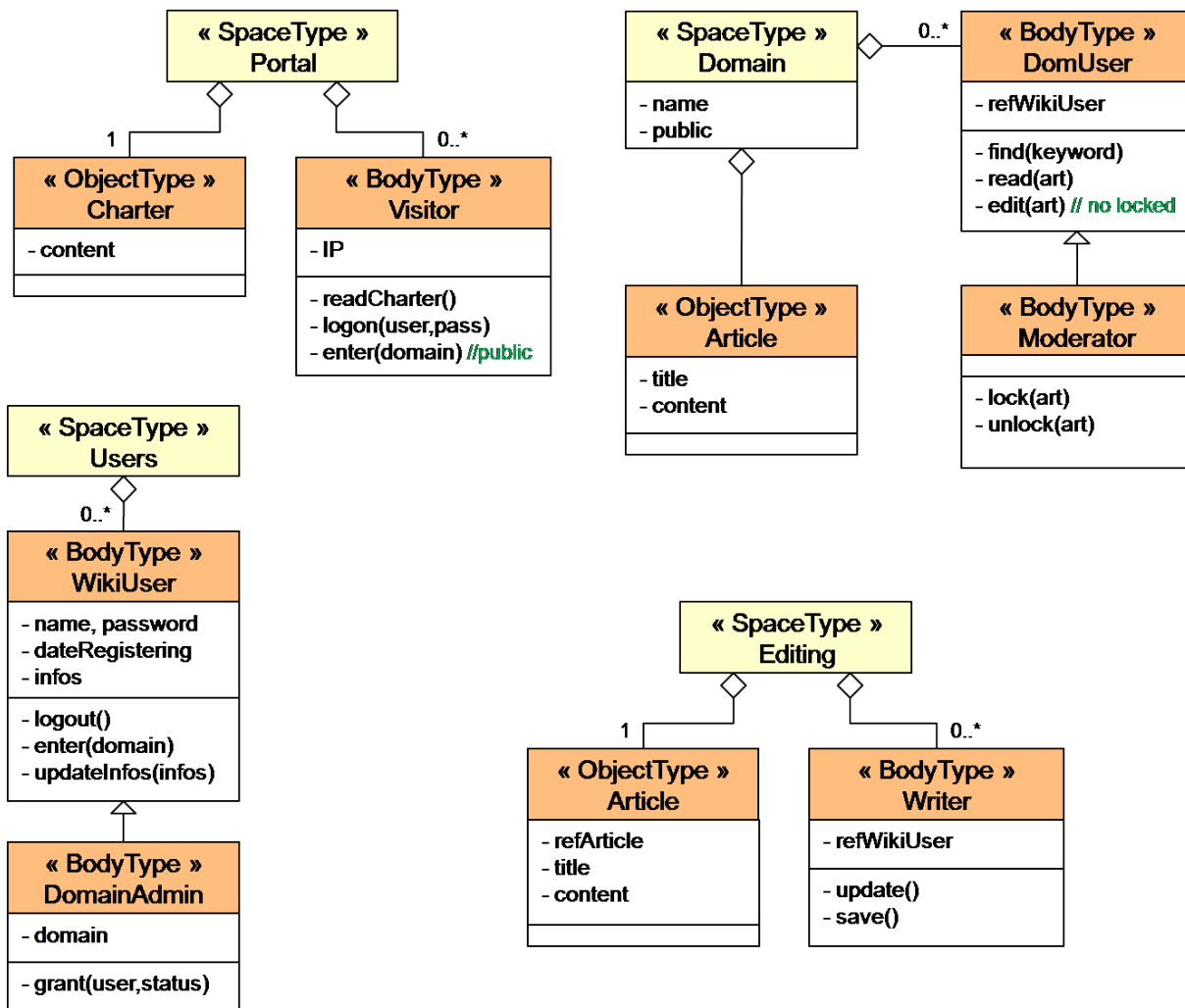


FIG. 5.4 – Diagramme UML des types d’espace brut, de corps et d’objet pour le wiki

```

<TypeDefinition>
  <SpaceType name="Portal">
    <ObjectType name="Charter">
      <Attribute name="content" type="text" />
    </ObjectType>
    <BodyType name="Visitor" public="true">
      <Attribute name="IP" type="string" />
      <Ability name="readCharter" />
      <Ability name="logon" param="user, password" />
      <Ability name="enter" param="domain" /> <!-- available only for public domain -->
    </BodyType>
  </SpaceType>
</TypeDefinition>

```

FIG. 5.5 – Extrait de la définition de type en XML

SpaceType, *ObjectType* et *BodyType*. Ce diagramme permet de représenter les types de corps et d'objets que peut contenir un espace brut d'un certain type, éventuellement avec des contraintes de cardinalités. Pour chacun de ces types sont représentés les attributs qui caractérisent leur état. La spécificité des types de corps est de comporter aussi un ensemble de capacités (notées ici comme des méthodes). Cette description des capacités correspond au nom d'une influence valide qui peut être envoyé par un esprit qui possède un corps de ce type, ainsi que la liste des paramètres qu'il doit fournir dans l'influence.

Enfin la relation d'héritage entre type d'objet/corps est classique, elle permet d'étendre les variables d'états et les capacités du type hérité.

La figure 5.5 présente un extrait de description de type MASQ en XML.

Séquence d'utilisation

Nous allons maintenant décrire les trois séquences d'utilisation du système suivantes :

1. authentification d'un utilisateur (figure 5.6) ;
2. entrée sur un domaine et consultation d'un article (figure 5.7) ;
3. édition d'un article (figure 5.8).

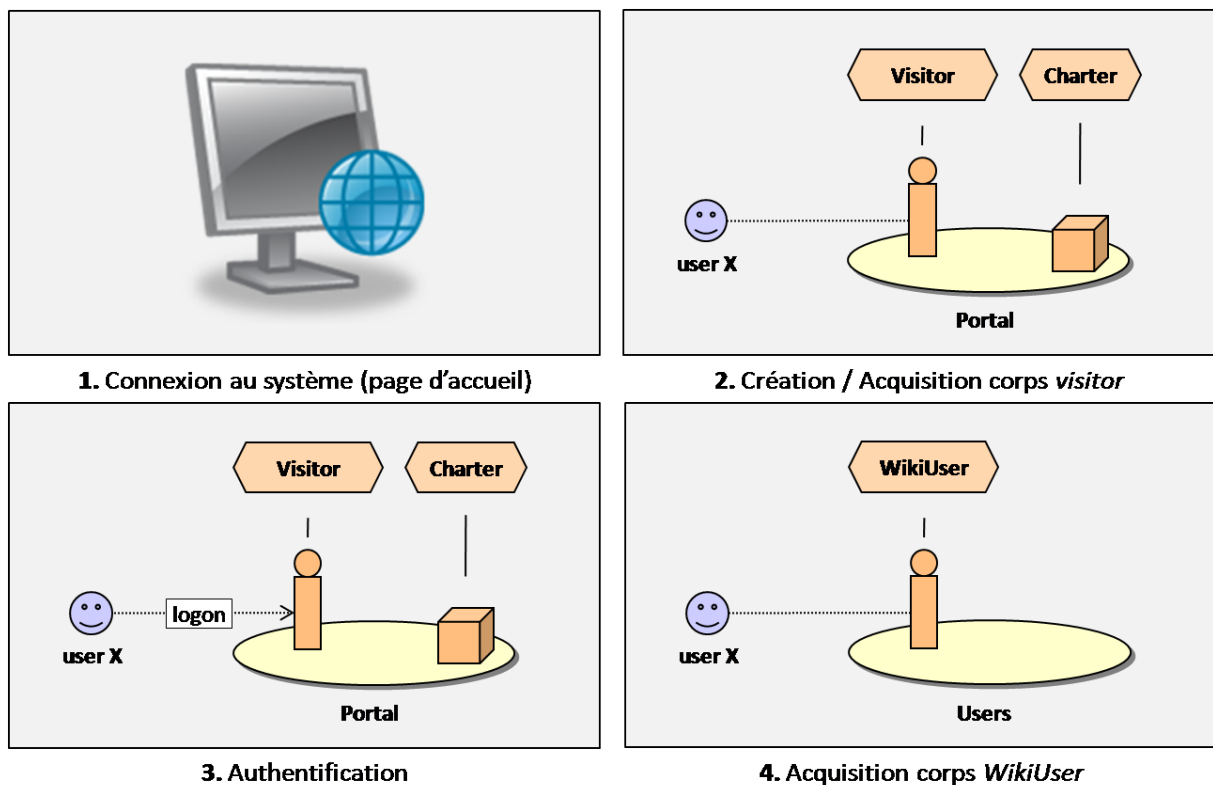


FIG. 5.6 – Séquence d'authentification

La figure 5.6 illustre la première séquence. Tout d'abord l'utilisateur doit se connecter au système via son navigateur (1). Cette étape engendre automatiquement la création d'un

corps de type `Visitor` dans l'espace brut `Portal` qui lui est affecté (2). L'utilisateur peut alors transmettre l'influence `login` pour s'authentifier (3). Si les informations transmises sont valides, le corps correspondant à son compte utilisateur dans l'espace brut `Users` lui sera affecté (4). Ce corps n'est pas créé car il est maintenu continuellement par le système pour stocker toutes les informations relatives au compte utilisateur.

