Philippe Pasquier

Novembre 2002

La cohérence cognitive comme fondement pour la pragmatique des communications entre agents

Directeur :
Prof. Brahim Chaib-draa

Comité d'encadrement :
Prof. Bernard Moulin
Prof. Mamadou Koné

Proposition de thèse

Doctorat en informatique
Département d'informatique et de génie logiciel
Université Laval, Ste-Foy, PQ, Canada

édité le 22 janvier 2003

Table des matières

Résumé

Les cadres interactionnels actuels pour les communications agent (protocoles, stratégies de conversation, jeux de dialogue, ...) garantissent la cohérence structurelle des conversations tenues. Pourtant, ce n'est pas tant l'habilité des agents à structurer leurs conversations qui nous intéresse que leurs aptitudes à tenir des conversations utiles quant à leurs objectifs individuels et collectifs. C'est à cette problématique issue d'un état de l'art fourni du domaine des théories de la communication en générale et de leurs applications aux systèmes multi-agents que nous souhaitons contribuer.

Pour ce faire, nous introduisons une théorie cognitive de l'utilisation de ces cadres interactionnels basée sur la notion de cohérence cognitive et fondée sur la psychologie sociale et la philosophie de l'esprit. En effet, notre proposition se structure méthodologiquement en deux étapes. La première consiste à unifier la théorie de la dissonance cognitive, une des théories majeures de psychologie cognitive avec la théorie de cohérence cognitive de Thagard. Cette unification nous permet de formuler notre approche en terme d'éléments et de contraintes, notions familières en informatique. La seconde étape consiste en une extension, afin d'y intégrer la communication, de la théorie de la cohérence cognitive obtenue. Sous les hypothèses de notre théorie, nous définissons ensuite une métrique de l'utilité des conversations. Nous montrons comment une telle théorie motivationnelle permet de répondre à certaines des problématiques habituelles des communications agent.

En particulier nous proposons une première application de notre théorie pour l'utilisation automatique par les agents de jeux de dialogue tels que définis conjointement avec Maudet et Chaib-draa dans le langage DIAGAL[DIAlogue Game based Agent communication Language]. Ce faisant, nous dégageons un certain nombre d'outils techniques pour l'automatisation des communications entre agents tout en précisant quels sont nos apports théoriques pour les SMAs et plus généralement pour les sciences cognitives.

1 Introduction générale

Comme l'a argumenté Craig [Craig, 1993], le grand nombre de théories relatives à la communication reflète la diversité des idées sur le sujet. Dès lors, si on ne peut trouver une théorie unificatrice des théories de la communication, il faut composer avec la multiplicité des approches. Selon Littlejohn, le but des recherches dans les domaines de la communication ne devrait plus être la recherche d'un hypothétique modèle standard qui rendrait le champ statique et « mort » [Littlejohn, 1992].

A l'inverse, dans le champ des SMAs[Systèmes Multi-Agents] à base d'agents cognitifs, le besoin d'un modèle de communication standard se fait sentir et les efforts se multiplient dans ce sens. Les technologies agents et multi-agents permettent de concevoir et de développer des applications complexes. La caractéristique fondamentale de celles-ci dans le paradigme actuel de l'informatique répartie est l'habileté des agents à communiquer entre eux de manière utile à leurs objectifs tant individuels que collectifs. Si de nombreux travaux se sont préoccupés de définir des langages de communication agents (notés ACL pour Agent Communication Language dans la suite), peu se sont concentrés sur les aspects cognitifs de la communication agent, c'est à dire sur les procédés par lesquels les agents vont utiliser dynamiquement et automatiquement ces langages de communication. En effet, si on suppose la souplesse et la cohérence structurale des dialogues engagés garanties par les contraintes fournies par le cadre interactionnel (ACL classiques à base d'actes de langage, protocoles ou jeux de dialogue), un certain nombre de problèmes subsistent[1] :

- *les problèmes liés à la gestion de la dynamique du dialogue* : Quand et quoi communiquer, à qui et pourquoi ? Comment choisir un interlocuteur ? Quel type de dialogue tenir ? À quel sujet ? Dans quel espoir ? Au sein d'un dialogue, comment sélectionner parmi les continuations possibles ? Comment choisir les degrés d'intensité des actes de langage utilisés ? Comment et pourquoi composer différents types de dialogues ?

- *les problèmes liés à l'indépendance des modèles cognitifs des agents et des cadres interactionnels pour la communication entre agents* : Comment garantir l'alignement de la satisfaction des aspects conventionnels des conversations, dictés par le cadre interactionnel, avec les aspects cognitifs des agents dialoguant ?

- *le problème original de la mesure de l'utilité du dialogue* : Quelle est l'utilité du dialogue pour un agent ? Quelle est l'utilité de la conversation pour le groupe d'agents impliqués ?

- *les problèmes liés à la gestion des conséquences sociales et cognitives des communications entre agents* : Quels sont les impacts des communications sur les agents au niveau cognitif de leurs états mentaux comme au niveau social de leurs accointances ?

Pour répondre ne serait-ce que partiellement à ces questions issues d'un état de l'art fourni (voir partie I et II), nous introduisons une théorie cognitive de l'utilisation de ces cadres interactionnels basée sur la notion de cohérence cognitive et fondée sur la psychologie sociale et la philosophie de l'esprit (Partie III). En effet, notre proposition se structure méthodologiquement en deux étapes. La première consiste à unifier la théorie de la dissonance cognitive [Festinger, 1957], une des théories majeures de psychologie cognitive et la théorie de cohérence cognitive de Thagard [Thagard, 2000], une des références de la philosophie de l'esprit

[1]Notons que ces problèmes ne sont pas indépendants les uns des autres.

et des sciences cognitives. Cette unification nous permet de formuler notre approche en terme d'éléments et de contraintes, notions familières en informatique. La seconde étape consiste en une extension de la théorie de la cohérence cognitive obtenue pour y intégrer la communication. Sous les hypothèses de notre théorie, nous définissons ensuite une métrique de l'utilité des conversations. Nous montrons comment une telle théorie motivationnelle permet de répondre à certaines des problématiques habituelles des communications agent.