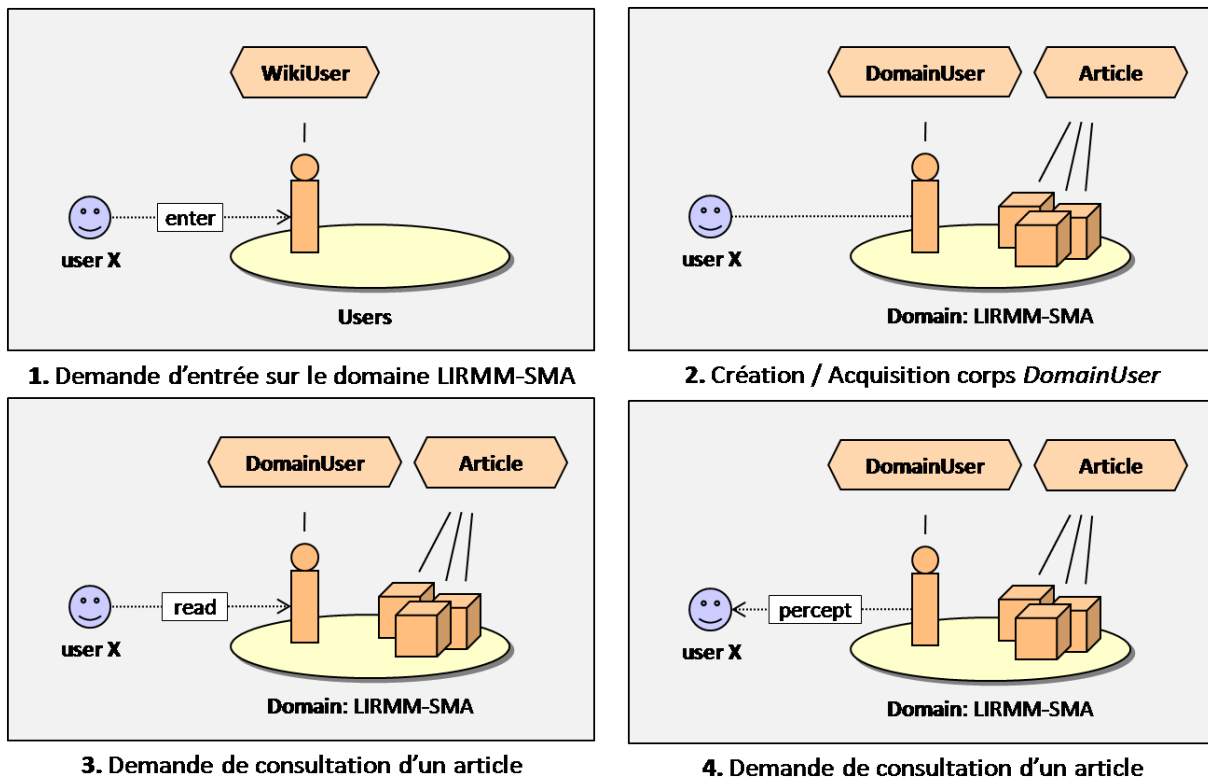


FIG. 5.7 – Séquence de consultation d'un article

La figure 5.7 illustre une séquence d'entrée sur un domaine et de consultation d'un article. Pour entrer sur un domaine, l'utilisateur émet une influence de type `enter` vers son corps d'utilisateur du wiki (1). Si ce corps possède la capacité d'entrer sur le domaine requis (i.e. que l'utilisateur possède le droit d'accès pour ce domaine), un corps d'utilisateur de domaine est alors créé et affecté à l'utilisateur (2). L'utilisateur peut alors envoyer une influence de type `read` à son corps d'utilisateur de domaine (3). Une perception de l'article sera alors retournée à l'utilisateur (4).

La figure 5.8 illustre la dernière séquence : l'édition d'un article. L'utilisateur envoie tout d'abord une influence de type `edit` vers son corps d'utilisateur du domaine pour passer en mode édition (1). Si l'article n'est pas verrouillé et que personne n'édite actuellement cet article, un espace brut de type `Editing` est alors créé (2). Dans ce nouvel espace brut sont aussi créés un corps de type `Writer` qui est affecté à l'utilisateur ainsi qu'un objet de type `Article` qui est une copie de l'article qui va être modifié. L'utilisateur peut alors exploiter son corps de rédacteur pour effectuer une série de modifications sur

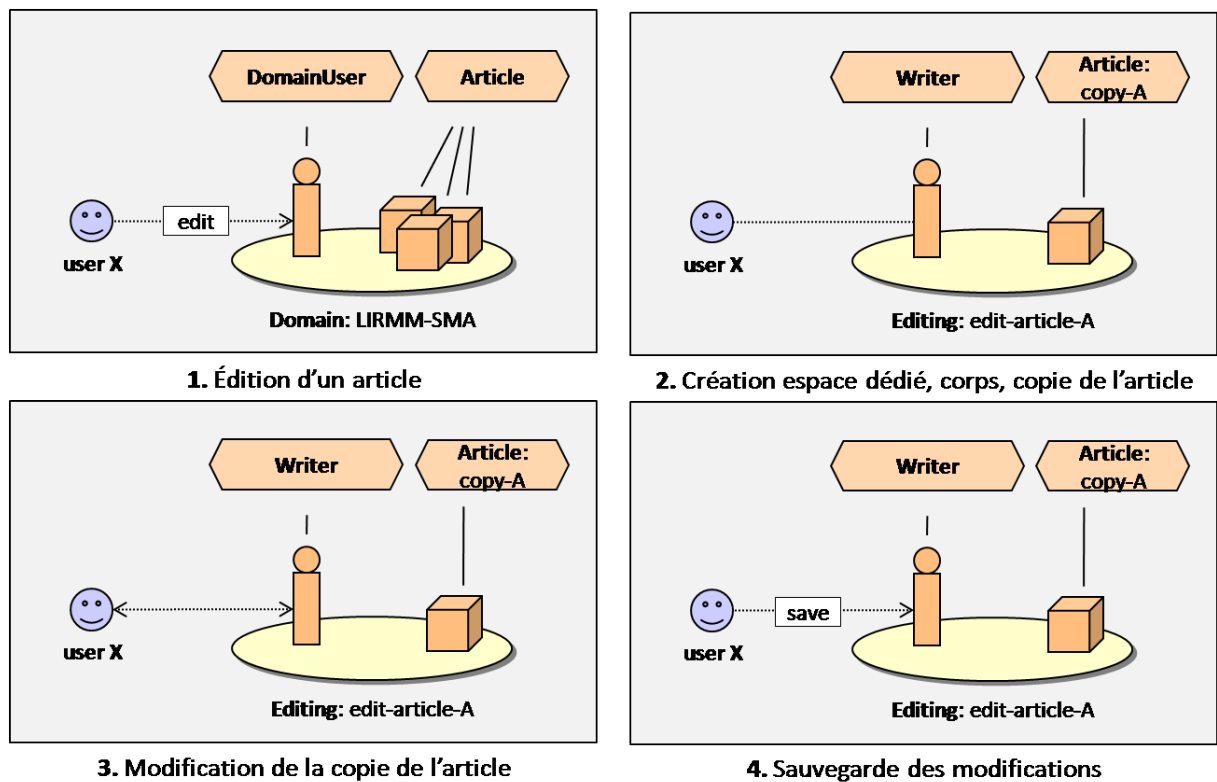


FIG. 5.8 – Séquence d'édition d'un article

l'article et visualiser ces différentes modifications (3). Enfin l'utilisateur peut sauvegarder ses modifications en émettant une influence de type `save` à son corps de rédacteur (4). L'espace brut dédié à l'édition est alors détruit ainsi que tous les objets qu'ils comportent, et la version originale de l'article dans l'espace brut domaine est mise à jour.

Aspects culturels

Sur un wiki, les aspects culturels jouent un rôle dans la rédaction, la consultation et la modération d'articles.

Si l'on prend l'exemple de la culture scientifique, celle-ci définit les critères qui permettent de considérer un article ou un rédacteur comme « scientifique ». Il s'agit dans ce cas d'un statut culturel et non d'un type d'objet social car le fait qu'un article soit ou non reconnu comme scientifique ne modifie en rien les interactions possibles avec celui-ci dans le système. Par contre cela peut affecter les personnes qui l'exploitent (certains scientifiques peuvent souhaiter n'exploiter que des articles scientifiques).

D'autre part, le wiki est associé à une culture spécifique dédiée à son utilisation. Celle-ci est établie par le ou les concepteurs du système et elle est réifiée dans la charte d'utilisation pour être accessible à tous les utilisateurs. Cette culture permet d'exprimer des normes telles que « il est interdit d'utiliser des propos injurieux » ou encore « un article doit contenir un minimum de référence bibliographique ». Cette culture pourra

influencer les utilisateurs et les modérateurs dans leurs démarches sur le wiki. L'usage de cette culture présente un intérêt dans deux situations différentes :

1. lorsque le ou les concepteurs ne sont pas en mesure de prévoir ou de décrire à l'avance toutes les situations nuisibles (par exemple ce qu'est un propos injurieux) ;
2. pour donner plus de souplesse au système afin de pouvoir s'adapter à des situations particulières (par exemple l'ajout d'article sans références bibliographiques pourrait être rendu impossible, mais cela peut s'avérer non pertinent dans certaines situations).

Le fait de placer des contraintes au niveau culturel et non au niveau de l'environnement permet de s'appuyer sur la rationalité des agents plutôt que sur un système automatique qui enrégimente le comportement des agents.

Outre la charte, d'autres éléments environnementaux peuvent être utilisés pour servir de support à une interprétation culturelle. Par exemple, dans Wikipédia, les modérateurs ont la possibilité de placer des *tags* sur les articles tels que : « cet article ne cite pas suffisamment de sources » ou « cet article provoque une controverse de neutralité ». Le fait de pouvoir placer un de ces tags sur un article nécessite d'en avoir la capacité, celle-ci n'étant confiée qu'aux modérateurs ; le système délègue donc ce jugement aux modérateurs. Ici encore, ce type de tag donne une information qui pourra être interprétée différemment selon les utilisateurs mais qui n'affecte pas les interactions possibles avec l'article concerné.

5.1.3 Synthèse

L'exemple du wiki qui vient d'être traité nous a permis d'illustrer l'utilisation que peut être faite des concepts du méta-modèle MASQ sur ce scénario. Notamment nous avons illustré l'utilisation des espaces bruts pour établir la *structure fonctionnelle* du système. Cette structuration permet de distinguer les différents contextes d'interaction du système (accueil, gestion des utilisateurs, consultation, édition) mais aussi d'assurer la sécurité des domaines dont l'accès est restreint. Nous avons illustré l'utilisation des corps pour établir les capacités qui sont données aux différents utilisateurs suivant le contexte et leur rôle dans le système. La définition de l'ensemble des espaces bruts, des corps et des objets permet de délimiter l'autonomie offerte aux agents dans le système. Enfin cet exemple permet d'illustrer les fonctionnalités qui sont mises en place pour permettre une autorégulation du système par un contrôle social (ici celui des modérateurs). Concernant cet aspect spécifique, nous avons illustré l'intérêt de la réification de la culture dans le monde brut avec l'exemple de la charte d'utilisation.

Cet exemple faisant intervenir des agents humains, associés à leurs propres cultures, cet exemple ne permet pas d'illustrer une opérationnalisation des esprits et des cultures.

Nous allons donc maintenant considérer un deuxième scénario qui, lui, fait intervenir des agents logiciels.

5.2 Le projet *Warbot*

5.2.1 Description

Le projet *Warbot* a été réalisé dans le cadre de la thèse de doctorat de [Michel, 2004]; nous reprenons ici la description de ce projet. Il a été initié dans l'objectif de concrétiser la distinction entre le corps et l'esprit d'un agent ainsi que le principe influence / réaction. Ce projet est aussi exploité d'un point de vue pédagogique dans le cadre du *master* recherche d'informatique (anciennement DEA) de Montpellier.

La plate-forme *Warbot* consiste en la simulation de plusieurs équipes de robots qui d'affrontent dans une arène : le but de chaque équipe est de détruire la base adverse. Pour cela, trois types de robots sont définis : les explorateurs, qui se meuvent rapidement et dispose d'un grand champ de perception, les guerriers, qui se déplacent plus lentement mais qui possèdent la capacité de lancer des missiles, et enfin les bases qui ne sont pas capables de se déplacer, mais qui possèdent leur propre champs de perception et ont, comme les autres types de robots, la faculté de communiquer.

La figure 5.9 présente une capture d'écran d'une partie de *Warbot*.

Implémentation actuelle

La plate-forme *Warbot* a été implémentée au-dessus de la plate-forme *Madkit*.