En particulier nous proposons une première application de notre théorie pour l'utilisation automatique par les agents de jeux de dialogue tels que définis conjointement avec Maudet et Chaib-draa dans le langage DIAGAL[DIAlogue Game based Agent communication Language] [Chaib-draa et al., 2002]. Ce faisant, nous dégageons un certain nombre d'outils techniques pour l'automatisation des communications entre agents tout en précisant quels sont nos apports théoriques pour les SMAs et plus généralement pour les sciences cognitives. Finalement, nous discutons d'un échéancier pour les différentes tâches en cours et avenir qui seraient susceptibles de nous conduire au terme de ce travail de thèse (partie IV).

Première partie

Revue de littérature : la communication dans les SMAs cognitifs

2 Introduction

Faire le point sur l'étude des conversations entre agents (les langages de communications entre agents, ...) tout en rendant compte de leurs fondements est une tâche délicate. En effet, les recherches dans ce domaine empruntent (comme c'est l'habitude dans le domaine des SMAs) des notions à de nombreux domaines des sciences cognitives ou autres. On trouvera pêle-mêle : philosophie du langage, philosophie de l'esprit (épistémologie), socio-linguistique, linguistique, psychologie sociale, sociologie, dialectique, IA[Intelligence Artificielle] et IAD[Intelligence Artificielle Distribuée]. La stupéfiante diversité des approches ainsi que la complexité des cadres conceptuels développés, si elles rendent compte de l'importance et de l'envergure du domaine, rendent également toute tentative de synthèse du domaine hasardeuse et partielle.

Une première critique serait d'ailleurs de constater que malgré leur intérêt scientifique théorique fondamental indéniable, ces études se sont bien souvent éloignées de la réalité informatique. On ne présente généralement pas d'algorithme, on n'analyse pas les approches en terme de complexité computationnelle et on discute peu de l'implémentation des idées présentées. La frontière entre l'étude du dialogue pour ce qu'il est chez les humains et sa modélisation dans un cadre explicitement multi-agent est souvent confuse. Voyons pourquoi cette critique serait bien malvenue en rappelant pourquoi le langage naturel est la métaphore privilégiée des travaux sur les communications entre agents cognitifs.

2.1 Discussion : langage naturel et SMA.

Dans les systèmes multi-agents la communication est un point clé. Mais est-ce que, pour autant, les modèles de communication des agents doivent être basés sur ceux issus de la recherche sur le langage naturel ? Il y a évidemment plusieurs courants au sein des recherches concernant les SMAs. Nous nous limiterons ici aux SMAs cognitifs, basés sur le paradigme de représentation symbolique des connaissances et de formalisation du raisonnement issu du courant cognitiviste des sciences cognitives. L'interaction dans ce type de système requiert des techniques de communication plus sophistiquées que les solutions traditionnelles (passage de données ou appel de procédures). Même si on se limite, comme c'est souvent le cas, à la classe des systèmes coopératifs, il faut prendre en compte :

1. *l'hétérogénéité des agents* : les messages doivent être mutuellement compréhensibles alors que les points de vues des agents ne sont pas forcément mutuellement consistants ;

2. *l'échange de savoir* : un agent coopératif doit pouvoir manipuler des croyances sur les autres et en particulier sur leurs comportements, croyances et intentions. En effet, un tel agent doit être capable d'identifier et expliciter les conflits qui font obstacle à la résolution de ses problèmes. Pour ce faire, il doit pouvoir exprimer ces différents types de connaissances et non de simples données.

3. *le contrôle local* : les agents doivent être autonomes. C'est-à-dire que leur comportement ne doit pas dépendre d'un planificateur central ni d'interactions pré-définies. L'agent doit être capable de développer sa propre stratégie de communication dynamiquement. Notons que ce n'est pas le cas dans les systèmes actuels.

4. *la structure organisationnelle* : Pour éviter l'explosion combinatoire des communications au sein du système, il est commun d'avoir recours à une structure transactionnelle qui distribue les rôles ainsi que les relations hiérarchiques et les comportements attendus qui leur sont associés.

Dès lors que les humains parviennent à intégrer ces dimensions, il n'est pas inutile de prendre exemple sur les modèles de la communication humaine pour élaborer ceux des SMAs. L'observation des conversations humaines est la base de l'élaboration de protocoles sophistiqués. Les humains ont développé des techniques d'interaction très perfectionnées qui s'accommodent de leur rationalité limitée. Puisque les agents artificiels sont eux-aussi limités dans leur rationalité, on peut s'en inspirer.

Il faut aussi garder à l'esprit que le langage naturel a le plus grand pouvoir d'expression ! Ceci n'est pas sans conséquence lorsque l'on souhaite l'interopérabilité des différentes architectures d'agents. Pouvoir garantir que le langage développé, à l'instar du langage naturel, permet de tout dire et sous certaines conditions d'être compris de tous est un aspect majeur. Doter un cadre de communication inter-agent de ce type de pouvoir expressif supprime la tentation pour un développeur de spécialiser son système de communication, ce qui rendrait ses agents incompréhensibles pour ceux crées par d'autres développeurs. En outre, cela donne un cadre unificateur à des recherches souvent difficiles à comparer autrement. Notons finalement que cela participe du projet de l'IA cognitive et plus généralement des sciences cognitives.

Les chercheurs qui travaillent sur le dialogue et les conversations avec des outils informatiques poursuivent (indépendamment) les objectifs suivants :

– élaboration de systèmes de dialogue homme-machine en langage naturel ;

– conception de systèmes d'aide au dialogue, de médiatisation (collecticiel) ;

– élaboration de la composante interactionnelle de systèmes multi-agents ;

– modélisation linguistique ;

– modélisation cognitive ;

– modélisation physiologique et neuro-biologique.

Cette simple liste permet de bien comprendre que ces chercheurs ont un certain nombre d'ambitions qui si elles ne sont pas toujours incompatibles sont bel et bien différentes. Certaines recherches visent l'universalité des résultats et souhaitent embrasser le genre humain, d'autres étudient des types de dialogues particuliers (orienté tâche, supervisée, ...) ou des types de systèmes particuliers (SMA, IHM[Interfaces Homme-Machines], simulations de systèmes sociaux ou biologiques, ...). De ce champ d'une grande diversité, nous tenterons donc d'extraire les éléments pertinents pour les communications agents. Pour cela, il convient de préciser le type d'agents dit « cognitif » dont il est question dans les SMAs.