Warbot fournit la simulation du monde physique, c'est-à-dire l'arène de combat, les différents robots ainsi que les objets (missiles et obstacles). *Warbot* a été implémenté en utilisant la séparation du *corps* et de l'*esprit* d'un agent, et le principe *influence / réaction*. Ainsi la composante physique des robots est conçue séparément des esprits (modules comportementaux). Un ensemble de fichiers de configuration permettent, au moment de l'exécution, d'associer un esprit avec un robot dans l'arène. *Warbot* fournit aussi la gestion des perceptions qui sont délivrés par la plate-forme aux esprits. Ces perceptions sont essentielles, non seulement pour que les esprits puissent concevoir une représentation du monde dans lequel ils évoluent, mais aussi pour qu'ils puissent vérifier leur comportement. En effet, suivant le principe *influence / réaction*, une influence émise par un esprit n'a pas nécessairement l'effet escompté. Par exemple, la demande de déplacement d'un esprit n'aura pas d'effet si son corps se trouve face à un obstacle.

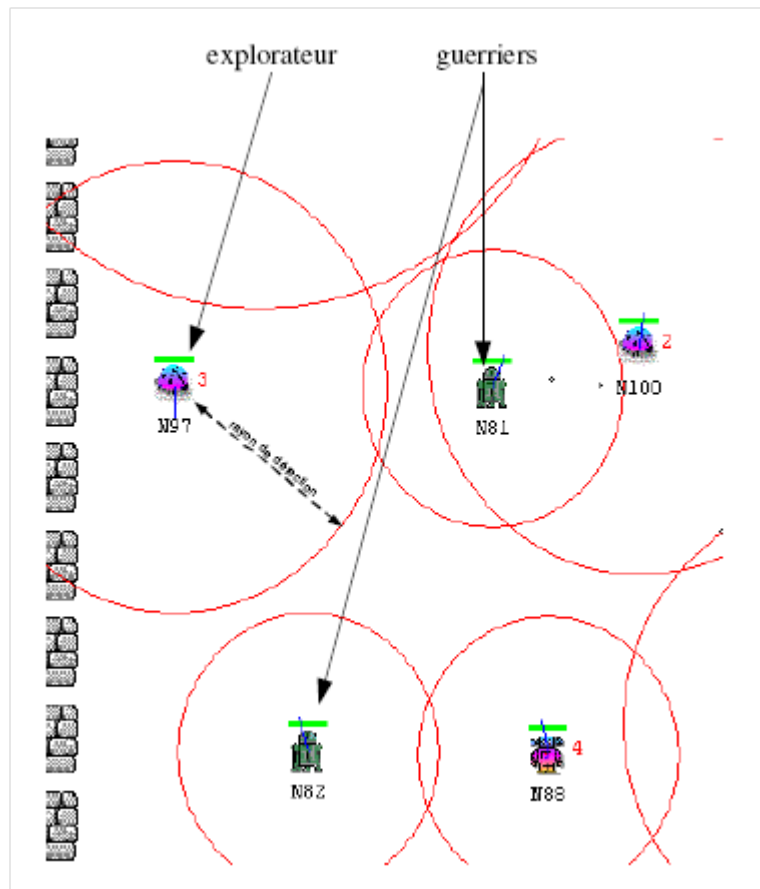


FIG. 5.9 – Capture d'écran d'une partie de *Warbot*

La plate-forme *Madkit* quant à elle fournit le support à l'organisation sociale des différentes équipes suivant le modèle AGR. Chaque équipe peut ainsi créer tous les groupes et les rôles nécessaires à la mise en place de sa stratégie. Ces groupes et ces rôles fournissent alors le support à la communication entre les membres d'une même équipe.

Dans le contexte pédagogique, l'intérêt de *Warbot* est de permettre aux étudiants de concevoir et expérimenter concrètement des stratégies et des mécanismes de coordination entre agents, et cela dans un contexte de compétition. Une fois les différents esprits conçus par des groupes d'étudiants, les équipes peuvent alors s'affronter dans l'arène.

On peut noter que l'un des intérêts de la séparation entre l'esprit et le corps des agents et de permettre l'utilisation de langages différents pour implémenter ces deux modules. Alors que *Warbot* a été implémenté en Java, les esprits peuvent être conçus dans différents langages (Java, Scheme, Python, etc.).

La figure 5.10 illustre l'implémentation du principe *influence / réaction*.

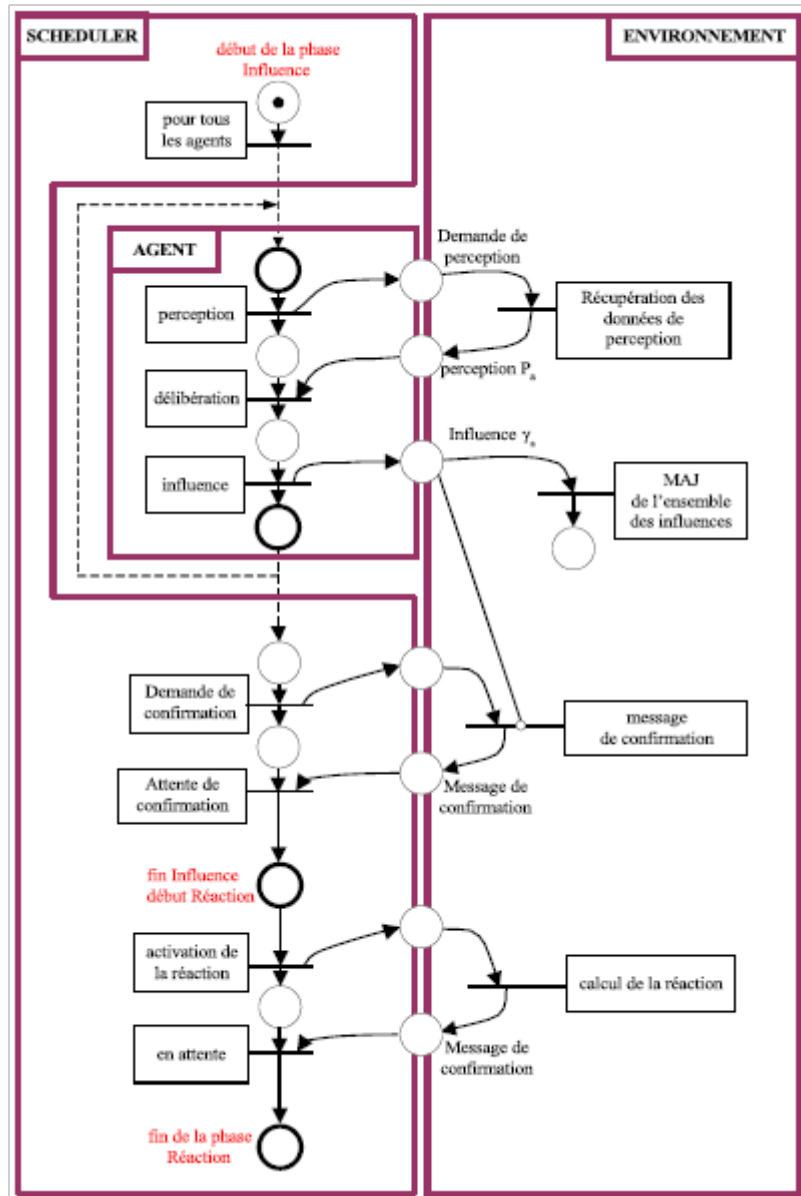


FIG. 5.10 – Implémentation d'un cycle Influence/Réaction

5.2.2 Modélisation MASQ de *Warbot*

La représentation MASQ de *Warbot* est donné sur la figure 5.11. On trouve sur cette figure autant d'esprits qu'il y a de robots qui vont participer au combat. Chacun de ces esprits est associé à un unique corps physique dans l'arène et à différents corps sociaux dans l'organisation de son équipe.

Nous allons maintenant analyser successivement la composante physique, la composante sociale et la composante culturelle de *Warbot*.

Composante physique

La modélisation de la composante physique de *Warbot* avec le méta-modèle MASQ est naturelle :

- un type d'objet `Rocket` pour les missiles ;
- un type d'objet `Obstacle` pour les obstacles ;
- un type de corps pour chaque type de robots (`Base`, `Explorer` et `RocketLauncher`) ;
- l'arène est modélisée par un unique espace brut qui traite toutes les interférences possibles entre les robots et les objets.

Composante sociale

En ce qui concerne la composante sociale de *Warbot*, elle s'exprime en termes de groupes et de rôles (au sens du méta-modèle AGR). Il nous faut donc modéliser ces deux notions en exploitant les concepts de MASQ. Ce qui nous permet d'illustrer le fait que le méta-modèle MASQ généralise le méta-modèle AGR.

Dans le méta-modèle AGR, ainsi que dans son implémentation *Madkit*, les possibilités qui sont offertes aux agents sont de :

- créer des groupes ;
- acquérir ou libérer un rôle dans un groupe ;
- communiquer par envoi de messages avec les autres agents en utilisant les rôles comme abstraction.

Suivant MASQ, la notion de groupe peut être modélisée par un type d'espace brut (noté `MadkitGroup`) et la notion de rôle par un type de corps (noté `MadkitRole`). De plus, les messages échangés entre agents peuvent être modélisés par des objets de type `MadkitMessage`.