2.2 Hypothèses : caractéristiques des agents cognitifs

Au fur et à mesure que le domaine des SMAs se développe, les agents se complexifient. Ainsi, on prête aux agents intelligents actuels un certain nombre de caractéristiques, voyons ici les hypothèses qui sont faites concernant les agents cognitifs qui sont capables d'utiliser des « formes » dialogique :

– représentation explicite des connaissances ;

– structuration des états mentaux (BDI, BDIO) ;

– maintien d'un modèle des autres aussi complet que possible ;

– capacités sociales : gestion des engagements (cela suppose le raisonnement temporel), capacités de communication (manipulation d'un cadre interactionnel) ;

– capacités cognitives : apprentissage (ou au moins actualisation des connaissances : mise à jour et révision des connaissances), planification garantissant la proactivité et l'autonomie.

Tous ces éléments sont assemblés en une théorie comportementale qui abstrait un modèle de fonctionnement interne de l'agent (formalisation du raisonnement en dehors de toute instanciation) basé sur un modèle cognitif (représentation et structuration des connaissances). Les théories comportementales de ce type doivent fournir pour un agent des éléments concernant par exemple : sa stratégie de raisonnement, son modèle déductif/inductif, sa théorie de l'action et de la causalité, ses méthodes de planification et de satisfaction de buts, son système de dynamique de croyance et de révision de croyance, ses capacités, ...
[2]

Finalement, la façon de communiquer d'un agent doit être compatible avec son fonctionnement interne. On peut même dire que la capacité de communication d'un agent fait partie intégrante de son modèle cognitif et comportemental. Pour un agent donné, une conversation est un processus dynamique qui met en jeu l'essentiel de ses ressources cognitives.

3 Généralités sur le dialogue

3.1 Définition

Dans les sociétés humaines, un acte de langage est rarement utilisé seul. Dans la communication utilisant le langage naturel, un système de signes parmi d'autres, on distingue le discours (un locuteur), du dialogue (deux locuteurs) et de la conversation (au moins deux intervenants). Ce qui différencie le dialogue de la simple communication (le discours), c'est la

[2] De telles théories peuvent être composées, par exemple :

– d'éléments sur le savoir et l'action (Moore [Moore, 1990]) ;
– d'une architecture BDI (Rao et Georgeff [Rao and Georgeff, 1995]) ;
– de savoir-faire et gestion du temps (Singh [Singh, 1994]) ;
– d'une théorie de l'intention (Cohen et Levesque [Cohen and Levesque, 1990a]) ;
– d'éléments de théories économiques, théorie des jeux (Werner [Werner, 1992])

recherche d'une *inter-compréhension*. Cette recherche impose aux interlocuteurs de s'assurer qu'ils se comprennent pour co-construire des interprétations communes. Par exemple, un dialogue de sourds n'est pas un dialogue. Si l'un des interlocuteurs n'intègre jamais la vision de l'autre dans la sienne, la co-construction est bloquée. Cette recherche d'inter-compréhension peut être synchrone (dialogue oral[3]) ou asynchrone (courrier, courriel, forum de discussion).

La co-construction d'une interprétation commune ne signifie pas que les interlocuteurs doivent être d'accord. Ils peuvent avoir l'interprétation commune de leur désaccord. Ajoutons que cette co-construction semble réussir plus ou moins selon un grand nombre de facteurs inhérents aux interlocuteurs (proximité culturelle et sociale, historique de leur relation, ...). La recherche d'une inter-compréhension optimale est facilitée par :

- *la recherche d'un langage commun* : sans langage commun, le dialogue est difficile. Notons que cette recherche se poursuit au cours du ou des dialogues par la construction de références communes [Clark and Wilkes-Gibbs, 1986] ;

- *la capacité à détecter les ambiguïtés et les incohérences dans le discours d'autrui* ;

- *la production et la recherche de retours (feedback) de compréhension* : pour s'assurer de leur compréhension réciproque, les interlocuteurs fournissent et cherchent chez l'autre des indices de compréhension. Ces indices peuvent être positifs (manifestations d'une bonne compréhension : signes d'acquiescement verbaux ou non-verbaux) ou négatifs (demande de clarification et d'explication), explicites ou implicites.

- *les échanges correctifs* : la qualité de l'inter-compréhension repose sur des interprétations. Comme il peut toujours y avoir erreur d'interprétation, la capacité à corriger sa compréhension ou celle d'autrui est fondamentale.

- *la capacité à méta-communiquer* : l'évitement et la résolution de conflits de compréhension peuvent nécessiter de dialoguer sur l'activité de dialogue en cours.

Toutes ces caractéristiques sont à la base d'une propriété plus générale du dialogue : l'interactivité. Il y a interactivité lorsque les actions des uns sont influencées par celles des autres. Cette propriété a une implication majeure : la non-prédictibilité globale de la structure du dialogue. Ces vingt dernières années, les linguistes structuralistes ont entrepris de nombreux travaux sur le dialogue. Les outils structuralistes étant des outils d'analyse structurelle, ces tentatives se sont soldées par des échecs. En effet, contrairement aux phrases ou même au discours, les conversations ne sont généralement pas l'expression de structures régulières, mais plutôt de co-constructions complexes. Néanmoins, on peut tout de même différencier les conversations orientées, cadrées ou contraintes, d'autres plus libres. Par exemple, des conversations orientées telles que les enchères, les entretiens d'embauche, les interrogatoires, les dialogues patient-docteur lors d'une consultation, ou encore les échanges entre élève et professeur sont plus structurés que des conversations libres entre amis. On dira que ces conversations contraintes sont orientées structure et les conversations libres orientées processus. Le tableau 1 récapitule les différences entre ces deux types de conversation en reprenant les caractéristiques majeures du dialogue contraint et du dialogue libre indiquées par Searle [Searle, 1992b].

[3]Éventuellement porté par écrit, comme les pièces de théâtre ou les exemples de ce texte.