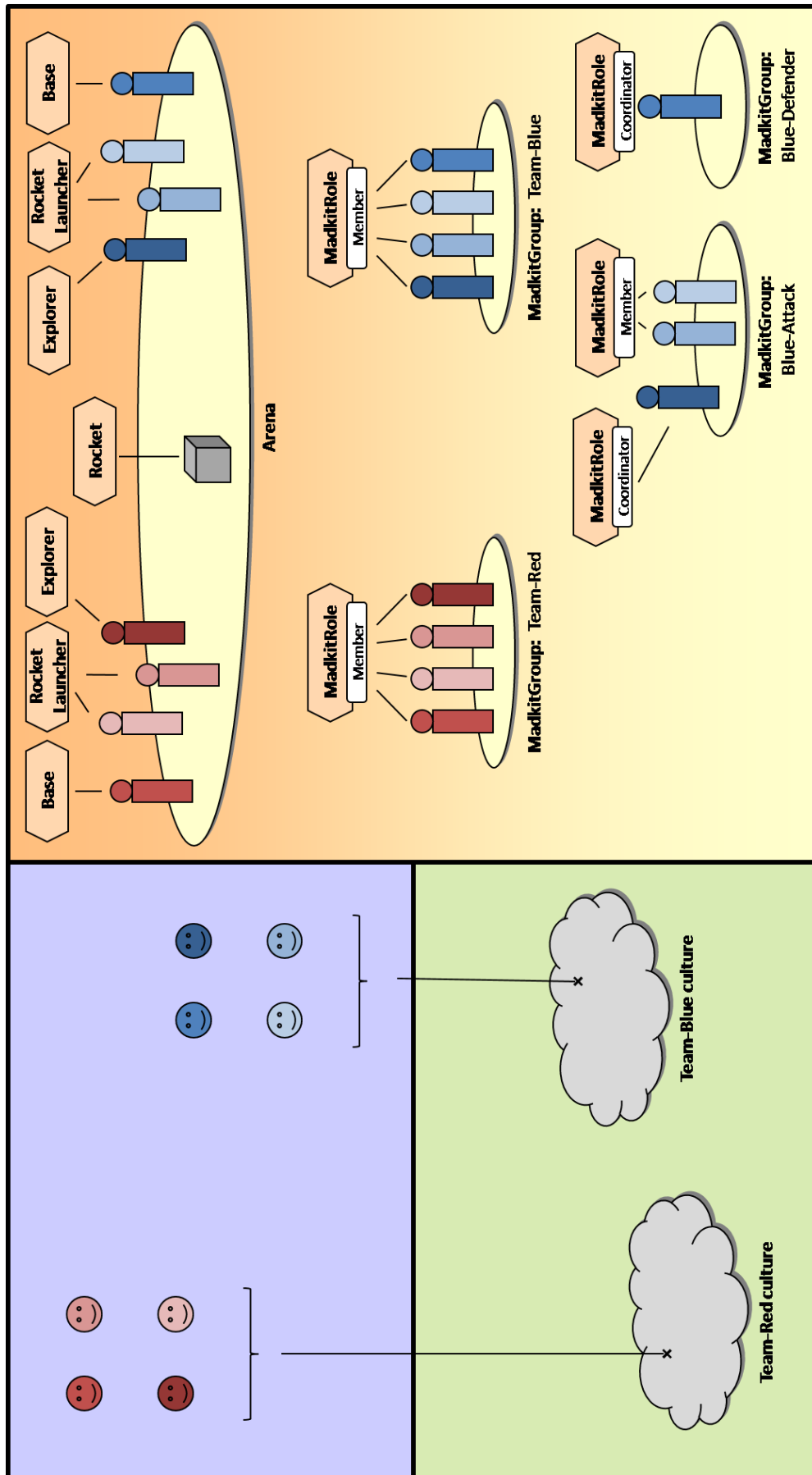


FIG. 5.11 – Représentation MASQ de warbot

Les propriétés des espaces bruts du type `MadkitGroup` sont :

- posséder un nom unique (`groupName`);
- ne contenir que des corps de type `MadkitRole` et des objets de type `MadkitMessage`;
- offrir les mécanismes (au travers de lois d'interférence) d'acheminement des messages.

Les propriétés des corps de type `MadkitRole` sont :

- posséder un nom (`roleName`);
- posséder la capacité de créer un nouvel espace brut de type `MadkitGroup`;
- posséder la capacité de créer et d'affecter à son esprit un corps de type `MadkitRole` dans tout espace brut de type `MadkitGroup` existant (nous ne traitons pas ici le cas des groupes privés qui nécessite l'envoi d'un mot de passe pour obtenir un rôle);
- posséder la capacité de s'auto-détruire (qui correspond à la libération du rôle);
- posséder les capacités nécessaires à l'envoi de message à tous les corps de type `MadkitRole` dans le même espace brut.

Suivant cette modélisation, deux rôles différents (par exemple le rôle d'enseignant et le rôle d'étudiant) sont modélisés par deux corps du type `MadkitRole`, la valeur de l'attribut `roleName` étant "enseignant" pour l'un et "étudiant" pour l'autre. Cette modélisation se justifie par le fait que tous les rôles induisent les mêmes capacités (envoi de message, acquisition/libération de rôle, création de groupe).

La sémantique d'un rôle spécifique n'est donc pas donné par le type mais par le nom du rôle (`roleName`); nous avons donc fait apparaître sur la figure 5.11 le nom de chaque rôle.

Sur la figure 5.11 nous avons représenté deux équipes (les rouges et les bleus) et nous avons illustré la structure organisationnelle mise en place par l'équipe bleue, à savoir un groupe d'attaquant qui comporte un coordinateur, et un groupe de défenseurs qui comporte aussi un coordinateur. Sur la figure, le groupe de défenseurs ne comporte pas d'autre membre que le coordinateur, mais cette situation pourra évoluer au cours du temps de la partie. Cette structuration spécifique est indépendante de la plate-forme *Warbot*, c'est le choix d'une équipe spécifique. La plate-forme *Warbot* fournit uniquement la composante physique (l'arène et ce qu'elle contient), et un espace de type `MadkitGroup` pour chaque équipe. C'est à partir du rôle initial de membre que chaque esprit a la possibilité de créer de nouveau groupe pour son équipe.

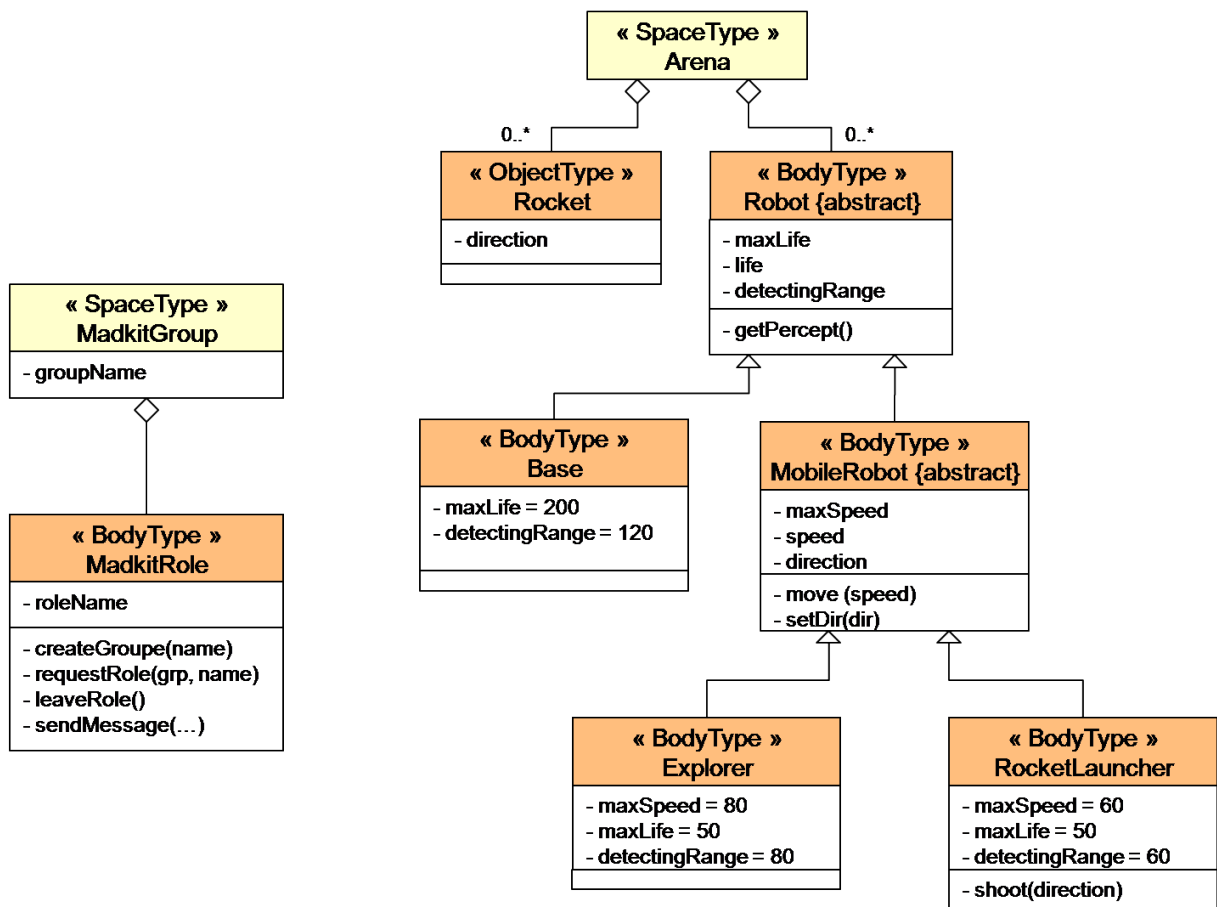


FIG. 5.12 – Diagramme UML des types d'espace brut, de corps et d'objet pour Warbot

Définition des types d'espaces bruts, d'objets et de corps

La figure 5.12 présente le diagramme UML de la définition de types MASQ pour la modélisation de *Warbot*. On peut remarquer sur ce diagramme deux types abstraits (**Robot** et **MobileRobot**) qui permettent de factoriser les éléments communs à plusieurs types concrets.

Composante culturelle

La version actuelle de *Warbot* ne comporte pas de dimension culturelle. Nous allons donc considérer ici quels sont les bénéfices à ajouter cette dimension à cette application spécifique.

Les esprits reçoivent continuellement de la plate-forme des perceptions qui leur permettent de se construire une représentation partielle du monde dans lequel leurs corps se trouvent. Ces perceptions sont de nature objective : elles contiennent par exemple la nature et la position relative des différentes entités qui se trouvent dans les champs de perception.

Pour produire leur comportement, les agents doivent interpréter ces informations objectives. Par exemple, une situation donnée pourra être considérée comme un danger pour la base. Les messages que s'échangent les membres d'une équipe doivent aussi être interprétés. Par exemple, un message donné peut signifier une demande de renfort.

Pour renforcer la coordination entre les membres d'une équipe, il est important que les règles d'interprétations soient communes à tous les membres de l'équipe.

Actuellement, ces interprétations sont réalisées par les différents agents en fonction des règles encodées dans leur structure interne par leur(s) concepteur(s). Les règles d'interprétations des esprits d'une même équipe sont communes car elles sont définies par un concepteur unique ou un groupe de concepteurs qui collaborent.

Cependant, ne pas séparer la conception de la partie décisionnelle d'un agent (esprit) de la partie interprétation commune d'une équipe (culture) relève d'une confusion de rôle entre l'interprétation individuelle et l'interprétation collective.

Considérons maintenant les avantages qu'il y a à séparer explicitement la dimension culturelle :

- *modularité* : la construction d'une interprétation de la réalité brute peut être réalisée indépendamment des mécanismes de raisonnement qui s'appuient sur cette interprétation pour produire un comportement ; il s'agit donc de découpler des parties de codes qui peuvent être conçues séparément ;
- *réutilisabilité* : une culture, si elle a été conçue séparément, peut être exploitée par différentes équipes ; il est de plus possible, au cours du temps, d'enrichir une culture donnée, ce qui permettrait, à terme, pour *Warbot*, de disposer de différentes cultures de bases. Ces cultures pourront alors être exploitées et enrichies par de nouvelles équipes ; cette réutilisation de la culture éviterait de redévelopper certains aspects nécessaires à toutes équipes tel que la définition d'un langage de communication minimal pour coordonner une attaque, un repli, etc.

5.2.3 Aspects implémentatoires

A ce jour, l'implémentation du méta-modèle MASQ est en cours de réalisation. Ce travail s'inscrit dans la définition d'une nouvelle version de la plate-forme *Madkit*. Alors que les différentes versions opérationnelles de *Madkit* se fondent sur le méta-modèle AGR, cette nouvelle version sera fondée sur le méta-modèle MASQ.

La figure 5.13 présente un diagramme UML de ce que pourrait être une implémentation de *Warbot* en utilisant le méta-modèle MASQ. Sur ce diagramme, les classes coloriées

en bleu sont celles du méta-modèle MASQ. Les classes coloriées en jaune sont celles de la plate-forme *Warbot*. Ces classes étendent les classes de MASQ; elles correspondent aux différents types qui ont été établis durant la modélisation. Les classes coloriées en rouge sont celles qui sont définies par le(s) concepteur(s) d'une équipe de *Warbot*.