Orienté processus	**Orienté structure**
Peu de contraintes sur le type de contribution	Beaucoup de contraintes sur le type de contribution
De multiples buts essentiellement locaux	Peu de buts globaux
Ordre d'intervention assez libre	Ordre d'intervention assez fixe
Rôle des participants pas très bien défini	Rôle des participants bien défini
Contribution dépendante des contributions passées	Contribution rarement dépendante des contributions passées
Principe d'organisation local	Principe d'organisation global

TAB. 1 – Dialogue orienté structure vs. orienté processus selon Chaib-draa et Vongkasem [Chaib-draa and Vongkasem, 2000].

3.2 La conversation comme action commune

Il existe de nombreux travaux sur la sémantique des actes de langage (voir l'annexe A). Mais peu d'études formelles traitent de la sémantique d'une conversation. On pourrait penser que la sémantique d'une conversation est obtenue par composition des sémantiques des actes de langage qui la compose. Il n'en est rien. Illustrons cela à l'aide d'un exemple. Le graphe de conversation pour l'action de Winograd et Flores [Winograd and Flores, 1986] montre qu'une simple demande est un processus complexe composé de plusieurs actes de langage impliquant le demandeur, mais aussi l'agent à qui la demande est faite (voir figure 1). Une telle demande est une action commune complexe et la sémantique des actes de langages qui composent le graphe ne saurait rendre compte de la sémantique de la conversation dans sa généralité (en particulier sa structure).

Cet exemple, parmi d'autres, montre que la communication entre agents se fait par conversation, séquence d'actes de langage dont la somme des sens isolés ne rend pas compte de la signification. C'est pourquoi, de nombreux chercheurs rendent aux conversations leur *dimension sociale* et tentent d'analyser le dialogue du niveau conversationnel vers le niveau des actes de langage plutôt que le contraire. Cohen et Levesque ont proposé leur point de vue avec les buts persistants et les croyances mutuelles [Cohen and Levesque, 1990b]. Grosz a proposé les plans partagés [Grosz and Kraus, 1996], Singh l'engagement commun [Singh, 1994] et Traum étudie les croyances mutuelles [Traum, 1997]. Chaib-draa et Vongkasem partent dans la même voie avec les idées de Searle et Clark qui voient la conversation comme une *activité commune* des agents qui y participent [Chaib-draa and Vongkasem, 2000]. Cette activité commune est une séquence d'actions dont la plupart sont des *actions communes* [Clark, 1996]. Pour Clark, une activité commune est :

1. une activité qui implique au moins deux participants ;
2. les participants jouent un rôle au sein de cette activité ;
3. les participants cherchent à accomplir certains buts collectifs ;

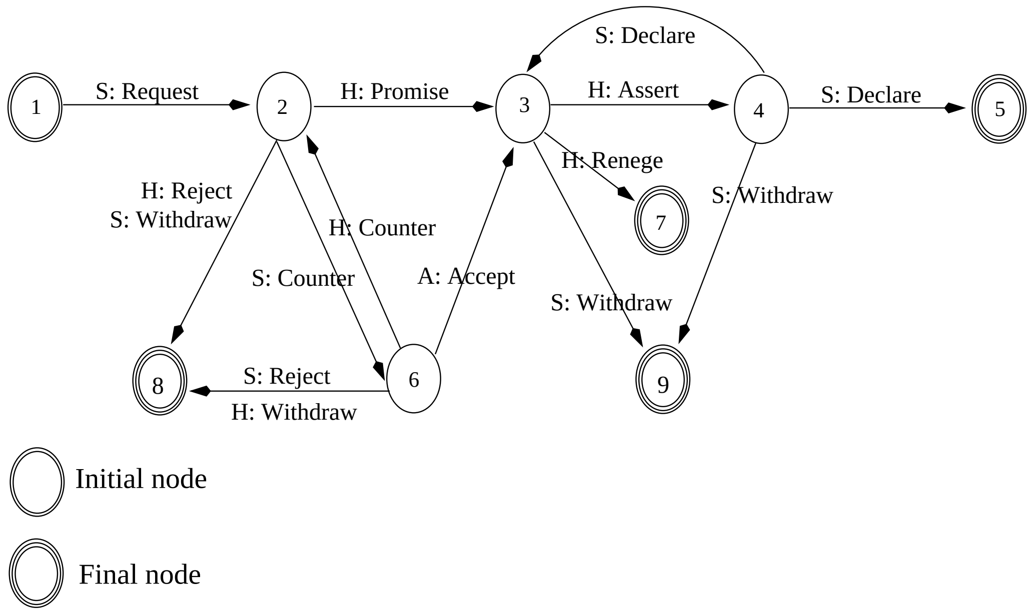

FIG. 1 – Graphe d'une requête par Winograd et Flores.

4. les participants cherchent à accomplir certains buts individuels/privés ;

5. uns activité commune se compose d'actions individuelles ou d'actions communes. Mais une action commune est plus que la somme des actions individuelles des participants. La coordination, notamment, est en plus. ;

6. les participants peuvent utiliser des procédures spécifiques pour atteindre leurs buts ;

7. les participants s'accordent sur le début et la fin de l'activité ;

8. l'exécution des procédures constituantes de l'activité peut être simultanée ou séquencée.

Comme exemple d'activité commune, citons le duo en musique et plus proche du langage linguistique : les dialogues de recherche d'information comme les paires question-réponse. Idéalement, tous les types de dialogues (voir section 3.4) possèdent les caractéristiques d'une activité commune. Pour approfondir cette idée d'action commune, la section suivante présente les idées majeures concernant les éléments et processus impliqués dans la mise en commun.

3.3 Arrière-plan, fond commun et établissement

La notion d'arrière-plan[4] est la clé de voûte de la philosophie Searlienne [Searle, 1992a]. L'idée en est simple : les mots et les phrases ne suffisent pas en eux-mêmes à générer une interprétation. Le même sens linguistique admettra des interprétations différentes selon les usages[5]. Par exemple, le verbe ouvrir s'interprète différemment dans « ouvrir les yeux », « ouvrir un

[4]Les termes arrière-plan, fond commun et établissement sont les traductions « fragiles » de : background, common ground et grounding dont il est difficile de rendre exactement compte en français.

[5]La notion d'arrière-plan inclut celle de contexte (voir en annexe A.3) mais ne s'y réduit pas.

restaurant », « ouvrir son coeur », « ouvrir une porte », « ouvrir le débat ». Pour Searle, tout ce qui est sens s'appuie sur un ensemble d'aptitudes, de dispositions et de capacités dites d'arrière-plan. L'arrière-plan ne fait pas partie du sens et pourtant le sens n'existe en tant que tel que par rapport à lui[6]. En première approximation, on réduit l'arrière-plan à ce qui est dit et communément accepté dans le cadre d'une conversation. A cet égard, la notion de croyance commune introduite par Clark et Schaeffer avec la notion d'établissement [Clark and Schaeffer, 1987] semble acceptable. Une information est d'abord présentée, comprise (retour de compréhension) puis acceptée : on dira alors qu'elle est socialement établie. Elle devient alors commune. Cette façon de voir évite le problème de la récursivité infinie des connaissances mutuelles[7] : je sais qu'il sait que je sais qu'il sait que je... Ce qui rend néanmoins difficile l'implémentation de ces notions, c'est la diversité des types de présentation (directe,co-construite,...) et d'acceptation (implicite par continuation ou explicite comme « OK » ou encore par une sous-conversation,...). Ainsi, il est souvent difficile de dire si un énoncé fait partie de la présentation ou de l'acceptation.

3.4 Typologie des dialogues

Une typologie exhaustive des dialogues est illusoire puisqu'il existe une infinité de types de dialogues [Wittgenstein, 1953]. Mais, en se restreignant aux dialogues à but discursif, Vanderveken propose une classification des dialogues discursifs basée sur les directions d'ajustement [Vanderveken, 1999] :

1. *descriptif* : les interlocuteurs dialoguent pour décrire l'état du monde ou se mettre d'accord sur cette description. C'est le cas des dialogues de persuasions, des investigations, des recherches d'informations, des enquêtes, des examens, ...

2. *délibératif* : les interlocuteurs dialoguent pour s'engager ou engager les autres à réaliser certaines actions. C'est le cas des dialogues de négociation ou de délibération, ...

3. *déclaratif* : les interlocuteurs dialoguent dans le but de réaliser une déclaration commune. C'est le cas pour les dialogues tenus par des assemblées ou par un jury, ...

4. *expressif* : les interlocuteurs dialoguent pour exprimer leurs attitudes. C'est le cas des hommages, des bravos, des félicitations, ...

Dans le même esprit que pour la théorie des actes de langage (présentée en annexe A.1), ces quatre types de dialogue sont raffinables selon leurs composantes : mode d'atteinte du but discursif, conditions thématiques, conditions d'arrière-plan, conditions de sincérité.

Une autre classification, issue des travaux en dialectique est celle proposée par Walton et Krabbe [Walton and Krabbe, 1995]. Ils définissent cinq types de dialogues principaux qui sont caractérisés par le but global et commun du dialogue et les buts privés de chacun des interlocuteurs :

1. *persuasion* : le but global est de résoudre un conflit, le but privé de chacun des intervenants est de convaincre son ou ses interlocuteur(s) ;

[6]Cette idée est déjà en germe chez les philosophes de l'école Gestalt. Notons en outre que Searle pose cet arrière-plan langagier comme infini et non représentable, ce qui l'amène à conclure l'existence d'une complexité irréductible de la conscience.

[7]En fait, le problème resurgit dès que l'on pense à distribuer cet objet commun.

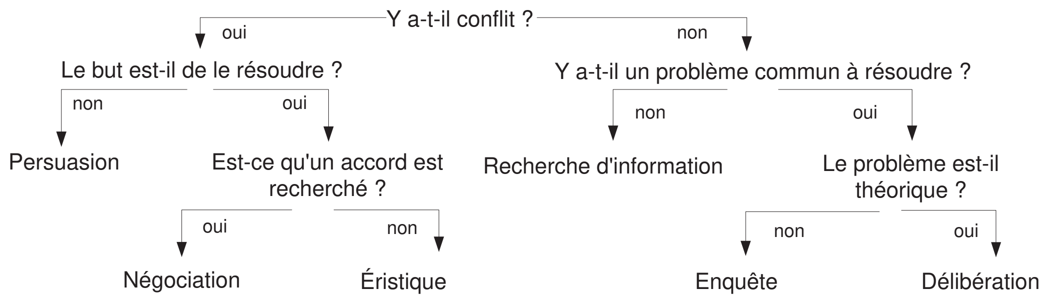

FIG. 2 – Situation initiale et type de dialogue.

2. *négociation* : le but global est de résoudre un conflit en atteignant un accord, chacun des agents ayant pour but privé de maximiser sa satisfaction ;

3. *investigation (ou enquête)* : le but global est d'établir la véracité d'un fait, les buts individuels sont de participer à ce processus ;

4. *délibération* : le but global est de s'entendre sur un plan, chacun travaillant à l'influencer à son avantage ;

5. *recherche d'informations* : le but global est de trouver de l'information. C'est le seul type de dialogue qui ne soit pas symétrique, seul un des agents a pour but privé d'obtenir de l'information.

On note notamment que certains types de dialogues mentionnés par Vanderveken sont absents de la classification de Walton et Krabbe (pas d'expressif). Walton et Krabbe ont aussi systématisé les conditions initiales de ces différents types de dialogues comme indiqué par la figure 2.

Ces classifications ne signifient pas que les dialogues soient d'un seul et même type tout au long de leur déroulement. Les dialogues observables dans la réalité seraient plutôt des compositions parfois complexes de ces types de base. On parle alors de *dialogue de types complexes*.

En outre, le choix du type de dialogue dépend du contexte. Bunt [Bunt, 1996] distingue les aspects statiques des aspects dynamiques du contexte. Pour les aspects statiques, il identifie :

– *le contexte physique et perceptuel* :

  – *co-présence* : c'est la situation de dialogue élémentaire (parce que la plus commune), le « face à face », dialogue oral, à un endroit donné et à un moment précis.
  – *non co-présence* : le dialogue est médiatisé (téléphone, courriel, courrier) par l'oral ou l'écrit.

– *le contexte social* : les rôles sociaux des interlocuteurs sont porteurs de droits et d'interdictions liés à l'activité dialogique (par exemple : prof-élève, policier-témoin, patient-médecin, directeur-employé, client-vendeur, . . .). Par exemple, l'interlocuteur qui a l'initiative est souvent celui qui guide la discussion. Dans le cas général, on parle de dialogues d'initiatives mixtes.