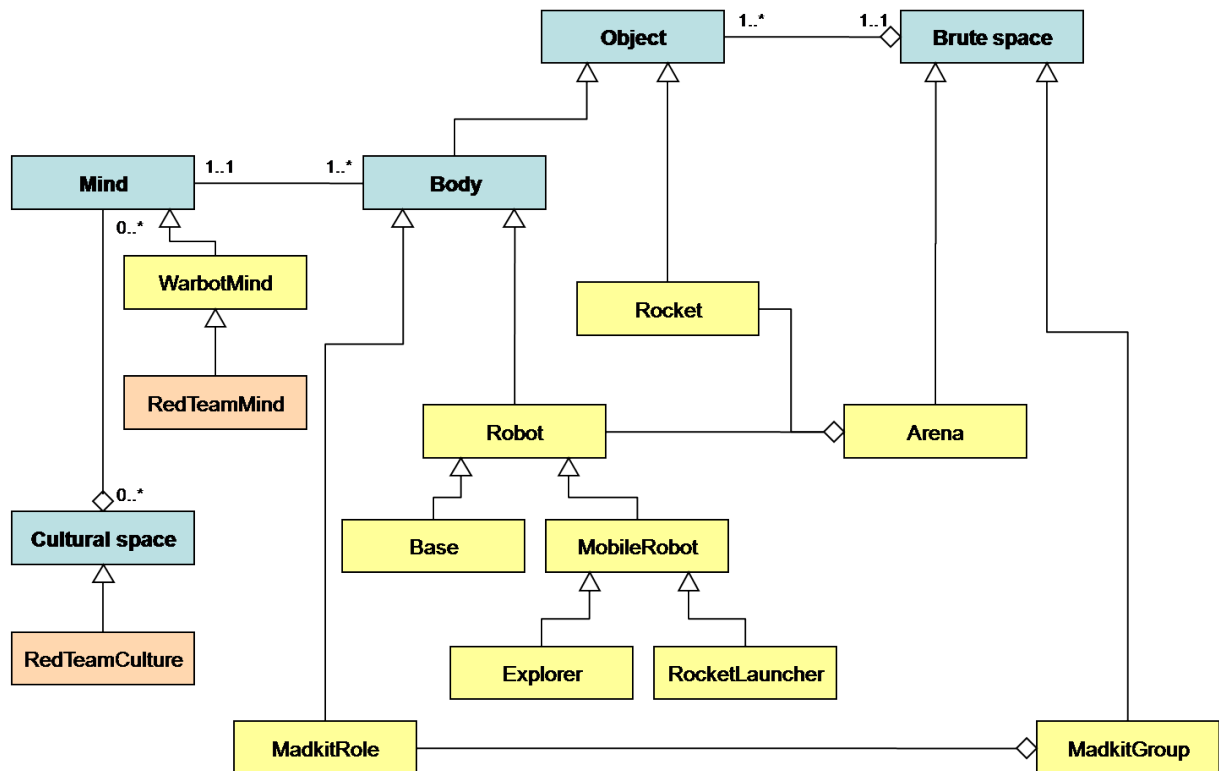


FIG. 5.13 – Warbot/MASQ : Diagramme UML

```

public class RedTeamMind extends WarbotMind {
...
    public void live() {
        Percept [] p = myRobotBody.getPercept();
        Message [] m = myMemberTeamBody.getMessage();
        ...
    }

    private void handleMessage(Message msg) {
        RedTeamCulture.interpretMessage(myCulturalReality, msg);
    }

    private void move(float speed, float direction) {
        myRobotBody.setSpeed(speed);
        myRobotBody.setDirection(direction);
    }

    private void informAll(Message msg) {
        myMemberTeamBody.broadcastMessage(msg);
    }
    ...
}

```

FIG. 5.14 – Extrait de code d'un esprit pour Warbot

La figure 5.14 présente un extrait de ce que serait le code d'un esprit *Warbot* implémenté suivant MASQ.

L'envoi d'une influence vers un corps s'effectue par un appel de méthode (par exemple `myRobotBody.setSpeed(speed)`). Cet appel de méthode n'engendre que la transmission de l'influence. Son traitement sera effectué par la suite, au moment de la réaction de l'espace brut.

La représentation du contexte culturel est stockée dans la variable `myCulturalReality`. L'interprétation d'un message est alors effectuée par un appel à la méthode statique `interpretMessage` de la classe `RedTeamCulture`. Cette méthode modifie par effet de bord la variable `myCulturalReality`. Par ailleurs, cette variable est exploitée durant la délibération de l'esprit (méthode `live`).

5.2.4 Conclusion

Dans cette section nous avons illustré comment la plate-forme *Warbot* peut être modélisée avec le méta-modèle MASQ. Au travers de cet exemple, nous avons montré comment l'activité physique et l'activité sociale pouvait être traitée de manière uniforme en utilisant les concepts d'espace brut, d'objet et de corps. Nous avons montré en quoi le méta-modèle MASQ généralise le méta-modèle AGR. Par ailleurs, nous avons motivé pour ce contexte spécifique l'introduction explicite de la dimension culturelle.

Un aspect intéressant de ce scénario est l'ouverture du système, au sens où les esprits sont conçus indépendamment de la plate-forme. Nous avons présenté comment la conception de l'environnement (physique et social) engendre les capacités matérielles conférées aux esprits. En particulier, nous avons vu que les esprits peuvent intervenir dans la structuration de l'environnement en créant dynamiquement des groupes et des rôles.

Nous concluons cette section sur la perspective suivante. Il serait intéressant d'expérimenter des combats de robots avec *Warbot* en formant des équipes à partir d'esprits conçus séparément. Actuellement la stratégie d'une équipe de robots est le résultat de la réflexion du ou des concepts de l'équipe. Si les équipes devaient être constituées au moment de l'exécution, la stratégie devrait être mise au point dynamiquement par les agents eux-mêmes. Un tel scénario serait difficile à mettre en œuvre avec l'implémentation actuelle de *Warbot* qui ne permet de définir des cultures. Il permettrait d'une part de mieux illustrer l'importance de la culture, et d'autre part de disposer d'un outil pour expérimenter des techniques destinées à favoriser la coordination en milieu ouvert.

Chapitre 6

Conclusion et perspectives

Dans cette thèse nous nous sommes intéressé au problème de l'intégration des concepts d'agent, d'environnement, d'organisation et d'institution. Nous avons constaté qu'il existait différentes écoles en ce qui concerne les SMA, chacune d'elles privilégiant soit le concept d'agent, soit le concept d'environnement, soit les concepts sociaux. Afin de pouvoir intégrer tous ces concepts de manière cohérente, nous avons mis en évidence le besoin d'établir un cadre conceptuel qui puisse réconcilier ces différentes approches.

6.1 Bilan

Dans cet objectif, nous avons proposé un nouveau cadre conceptuel inspiré de la vision intégrale proposée par Wilber. Ce cadre conceptuel repose sur quatre dimensions fondamentales (les quatre quadrants) : intentionnelle (quadrant I-I), comportementale (quadrant E-I), sociale (quadrant E-C), et culturelle (quadrant I-C). Ces quatre dimensions résultant de la distinction entre la perspective interne et la perspective externe d'une part, et entre la perspective individuelle et la perspective collective d'autre part.

Comparaison avec l'approche Voyelles

Dans l'introduction de cette thèse, nous avons considéré les dimensions qui sont proposées par l'approche VOYELLES pour aborder un SMA : l'agent, l'environnement, l'interaction et l'organisation (ou la dimension sociale). Étant donné que nous proposons quatre dimensions différentes (les quadrants), il est important de comparer les deux approches (table 6.1 et figure 6.1).

Sur la table 6.1, nous présentons pour chaque dimension de VOYELLES les quadrants qui leur correspondent. Ce que cette table illustre, c'est qu'il est possible d'exprimer chaque dimension du modèle VOYELLES en termes des quadrants. Par contre, la réciproque n'est pas vraie. L'approche VOYELLES a le mérite de couvrir tous les aspects

fondamentaux d'un SMA, mais elle ne permet pas de distinguer entre les quadrants E-C et I-C (social et culturel), et les quadrants E-I et E-C (structures individuelles et structures collectives). De plus, dans cette approche, l'interaction est considérée comme une dimension à part entière ; nous considérons au contraire l'interaction comme un élément qui intervient dans tous les quadrants mais qui n'ajoute rien de plus à ceux-ci.

La figure 6.1 illustre le placement des dimensions de l'approche VOYELLES sur les quadrants. Cette figure montre le chevauchement de ces dimensions. Ainsi, si les quadrants engendrent quatre dimensions orthogonales, les dimensions de VOYELLES elles ne le sont pas. C'est pourquoi nous proposons d'utiliser comme cadre de référence les quatre quadrants, et non les concepts d'agent, d'environnement, d'interaction et d'organisation.

Voyelles	Quadrants
Agent	I-I ¹
Environnement	E-I + E-C
Interaction	I-I + E-I + E-C + I-C
Organisation	E-C + I-C ²

TAB. 6.1 – Comparaison VOYELLES – Quadrants

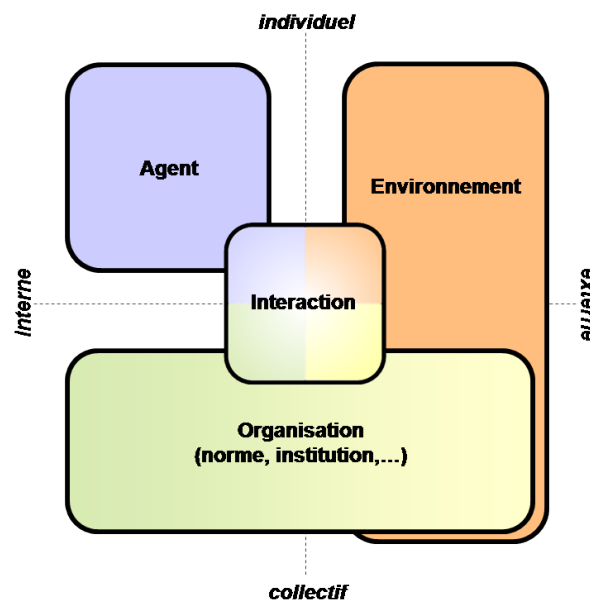


FIG. 6.1 – Les dimensions de l'approche VOYELLES situées sur les quadrants

¹En considérant que la dimension de l'agent dans VOYELLES ne comprend que la partie décisionnelle de l'agent.

²Nous incluons ici le quadrant I-C car dans VOYELLES la dimension de l'organisation comprend aussi les aspects normatifs.

Différenciation des concepts fondamentaux

Le cadre conceptuel que nous proposons introduit une nouvelle terminologie, celle de Wilber, ce qui permet de réduire les confusions possibles. Alors que certains concepts ont des acceptions différentes (comme par exemple l'organisation ou l'institution), les quadrants ont une sémantique bien précise qui permet dans certains cas de différencier des concepts qui sont employés dans des sens différents.