– *le contexte cognitif* : ce contexte contient les différentes attitudes mentales des interlocuteurs. Il est donc essentiellement dynamique. Les aspects statiques qui peuvent être isolés en début de dialogue sont : les identités des interlocuteurs, les capacités qu'on leur prête (on attribue certaines capacités, différentes, à un enfant ou à une machine, ...) et le but principal de la conversation qui explique l'entrée en dialogue.

On peut rendre explicite une hypothèse implicite à tous les travaux présentés dans la suite de ce document : on ne considère que les dialogues oraux médiatisés (cas de non co-présence), en laissant de côté les gestes, les mimiques et autres postures propres aux humains.

3.5 De l'énoncé au dialogue

Pour ce qui est de la représentation des énoncés, c'est la théorie des actes de langage qui est communément admise et utilisée. L'annexe A présente les théories des actes de langage et leurs instanciations dans le cadre des systèmes multi-agents : les langages de communications agents, notés ACL[Agent Communication Language] dans le reste du texte.

Classiquement, on considère deux grandes familles d'approches pour caractériser la structuration des conversations : les approches intentionnelles et les approches conventionnelles. Avant de les détailler (sections 4 et 5), il nous faut introduire quelques éléments qui permettent de bien comprendre les difficultés du passage de l'énoncé au dialogue.

Historiquement, les unités conversationnelles ont été étudiées avant le cadre conversationnel dans lequel elles doivent prendre place. Lorsqu'il a fallu les utiliser pour les conversations, de nombreux problèmes sont apparus. Les problèmes de la théorie des actes de langage apparaissent lorsque l'on considère certaines caractéristiques du dialogue :

– *le caractère non monologique du dialogue* : l'interlocuteur est plus qu'un simple auditeur ;

– *le caractère multidimensionnel du dialogue* : en plus d'être utilisé pour effectuer une tâche, il est également utilisé pour se contrôler lui-même.

De nombreux auteurs ont reproché aux actes de langage leur caractère monologique. Pourtant, il y a chez Austin la notion d'uptake [Austin, 1962a].

Exemple :
A : Vous allez à Berlin ?
B' *: Oui*
B'' *:Oui, montez !*
B''' *: Vous pouvez monter.*

A formule un acte de langage typiquement indirect (question fermée ou requête) et c'est la réponse (l'uptake) de B qui détermine le sens (la force illocutoire, notamment) de l'acte de A. Malheureusement, cette notion d'uptake n'a pas été formalisée dans la théorie des actes de langage.

La première idée face à ces manques de la théorie des actes de langage fut d'adapter celle-ci au dialogue. Trois types de solutions pour prendre en compte les spécificités du dialogue dans la théorie des actes de langage ont été proposées : (1) dialogisation des actes de langage, (2) actes de dialogue et (3) actes multi-niveaux.

3.5.1 Dialogisation des actes de langage

Plusieurs extensions de la théorie des actes de langage ont été proposées en vue d'utiliser les actes de langages dans un cadre dialogique :

- Allwood [Allwood, 1976] introduit la notion d'effet perlocutoire attendu (fonction expressive et fonction évocatrice) ;
- Brassac [Brassac, 1994] dialogise la logique illocutoire et en fait une logique inter-locutoire ;
- Sadek [Sadek, 1991a] introduit la notion d'effet rationnel ;
- Bunt [Bunt, 1996] et les partisans de l'approche contextuelle distinguent les actes qui vont de l'avant (nommés « forward-looking » et qui sont plutôt initiatifs) des actes qui suivent (traduit de « backward-looking » et qui sont plutôt réactifs). Le problème c'est que l'étude des signaux de retour (feedback) montrent qu'ils sont généralement ambigus, essentiellement à cause des différents niveaux de l'interaction (voir les actes multi-niveaux présentés ci-dessous).

3.5.2 Actes de dialogues

Bunt [Bunt, 1996] a proposé une théorie des actes de dialogues. Selon Bunt, on distingue les actes qui font progresser vers l'accomplissement de la tâche sous-jacente (TO-acts, pour task-oriented acts) des actes de contrôle du dialogue (DC-acts, pour dialogue control acts). Ces derniers sont de trois types : les retours (feedbacks), les obligations sociales et la gestion de l'interaction (tour de parole, structuration du dialogue, contact entre les interlocuteurs).

3.5.3 Les actes multi-niveaux

Traum et Hinkelman [Traum and Hinkelman, 1992] ont proposé (et utilisé) les actes de conversation multi-niveaux. On y retrouve des actes de langage noyaux qui sont les actes de langages classiques, mais qui ne prennent leur plein effet qu'après établissement (proposition et acceptation, voir section 3.2). Ils sont considérés comme des actions communes qui sont ajoutées au fond commun plutôt que comme de simples actions du locuteur. Cette théorie des actes conversationnels suppose en outre trois autres types d'actes de langage : les actes de tour de parole, les actes d'établissement et les actes d'argumentation, plus complexes, car ils mettent en jeu plus d'un acte noyau. Les actes conversationnels sont récapitulés dans le tableau 2.

3.6 Conclusion

Les propositions de modèles de structuration des dialogues faites jusqu'alors peuvent être regroupées en deux familles :

- les approches intentionnelles : les enchaînements d'énoncés résultant des intentions des interlocuteurs supposés coopératifs ;
- les approches conventionnelles : qui mettent l'accent sur les contraintes structurelles liées aux dialogues, que les interlocuteurs sont socialement poussés à observer.

les sections suivantes détaillent chacune de ces familles d'approches.

Niveau de discours	Type d'actes	Exemples
SUB-UU	tour de parole	prendre-tour, garder-tour, donner-tour, assigner-tour
UU	établissement	initier, continuer, valider, réparer, demande-réparer, demande-valider, accepter, rejeter
Unité de discours (DU)	actes-noyau	informer, demander, s'engager
Multiple DUs	argumentation	élaborer, résumer, clarifier, question-réponse, convaincre, trouver un plan

Tab. 2 – Actes conversationnels [Traum and Poesio, 1997].