Nous avons vu qu'il était possible d'exprimer les concepts d'agent, d'environnement, d'interaction, d'organisation, de norme, d'engagement social ou encore d'institution en termes des quatre quadrants ce qui nous a permis de déterminer plus précisément le rôle de chacun de ces concepts.

De plus, cette approche permet de lever certaines confusions de rôle entre les concepts :

- confusion entre la partie décisionnelle et la partie active d'un agent ;
- confusion entre les responsabilités de l'environnement et celles des agents ;
- confusion entre la structuration objective d'une organisation et les interprétations subjectives induites par les normes, les engagements sociaux et les institutions ;
- confusion dans les composantes de l'interaction (par exemple la sémantique publique d'un langage qui est définie en termes des états mentaux).

Intégration des concepts fondamentaux

Après avoir différencié les concepts fondamentaux, nous avons montré comment il était possible de les intégrer dans un cadre cohérent. Pour cela nous avons exploité les principes suivants :

- le principe de la séparation esprit/corps ;
- le principe influence/réaction ;
- une représentation explicite de la dynamique de l'environnement (lois de réaction, d'interférence et d'évolution) ;
- les règles constitutives définies par Searle.

Unification de l'activité physique/sociale

Dans cette approche, nous avons distingué la perspective subjective et la perspective objective plutôt que la perspective sociale et la perspective physique. L'intérêt de cette unification entre environnement physique et environnement social est de réconcilier les approches qui exploitent des agents situés et celles qui exploitent des agents communicants.

Un cadre d'analyse

Nous avons vu que la vision intégrale peut guider l'analyse de problème que l'on souhaite modéliser avec une approche multi-agent. Le fait que les quadrants puissent être exploités à la fois pour l'analyse et pour la conception est un avantage indéniable de cette approche.

Une approche de conception pour les systèmes ouverts

Un aspect particulièrement intéressant de cette approche est de permettre une conception séparée des quadrants pour un SMA. L'environnement¹ (quadrant E-I et E-C), les agents (quadrant I-I) et les cultures (quadrant I-C) peuvent être conçus indépendamment les uns des autres.

De plus, la dimension culturelle permet de proposer des solutions pour traiter de la coordination entre agents hétérogènes.

C'est pour ces deux raisons que nous pouvons dire que cette approche est adaptée à la conception de systèmes ouverts.

MASQ : un méta-modèle formel basé sur les quadrants

Nous avons proposé le méta-modèle MASQ pour permettre de décrire un SMA de manière formelle suivant chaque quadrant. La formalisation que nous avons donnée permet d'avoir une spécification précise des concepts qui sont mis en œuvre. De plus nous avons essayé de donner une formalisation la plus générique possible pour être en mesure d'intégrer des modèles spécifiques. L'objectif n'étant pas de déterminer quel modèle d'agent, d'environnement, d'organisation ou d'institution il faut utiliser, mais de fournir des moyens d'intégrer des modèles existants de ces différents concepts.

6.2 Perspectives

Cette thèse propose un cadre intégrateur très général, les perspectives de mises en œuvre concrètes de cette approche sont donc nombreuses. Hormis les mises en œuvre possibles, deux perspectives qui découlent de ce travail nous intéressent tout particulièrement : les aspects implémentatoires et l'évolution de la dimension culturelle.

¹Avec le méta-modèle MASQ, il est possible de concevoir séparément les objets et les espaces bruts. L'intérêt ici est d'avoir une conception modulaire de l'environnement, qui permet aussi la réutilisation des objets définis dans différents espaces bruts.

Aspects implémentatoires

Dans cette thèse, nous avons apporté une réflexion sur plan conceptuel ainsi qu'une formalisation des concepts utilisés. Cependant nous n'avons pas abordé de considérations implémentatoires. Bien évidemment, il serait intéressant d'analyser les répercussions de cette approche sur la manière d'implémenter un SMA et de proposer des outils pour faciliter cette tâche.

Évolution de la dimension culturelle

Dans cette thèse nous avons largement insisté sur l'importance de la culture pour la conception de systèmes ouverts. Cependant nous avons aussi discuté des limites de notre approche en ce qui concerne cet aspect. Une réflexion supplémentaire est nécessaire pour comprendre tous les enjeux de cette dimension. Notamment, notre compréhension du changement culturel est encore insuffisante pour permettre de mettre en œuvre des solutions opérationnelles. Donner la possibilité aux agents de faire évoluer leur culture, c'est leur donner la possibilité de modifier eux-mêmes la régulation de leurs interactions. Ils pourraient par exemple faire évoluer, en fonction de leurs besoins, la sémantique des langages qu'ils utilisent.

La tendance actuelle des systèmes informatiques est d'aller vers toujours plus d'ouverture. La dimension culturelle est donc un aspect qu'il sera essentiel de développer pour les agents informatiques.

Table des figures

1.1	Interaction entre un agent et son environnement	10
1.2	L'approche VOYELLES	12
2.1	L'agent, un processus à 3 phases : <i>perception, délibération</i> puis <i>action</i>	18
2.2	Modèle 3-tier pour les SMA	22
2.3	Structuration physique de l'environnement	26
2.4	Le protocole d'interaction FIPA Contract Net	36
2.5	UML : Le méta-modèle AGR	43
2.6	UML : Le méta-modèle AGR	44
2.7	Cycle de vie d'un engagement social	52
3.1	Les 4 quadrants	71
3.2	Les quatre quadrants : quatre points de vue	73
3.3	Les concepts fondamentaux d'un SMA situés sur les quadrants	75
3.4	Un rôle est l'association entre un corps social et un statut	78
3.5	Relations entre les quadrants	87
3.6	Quadrants et comportements	89
3.7	Quadrants et capacités	89
4.1	Les concepts primitifs : <i>esprit, objet, espace brut</i> et <i>culture</i>	92
4.2	La relation entre un esprit et son environnement	96
4.3	Lien esprit – corps	101
4.4	Structuration de l'environnement par des espaces bruts	106
4.5	Collision de deux billes	110
4.6	Le cycle à 3 phases d'un espace brut	112
4.7	Relation d'identité entre objets de différents espaces bruts	117
4.8	Communication par passage de messages entre différents espaces	119
4.9	Modélisation holonique	121
4.10	Architecture client–serveur	122
4.11	Architecture pair-à-pair (P2P)	122

4.12	Architecture <i>super-peer</i>	122
4.13	MASQ : les concepts	136
4.14	MASQ : les lois	136
4.15	MASQ : les relations	137
4.16	Diagramme UML du méta-modèle MASQ	138
5.1	Diagramme UML du wiki	140
5.2	Les cas d'utilisations du wiki	142
5.3	Représentation MASQ du wiki	144
5.4	Diagramme UML des types d'espace brut, de corps et d'objet pour le wiki	146
5.5	Extrait de la définition de type en XML	146
5.6	Séquence d'authentification	147
5.7	Séquence de consultation d'un article	148
5.8	Séquence d'édition d'un article	149
5.9	Capture d'écran d'une partie de <i>Warbot</i>	152
5.10	Implémentation d'un cycle Influence/Réaction	153
5.11	Représentation MASQ de warbot	155
5.12	Diagramme UML des types d'espace brut, de corps et d'objet pour Warbot	157
5.13	Warbot/MASQ : Diagramme UML	159
5.14	Extrait de code d'un esprit pour Warbot	159
6.1	Les dimensions de l'approche VOYELLES situées sur les quadrants	162

Liste des tableaux

2.1	Les paramètres des messages FIPA-ACL	34
2.2	Les notions juridiques opposées	47
2.3	Les notions juridiques duales	47
6.1	Comparaison VOYELLES – Quadrants	162

Bibliographie

- [Artikis *et al.*, 2002] Alexander Artikis, Jeremy Pitt, et Marek Sergot. Animated specifications of computational societies. In *Proceedings of the first international joint conference on Autonomous agents and multiagent systems (AAMAS '02)*, pages 1053–1061, New York, USA, 2002. ACM Press.
- [Austin, 1962] J. Austin. *How to Do Things with Words ?* Clarendon Press, London, 1962.
- [Bailhache, 1991] Patrice Bailhache. *Essai de logique déontique*. Cambridge University Press, 1991.
- [Balbo *et al.*, 2002] F. Balbo, A. El Fallah Seghrouchni, et S. Pinson. *Organisation et applications des SMA*, chapitre La coordination d’actions par planification Multi-Agents. Hermès, 2002.
- [Báez-Barranco *et al.*, 2006] José-Antonio Báez-Barranco, Tiberiu Stratulat, et Jacques Ferber. A unified model for physical and social environments. In *the Proceedings of Environments for Multiagent Systems (E4MAS)*, éditeurs Danny Weyns, H. Van Dyke Parunak, et Fabien Michel, volume 4389 de *Lecture Notes in Computer Science*, pages 41–50. Springer, 2006.
- [Boella et Torre, 2004] Guido Boella et Leendert van der Torre. Structuring organizations by means of roles using the agent metaphor. In *Proceedings of the Workshop dagli Oggetti agli Agenti (WOA'04)*, Pitagora, 2004.
- [Carabelea et Boissier, 2006] Cosmin Carabelea et Olivier Boissier. Coordinating agents in organizations using social commitments. *Electronic Notes in Theoretical Computer Science*, 150(3) :73–91, 2006.
- [Castelfranchi, 1995] Cristiano Castelfranchi. Commitments : From individual intentions to groups and organizations. In *the First International Conference on Multiagent Systems (ICMAS)*, pages 41–48, San Francisco, California, USA, June 1995. The MIT Press.
- [Chaib-draa *et al.*, 2006] Brahim Chaib-draa, Marc-André Labrie, Mathieu Bergeron, et Philippe Pasquier. Diagal : An agent communication language based on dialogue games and sustained by social commitments. *Autonomous Agents and Multi-Agent Systems*, 13(1) :61–95, 2006.