4 Approches intentionnelles

4.1 Fondements philosophiques

4.1.1 Intention

L'intention est une notion très étudiée en philosophie de l'esprit. Les travaux les plus influents en IA sont ceux de Searle [Searle, 1983] et Bratman [Bratman, 1987]. On distingue classiquement : l'intention dirigée vers le futur (avoir l'intention de faire quelque chose) et action intentionnelle (faire quelque chose intentionnellement). Ces deux notions sont liées puisque l'intention dirigée vers le futur mène généralement à la réalisation d'actions intentionnelles. L'intention qui nous intéresse ici est l'intention dirigée vers le futur.

Il faut se garder de confondre désir et intention. Les désirs représentent des états souhaités du monde, ils peuvent être irréalisables ou bien contradictoires. Le processus qui sélectionne les désirs qui pourront être poursuivis est *la délibération*. La délibération concerne la production des buts (les liens entre les notions de but et d'intention soulèvent de nombreux débats) et a été beaucoup étudiée en IA.

En adjoignant à ces deux ingrédients les croyances, nous nous retrouvons dans le cadre intentionnel classique BDI[Belief, Desire, Intention]. Dans la délibération de sens commun (encore appelé raisonnement pratique), les buts viennent des désirs et les désirs peuvent être inaccessibles. En outre, les buts peuvent ne pas être consistants.

4.1.2 Signification non-naturelle

En linguistique, la notion de signification a longtemps été considérée comme le seul fait des énoncés. Dans la tradition « codique » de la communication, le locuteur « code » un sens dans un énoncé qu'il transmet à un auditeur qui le « décode ». Or, l'énoncé ne représente pas la totalité de la signification. La thèse de Grice [Grice, 1957], père de la vision inférentielle de la communication, est qu'il existe un rapport entre le vouloir dire du locuteur et ses intentions. C'est ce qu'il appelle la signification non-naturelle en opposition à la relation naturelle qui existe entre signe et signification. La formulation de Levinson [Levinson, 1983] est la suivante. Le locuteur A a voulu dire z en exprimant e, si et seulement si :

1. A a l'intention que e provoque l'effet z chez B ;

2. A a l'intention que l'intention dont il est question en 1 soit réalisée par B en reconnaissant celle-ci (cette seconde intention est aussi appelée *intention communicative*).

L'intention est donc placée doublement au coeur du processus communicatif. On considère, d'une part, l'intention qu'a le locuteur en produisant l'énoncé et d'autre part, la reconnaissance de cette intention par l'auditeur. Lorsque, par exemple, je demande si p, (1) j'ai l'intention de savoir si p et (2) je souhaite que ceci soit réalisé par la reconnaissance de cette intention. Dans la littérature, les situations dans lesquelles l'intention (le but) du locuteur n'a pas pour vocation d'être reconnue sont appelées « non manifestes » (covert). C'est le cas, par exemple, lorsque le locuteur essaie d'impressionner l'interlocuteur.

4.1.3 Coopération

Grice a le premier défini la notion de coopération dans le cadre du dialogue [Grice, 1975]. Des interlocuteurs rationnels et coopératifs, désireux de maximiser l'échange d'information lors d'un dialogue devront se conformer au *principe de coopération* :

> Que votre contribution conversationnelle corresponde à ce qui est exigé de vous, au stade atteint par celle-ci, par le but ou la direction acceptée de l'échange dans lequel vous êtes engagé[8].

Ce principe peut se décliner sous la forme de quatre maximes de coopération :

- *maxime de qualité* : ne pas dire ce que vous croyez faux ou ce pourquoi vous manquez de preuve ;
- *maxime de quantité* : faites que votre contribution soit aussi informative que le but de l'échange le requiert, mais pas plus ;
- *maxime de pertinence* : soyez pertinent ;
- *maxime de manière* : soyez bref et ordonné, évitez les ambiguïtés.

Ces maximes paraissent simples, simplistes même, mais elles ne doivent pas être comprises comme strictement normatives. En effet, c'est en se basant sur le fait qu'elles sont régulièrement transgressées que peut être amorcé un processus inférentiel (appelé implicature) qui va permettre de retrouver le véritable sens des énoncés. Lorsqu'une (ou plusieurs) maxime est violée, les interlocuteurs réagissent coopérativement (le principe de coopération s'applique toujours) et essaient de comprendre/interpréter la violation : ce sont *les implicatures*. Les interlocuteurs cherchent à savoir ce qui est impliqué par cette violation apparente et manifeste.

Lorsque, par exemple, on dit : « Il pleut des cordes. », on viole la maxime de qualité, mais les autres comprennent bien que l'on parle métaphoriquement. C'est un exemple d'implicature conversationnelle. Les actes indirects sont un autre type d'implicature. Si on me demande : « Combien ça t'a coûté ? », et que je répond : « Assez cher. ». Un interlocuteur comprendra, en supposant que je reste coopératif, que ce viol de la maxime de quantité est à interpréter comme signifiant : « Ce ne sont pas tes oignons ». Souvent, les violations sont indiquées, anticipées explicitement. Par exemple, on dit : « J'exagère peut-être un peu, mais… » pour prévenir

[8]Traduction libre de [Grice, 1975].

d'un viol de la maxime de quantité, ou encore : « Tu vois ce que je veux dire... » pour signifier que l'on a violé la maxime de manière.

Lorsque le locuteur ne suit pas le principe de coopération, mais qu'il fait en sorte que les autres croient qu'il le suit, les violations peuvent servir à tromper les autres. Pour mentir, on viole discrètement la maxime de qualité, pour obscurcir ou embrouiller, on viole la maxime de quantité et pour distraire, on viole la maxime de pertinence.

Une des approches néo-gricéenne célèbre est la théorie de la pertinence de Sperber et Wilson [Sperber and Wilson, 1986] qui ne conserve que la maxime de pertinence pour en faire un principe : *le principe de présomption optimale de pertinence*. Dans leur théorie, la pertinence d'un énoncé est évaluée en fonction du rapport entre les effets qu'il produit et les efforts que la production de ces effets demande. Les éléments qui permettent d'évaluer cette pertinence sont malheureusement difficiles à déterminer.

Ajoutons pour clore cette introduction aux travaux sur le concept de coopération qu'il ne faut pas confondre la coopération dialogique et la coopération comportementale. La première est nécessaire pour assurer la cohérence conversationnelle du dialogue alors que la seconde est une hypothèse sur le comportement extra-linguistique de l'agent [Airenti et al., 1993].