- [Chang *et al.*, 2004] Paul Hsueh-Min Chang, Kuang-Tai Chen, Yu-Hung Chien, Edward Chao-Chun Kao, et Von-Wun Soo. From reality to mind : A cognitive middle layer of environment concepts for believable agents. In *the Proceedings of Environments for Multiagent Systems (E4MAS)*, pages 57–73, 2004.
- [Chellas, 1980] B. F. Chellas. *Modal Logic – an Introduction*. Cambridge University Press, 1980.
- [Cohen et Levesque, 1990] Philip R. Cohen et Hector J. Levesque. Intention is choice with commitment. *Artificial Intelligence*, 42(2-3) :213–261, 1990.
- [Colombetti *et al.*, 2003] Marco Colombetti, Nicoletta Fornara, et Mario Verdicchio. A social approach to communication in multiagent systems. In *the Proceedings of Declarative Agent Languages and Technologies (DALT)*, éditeurs João Alexandre Leite, Andrea Omicini, Leon Sterling, et Paolo Torroni, volume 2990 de *Lecture Notes in Computer Science*, pages 191–220. Springer, 2003.
- [Corbara *et al.*, 1993] B. Corbara, A. Drogoul, D. Fresneau, et S. Lalande. Simulating the sociogenesis process in ant colonies with manta. In *Towards a Practice of Autonomous Systems II*, Cambridge, 1993. MIT Press.
- [Davis et Smith, 1983] Randall Davis et Reid G. Smith. Negotiation as a metaphor for distributed problem solving. *Artificial Intelligence*, 20 :63–109, 1983.
- [Demazeau et Costa, 1996] Y. Demazeau et A.R. Costa. Populations and organizations in open multi-agent systems. In *the 1st National Symposium on Parallel and Distributed AI (PDAI96)*, 1996.
- [Demazeau, 2001] Yves Demazeau. *VOYELLES*. Mémoire d’Habilitation à Diriger des Recherches, Institut National Polytechnique de Grenoble INPG, Avril 2001.
- [Dignum *et al.*, 2000] F. Dignum, D. Morley, et E.A. Sonenberg. Towards socially sophisticated bdi agents. In *the Fourth International Conference on MultiAgent Systems (ICMAS’00)*, page 111, Washington, DC, USA, 2000. IEEE Computer Society.
- [Dignum *et al.*, 2002a] V. Dignum, J.-Ch. Meyer, H. Wiegand, et F. Dignum. An organisational-oriented model for agent societies. In *the International Workshop on Regulated Agent-Based Social Systems : Theories and Applications (RASTA ’02)*, éditeurs G. Lindemann, D. Moldt, M. Paolucci, et B. Yu, volume 318, pages 31–50, Hamburg, July 2002.
- [Dignum *et al.*, 2002b] V. Dignum, H. Wiegand, et L. Xu. Agent societies : Towards framework-based design. In *the 2nd Workshop on Agent-Oriented Software Engineering, Autonomous Agent*, Montreal, Canada, 2002. Lecture Notes in Computer Science 2222, Springer-Verlag.

-
- [Dignum *et al.*, 2002c] Virginia Dignum, John-Jules Ch. Meyer, Frank Dignum, et Hans Weigand. Formal specification of interaction in agent societies. In *the Proceedings of Workshop on Formal Approaches to Agent-Based Systems (FAABS)*, pages 37–52, 2002.
- [Eijk, 2002] Rogier M. van Eijk. Semantics of agent communication : An introduction. In *the UKMAS Workshop on Foundations and Applications of Multi-Agent Systems*, pages 152–168, London, UK, 2002. Springer-Verlag.
- [El Fallah Seghrouchni, 2001] A. El Fallah Seghrouchni. *Principes et architecture des systèmes multi-agents*, chapitre Les modèles de coordination d’agents cognitifs. Hermès, 2001.
- [Esteva *et al.*, 2001] Marc Esteva, Juan A. Rodríguez-Aguilar, Carles Sierra, Pere Garcia, et Josep Lluís Arcos. On the formal specifications of electronic institutions. In *Agent Mediated Electronic Commerce, The European AgentLink Perspective.*, pages 126–147, London, UK, 2001. Springer-Verlag.
- [Esteva *et al.*, 2002] Marc Esteva, Julian A. Padget, et Carles Sierra. Formalizing a language for institutions and norms. In *the 8th International Workshop on Intelligent Agents VIII in ATAL '01*, pages 348–366, London, UK, 2002. Springer-Verlag.
- [Ferber *et al.*, 2003] Jacques Ferber, Olivier Gutknecht, et Fabien Michel. From agents to organizations : an organizational view of multi-agent systems. In *Agent Oriented Software Engineering held at AAMAS*, Melbourne, 2003.
- [Ferber *et al.*, 2004] Jacques Ferber, Fabien Michel, et José-Antonio Báez-Barranco. Agre : Integrating environments with organizations. In *the Proceedings of Environments for Multiagent Systems (E4MAS)*, éditeurs Danny Weyns, H. Van Dyke Parunak, et Fabien Michel, volume 3374 de *Lecture Notes in Computer Science*, pages 48–56. Springer, 2004.
- [Ferber et Gutknecht, 1998] J. Ferber et O. Gutknecht. A meta-model for the analysis and design of organizations in multi-agent systems. In *the 3rd International Conference on Multi Agent Systems (ICMAS '98)*, Washington, DC, USA, 1998. IEEE Computer Society.
- [Ferber et Müller, 1996] Jacques Ferber et Jean-Pierre Müller. Influences and reaction : a model of situated multi-agent systems. In *the 2nd International Conference on Multi-agent Systems (ICMAS-96)*, éditeur Mario Tokoro, pages 72–79. The AAAI Press, 1996.
- [Ferber, 1995] J. Ferber. *Les systèmes multi-agents : vers une intelligence collective*. InterEditions, Paris, 1995.
- [Ferber, 1999] Jacques Ferber. *Multi-Agent Systems : An Introduction to Distributed Artificial Intelligence*. Addison-Wesley Longman Publishing Co., Inc., Boston, MA, USA, 1999.

- [Ferguson, 1992] I. A. Ferguson. *TouringMachines : An Architecture for Dynamic, Rational, Mobile Agents*. PhD thesis, University of Cambridge, UK, 1992.
- [Finin *et al.*, 1994] Tim Finin, Richard Fritzson, Donald P McKay, et Robin McEntire. KQML as an agent communication language. In *the third international conference on Information and knowledge management*, pages 456–463. ACM Press, October 1994.
- [FIPA, 2001] FIPA. *FIPA ACL Message Structure Specification*. FIPA, 2001.
- [Flores et Kremer, 2002] Roberto A. Flores et Rob Kremer. Bringing coherence to agent conversations. In *the Second International Workshop on Agent-Oriented Software Engineering II (AOSE '01)*, pages 50–67, London, UK, 2002. Springer-Verlag.
- [Fornara *et al.*, 2007] Nicoletta Fornara, Francesco Viganò, et Marco Colombetti. Agent communication and artificial institutions. *Autonomous Agents and Multi-Agent Systems*, 14(2) :121–142, 2007.
- [Fornara et Colombetti, 2003] Nicoletta Fornara et Marco Colombetti. Defining interaction protocols using a commitment-based agent communication language. In *Proceedings of the international joint conference on Autonomous agents and multiagent systems (AAMAS)*, pages 520–527. ACM, 2003.
- [Fox, 1988] Mark S. Fox. An organizational view of distributed systems. *Distributed Artificial Intelligence*, pages 140–150, 1988.
- [Gasser *et al.*, 1987] Les Gasser, Carl Braganza, et Nava Herman. Mace : A flexible test-bed for distributed ai research. *Distributed AI*, 1987.
- [Gâteau *et al.*, 2005] Benjamin Gâteau, Olivier Boissier, Djamel Khadraoui, et Eric Dubois. Moiseinst : An organizational model for specifying rights and duties of autonomous agents. In *the Third European Workshop on Multi-Agent Systems (EUMAS'05)*, éditeurs Marie Pierre Gleizes, Gal A. Kaminka, Ann Nowé, Sascha Ossowski, Karl Tuyls, et Katja Verbeeck, pages 484–485. Koninklijke Vlaamse Academie van Belie voor Wetenschappen en Kunsten, 2005.
- [Grassé, 1959] P. Grassé. La reconstruction du nid et les coordinations inter-individuelles chez *bellicositermes natalensis* et *cubitermes* sp. la théorie de la stigmergie : Essai d'interprétation des termites constructeurs. *Insectes Sociaux*, 6 :41–84, 1959.
- [Gutknecht et Ferber, 2000] Olivier Gutknecht et Jacques Ferber. Madkit : a generic multi-agent platform. In *the fourth international conference on Autonomous agents (AGENTS '00)*, pages 78–79, New York, NY, USA, 2000. ACM Press.
- [Gutknecht, 2001] Olivier Gutknecht. *Proposition d'un modèle organisationnel générique de systèmes multi-agents*. Thèse de doctorat, Université Montpellier II, Septembre 2001.

-
- [Hannoun *et al.*, 2000] Mahdi Hannoun, Olivier Boissier, Jaime Simão Sichman, et Claudette Sayettat. *Moise : An organizational model for multi-agent systems*. In *the International Joint Conference, 7th Ibero-American Conference on AI (IBERAMIA-SBIA'00)*, pages 156–165, London, UK, 2000. Springer-Verlag.
- [Helleboogh *et al.*, 2007] Alexander Helleboogh, Giuseppe Vizzari, Adelinde Uhrmacher, et Fabien Michel. Modeling dynamic environments in multi-agent simulation. *Autonomous Agents and Multi-Agent Systems*, 14(1) :87–116, 2007.
- [Hohfeld, 1917] Wesley Newcomb Hohfeld. Fundamental legal conceptions as applied in judicial reasoning. *The Yale Law Journal*, 26(8) :710–770, 1917.
- [Hübner *et al.*, 2002] Jomi Fred Hübner, Jaime Simão Sichman, et Olivier Boissier. *Moise+ : towards a structural, functional, and deontic model for mas organization*. In *the first international joint conference on Autonomous agents and multiagent systems (AAMAS '02)*, pages 501–502, New York, NY, USA, 2002. ACM Press.
- [Jennings, 2000] Nicholas R. Jennings. On agent-based software engineering. *Artif. Intell.*, 117(2) :277–296, 2000.
- [Jones et Sergot, 1996] A. Jones et M. Sergot. A formal characterization of institutionalized power. *Journal of the IGPL*, 4(3) :429–445, 1996.
- [Jonquet, 2006] Clément Jonquet. *Dynamic Service Generation : Agent interactions for service exchange on the Grid*. Thèse de doctorat, Université Montpellier II, Novembre 2006.
- [Kaminka *et al.*, 2002] Gal A. Kaminka, David V. Pynadath, et Milind Tambe. Monitoring teams by overhearing : A multi-agent plan-recognition approach. *Journal of Artificial Intelligence Research (JAIR)*, 17 :83–135, 2002.
- [Kanger, 1957] S. Kanger. New foundations for ethical theory. *Deontic Logic : Introductory and Systematic Readings*, pages 36–58, 1957.
- [Koning et Pesty, 2001] J.L. Koning et S. Pesty. *Modèles de communication*, chapitre Systèmes Multi-agents. Hermes, 2001.
- [Kripke, 1963] S. Kripke. Semantical considerations on modal logic. *Acta Philosophica Fennica*, 16 :83–94, 1963.
- [Lemaitre et Excelente, 1998] C. Lemaitre et C. B. Excelente. Multi-agent organization approach. In *the II Iberoamerican Workshop on DAI and MAS*, éditeurs F. J. Garijo et C. Lemaitre, 1998.
- [Magnin, 1996] Laurent Magnin. *Modélisation et simulation de l'environnement dans les systèmes multi-agents : application aux robots footballeurs*. Thèse de doctorat, Université Paris VI, Novembre 1996.

- [Martindale, 1978] D. Martindale. *Social Control for the 1980s : A Handbook for Order in a Democratic Society*, chapitre The theory of social control, pages 46–58. Greenwood Press, 1978.
- [Mertens *et al.*, 2004] Koenraad Mertens, Tom Holvoet, et Yolande Berbers. Adaptation in a distributed environment. In *Environments for Multiagent Systems*, pages 49–59, 2004.
- [Michel, 2004] Fabien Michel. *Formalisme, outils et éléments méthodologiques pour la modélisation et la simulation multi-agents*. Thèse de doctorat, Université Montpellier II, Décembre 2004.
- [Müller, 1996] J. Müller. *An Architecture for Dynamically Interacting Agents*. PhD thesis, DFKI, University of Saarlandes, Saarbrücken, 1996.
- [Morin, 1977] Edgar Morin. *La méthode 1 : La nature de la nature*. Éditions du Seuil, 1977.
- [North, 1990] Douglass C. North. *Institutions, Institutional Change, and Economic Performance*. Cambridge University Press, Cambridge, 1990.
- [Odell *et al.*, 2002] J. Odell, H.V.D. Parunak, M. Fleisher, et S. Breuckner. Modeling agents and their environment. In *Agent-Oriented Software Engineering III*, éditeurs F. Giunchiglia, J. Odell, et G. Weiss, volume 2585 de *Lecture Notes in Computer Science*. Springer-Verlag, 2002.
- [Pasquier *et al.*, 2004] P. Pasquier, R. A. Flores, et B. Chaib-draa. Modelling flexible social commitments and their enforcement. In *the Fifth International Workshop Engineering Societies in the Agents World (ESAW'04)*, volume 3451, pages 153–165. Springer-Verlag, 2004.
- [Pnueli, 1986] Amir Pnueli. Specification and development of reactive systems (invited paper). In *IFIP Congress*, pages 845–858, 1986.
- [Posner et Rasmusen, 1999] Richard A. Posner et Eric B. Rasmusen. Creating and enforcing norms, with special reference to sanctions. *International Review of Law and Economics*, 19(3) :369–382, September 1999.
- [Rao et Georgeff, 1992] A. Rao et M. Georgeff. An abstract architecture for rational agents. In *the International Workshop on Knowledge Representation, KR'92*, éditeurs C. Rich, W. Swartout, et B. Nebel, pages 439–449, 1992.
- [Rao et Georgeff, 1995] A. S. Rao et M. P. Georgeff. BDI-agents : from theory to practice. In *the First Intl. Conference on Multiagent Systems*, San Francisco, 1995.
- [Rodríguez-Aguilar, 2001] J.A. Rodríguez-Aguilar. *On the Design and Construction of Agent-mediated Electronic Institutions*. PhD thesis, Institut d'Investigació en Intelligència Artificial (IIIA), 2001.

-
- [Russell et Norvig, 2003] Stuart J. Russell et Peter Norvig. *Artificial Intelligence. A Modern Approach*. Prentice Hall, 2003.
- [Searle, 1969] J. Searle. *Speech Acts : An Essay in the Philosophy of Language*. Cambridge University Press, Cambridge, 1969.
- [Searle, 1997] John R. Searle. *The Construction of Social Reality*. Free Press, January 1997.
- [Searle, 2005] John R. Searle. What is an institution ? *Journal of Institutional Economics*, 1(1) :1–22, 2005.
- [Shannon et Weaver, 1949] C. E. Shannon et W. Weaver. *The Mathematical Theory of Communication*. University of Illinois Press, Urbana, Illinois, 1949.
- [Singh, 1991] Munindar P. Singh. Social and psychological commitments in multiagent systems. *AAAI Fall Symposium on Knowledge and Action at Social and Organizational Levels*, pages 104–106, 1991.
- [Singh, 1997] Munindar P. Singh. Commitments among autonomous agents in information-rich environments. In *the 8th European Workshop on Modelling Autonomous Agents in a Multi-Agent World*, pages 141–155, London, UK, 1997. Springer-Verlag.
- [Singh, 1998] M. P. Singh. Agent communication languages : Rethinking the principles. *IEEE Computer*, 31(12) :40–47, 1998.
- [Singh, 1999] M. Singh. An ontology for commitments in multiagent systems : towards a unification of normative concepts. *Artificial Intelligence and Law*, pages 97–113, 1999.
- [Singh, 2000] Munindar P. Singh. A social semantics for agent communication languages. In *Agent Communication*, éditeurs Frank Dignum et Mark Greaves, pages 31–45. Springer-Verlag : Heidelberg, Germany, 2000.
- [Soulié, 2001] Jean-Christophe Soulié. *Vers une approche multi-environnements pour les agents*. Thèse de doctorat, Université de la Réunion, Décembre 2001.
- [Stratulat, 2002] Tiberiu Stratulat. *Systèmes d’agents normatifs : concepts et outils logiques*. Thèse de doctorat, Université de Caën, Décembre 2002.
- [Tranier et al., 2003] John Tranier, Jacques Ferber, et Frédéric Koriche. Un modèle efficace de normes pour les systèmes multi-agents organisationnels. *Technique et Science Informatiques*, 22(4) :205–217, 2003.
- [Tuomela, 1995] Raimo Tuomela. The importance of us : A philosophical study of basic social notions. *Stanford Series in Philosophy*, 1995.
- [Venkatraman et Singh, 1999] Mahadevan Venkatraman et Munindar P. Singh. Verifying compliance with commitment protocols. *Autonomous Agents and Multi-Agent Systems*, 2(3) :217–236, 1999.

- [Vold *et al.*, 2002] G. B. Vold, T. J. Bernard, et J. B. Snipes. *Theoretical Criminology*. Oxford University Press, 5th edition, 2002.
- [von Wright, 1951] G. H. von Wright. Deontic logic. *Mind*, 60 :1–15, 1951.
- [Vázquez-Salceda *et al.*, 2005] Javier Vázquez-Salceda, Virginia Dignum, et Frank Dignum. Organizing multiagent systems. *Autonomous Agents and Multi-Agent Systems*, 11(3), Novembre 2005.
- [Vázquez-Salceda et Dignum, 2003] Javier Vázquez-Salceda et Frank Dignum. Modelling electronic organizations. In *CEEMAS*, éditeurs Vladimír Marík, Jörg P. Müller, et Michal Pechoucek, volume 2691 de *Lecture Notes in Computer Science*, pages 584–593. Springer, 2003.
- [Weyns *et al.*, 2004a] Danny Weyns, H. Van Dyke Parunak, Fabien Michel, Tom Holvoet, et Jacques Ferber. Environments for multiagent systems state-of-the-art and research challenges. In *E4MAS*, éditeurs Danny Weyns, H. Van Dyke Parunak, et Fabien Michel, volume 3374 de *Lecture Notes in Computer Science*, pages 1–47. Springer, 2004.
- [Weyns *et al.*, 2004b] Danny Weyns, Elke Steegmans, et Tom Holvoet. Towards active perception in situated multi-agent systems. *Applied Artificial Intelligence*, 18(9-10) :867–883, 2004.
- [Wilber, 2001] Ken Wilber. *A Theory of Everything : An Integral Vision for Business, Politics, Science and Spirituality*. Shambhala, October 2001.
- [Wooldridge *et al.*, 2000] Michael Wooldridge, Nicholas R. Jennings, et David Kinny. The gaia methodology for agent-oriented analysis and design. *Autonomous Agents and Multi-Agent Systems*, 3(3) :285–312, 2000.
- [Wooldridge et Jennings, 1995] Michael Wooldridge et Nicholas R. Jennings. Agent theories, architectures, and languages : a survey. In *the workshop on agent theories, architectures, and languages on Intelligent agents (ECAI-94)*, pages 1–39, New York, NY, USA, 1995. Springer-Verlag New York, Inc.
- [Yang et Garcia-Molina, 2003] Beverly Yang et Hector Garcia-Molina. Designing a super-peer network. In *ICDE*, éditeurs Umeshwar Dayal, Krithi Ramamritham, et T. M. Vijayaraman, pages 49–. IEEE Computer Society, 2003.
- [Yolum et Singh, 2001] Pinar Yolum et Munindar P. Singh. Commitment machines. In *the Eighth International Workshop on Agent Theories, Architectures, and Languages (ATAL-2001)*, éditeurs John-Jules Meyer et Milind Tambe, pages 245–257, 2001.
- [Yolum et Singh, 2002] Pinar Yolum et Munindar P. Singh. Flexible protocol specification and execution : Applying event calculus planning using commitments. In *the 1st International Joint Conference on Autonomous Agents and MultiAgent Systems (AAMAS)*, pages 527–534. ACM Press, July 2002.

[Zambonelli *et al.*, 2003] Franco Zambonelli, Nicholas R. Jennings, et Michael Wooldridge. Developing multiagent systems : The gaia methodology. *ACM Trans. Softw. Eng. Methodol.*, 12(3) :317–370, 2003.

Résumé.

Dans cette thèse nous mettons en avant l'importance de considérer un SMA suivant tous ses aspects : agent, environnement, interaction, organisation et institution. Jusqu'à maintenant, ces aspects ont principalement été traités séparément, et il est difficile d'intégrer des modèles qui portent sur des aspects différents. De plus, nous constatons qu'il existe parfois une ambiguïté sur le rôle de chacun de ces aspects d'un SMA. Il y a donc un besoin de les clarifier et de permettre de tous les intégrer de manière cohérente. Dans cet objectif, nous proposons un cadre conceptuel original pour les SMA qui est fondé sur les quatre quadrants de la vision intégrale de Wilber. Ces quadrants résultent de l'intersection de deux axes d'analyse : l'axe interne – externe et l'axe individuel – collectif. Un intérêt de ce cadre conceptuel est de mettre en évidence le champ d'application de modèles existants, et de faciliter l'intégration de modèles complémentaires. De plus, nous montrons qu'il est adapté à la conception de systèmes ouverts. Enfin nous proposons le méta-modèle MASQ, qui est une formalisation de cette approche conceptuelle. MASQ permet de décrire un SMA à partir de quatre concepts fondamentaux (esprit, objet-corps, espace brut et culture), de relations entre ces concepts, et de lois d'évolution qui déterminent leur dynamique. Ce méta-modèle a pour objectif de mettre en relation des modèles spécifiques complémentaires pour la conception d'un SMA.

Mots-clés: système multi-agent, environnement, organisation, institution, vision intégrale

Abstract.

In this thesis we emphasize the importance of considering a MAS following all its aspects : agent, environment, interaction, organization and institution. Up to now, these aspects were treated in most cases separately, and it was difficult to find models that combined many of them. Besides, we note that there is sometimes an ambiguity on the role of each of these aspects of a MAS. It is therefore important to clarify them separately and to try to integrate them in a consistent way. Our objective is to provide an original conceptual framework for MAS which is founded on the four quadrants of Wilber's integral vision. The quadrants result from the intersection of two axes of analysis : the internal/external axis and the individual/collective axis. One of the advantages offered by the framework we propose is that it shows for any existent model the aspects it covers and gives hints on how to integrate it with complementary models. We also show that it is well adapted to the design of open systems. Finally, we propose the meta-model MASQ, which is a formalization of the conceptual framework. MASQ allows someone to describe a MAS using four fundamental concepts (i.e. mind, object-body, physical space and culture), the relations between these concepts, and the laws of evolution which determine their dynamics. The main objective of MASQ is to make possible to assemble specific and complementary models of MAS.

Keywords: multiagent system, environment, organization, institution, integral vision