4.2 Structure intentionnelle du dialogue

L'idée qui sous-tend toutes les approches intentionnelles est la suivante : la structure du dialogue n'est qu'un épiphénomène qui résulte des intentions (et éventuellement de la coopérativité) des interlocuteurs. Cette idée a été généralisée par Grosz et Sidner [Grosz and Sidner, 1986] pour qui les trois éléments constitutifs de la structure du dialogue sont :

- *la structure linguistique* : ce sont les énoncés d'un dialogue agrégés en segments de dialogue (DS, Dialogue Segment) ; plus éventuellement des relations d'emboîtement entre eux ;
- *la structure intentionnelle* : il est possible de distinguer pour chaque segment de dialogue un but propre (DSP, Discourse Segment Purpose), sous-but du but global du dialogue. Les DSPs peuvent être liés entre eux par deux types de liens : (1) les liens de domination et (2) les liens de satisfaction ;
- *l'état attentionnel* : c'est ce sur quoi se focalise dynamiquement l'attention à un moment du dialogue. Cet état est généralement modélisé par une pile.

Selon Grosz et Sidner, la structure linguistique traduit la structure intentionnelle, en particulier, les relations d'emboîtement de la structure linguistique sont le reflet de cette structure intentionnelle. Ces auteures précisent toutefois : « la structure intentionnelle n'est pas isomorphe à la structure de la tâche sous-jacente ».

De nombreuses recherches ont été effectuées dans l'espoir de formaliser cette structure intentionnelle. Le domaine de l'intelligence artificielle concerné par la coordination d'actions en vue d'atteindre un but est la planification. C'est donc une approche par la planification qui s'est naturellement imposée comme première approche pour formaliser la structure intentionnelle.

4.2.1 Approches par la planification classique

Le premier système qui utilise explicitement la planification pour le dialogue est celui de Allen, Cohen et Perrault [Cohen and Perrault, 1979, Allen and Perrault,], inspiré des travaux

TAB. 3 – Caractéristiques de l'action $inform$.

Entête	$Inform(A, B, p)$
Préconditions	$Know(A, p)$
Corps	$MB(A, B, Want(A, Know(B, p)$
Effets	$Know(B, Know(A, p))$ et $Know(B, p)$

TAB. 4 – Éléments pour l'inférence de plans.

X	Y
Action	Effet
Précondition	Action
Corps	Action

de Bruce [Bruce, 1975] qui fut le premier à utiliser les actes de langage pour des travaux d'IA sur les actions et les plans. Dans ce système, les intervenants sont dotés d'états mentaux ($Want$ pour l'intention, Bel pour la croyance, $Know$ pour la connaissance, c'est-à-dire la croyance justifiée et MB pour la croyance mutuelle). Les actes de langages sont représentés comme des actions quelconques qui affectent les croyances des interlocuteurs, comme l'illustre le tableau 3 pour l'opérateur $inform$.

Le système de planification utilisé est une version modifiée du STRIPS de Fikes et Nilsson [Fikes and Nilsson, 1971]. L'hypothèse centrale est que les actes de langage sont planifiés au même titre que les autres actions. Les plans sont reconstruits par l'interlocuteur à partir de la connaissance des opérateurs et de règles de construction de plan de la forme :

$Want(A, X) \Rightarrow Want(A, Y)$ où X et Y sont donnés par le tableau 4.

On serait donc tenté de dire que la reconnaissance d'intention implique simplement la mise en oeuvre de règles permettant à l'interlocuteur, sur la base de ces règles de construction d'inférer le plan du locuteur, comme avec la règle suivante :

$Bel(B, Want(A, X)) \Rightarrow Bel(B, Want(A, Y))$ où X et Y sont de nouveau donnés par le tableau 4

C'est vrai, mais ce n'est pas si simple, car il convient de faire une distinction entre deux types de reconnaissance de plans :

- *la reconnaissance à l'insu* : en inversant les règles d'élaboration de plan, l'auditeur peut retrouver le plan du locuteur. Cette reconnaissance est indépendante de l'intention communicative du locuteur. C'est un aspect très important de la communication puisque c'est ce qui permet de devancer le locuteur en donnant des réponses pleinement coopératives.
- *la reconnaissance d'intention communicative* : la reconnaissance à l'insu n'est pas suffisante pour reconnaître le plan que le locuteur souhaite transmettre. Celle-ci nécessite

un niveau d'imbrication supplémentaire : $Bel(B, Want(A, Bel(B, Want(A, X))))$. Cette reconnaissance est nécessaire pour prendre en compte les actes de langage indirects.

Le principe de coopération prend la forme de l'*adoption de but* : après avoir inféré le but du locuteur et son plan pour le réaliser, un auditeur coopératif donnera l'information manquante du plan de manière à ce que le locuteur puisse réaliser son but.

Cette approche d'une grande importance historique souffre d'au moins deux problèmes :

– on se cantonne à l'analyse d'un seul énoncé (une paire d'adjacence si on considère la réponse donnée). Or, la reconnaissance de plan est souvent le résultat de véritables échanges (reconnaissance incrémentale, [Carberry, 1990])

– avec cette approche, la structure du dialogue est calquée sur la structure de la tâche sous-jacente. De nombreux dialogues ne respectent pas cet isomorphisme. C'est notamment le cas lorsqu'un sous-dialogue de clarification est entamé (il n'est généralement pas prévu dans la tâche, ...).

Cette seconde difficulté a donné lieu à la différenciation explicite par Litman et Allen de deux types de plans [Litman and Allen, 1990] :

– les *plans de domaine* modélisent les tâches extra-linguistiques sous-jacentes au dialogue ;

– les *plans de discours* sont des méta-plans qui permettent de manipuler la structure d'autres plans. Ils y a trois types de relations entre les plans du discours et ceux du domaine : (1) relation de continuation (permet de commencer l'exécution du plan de domaine ou d'en poursuivre le déroulement), (2) relation de clarification (identifier un paramètre, proposer une correction du plan courant, ...) et (3) relation de changement de sujet (permet d'introduire un nouveau plan de domaine). Sur la base de ces trois relations, cinq méta-plans ont été proposés : `suivre plan`, `identifier paramètre`, `corriger plan`, `introduire plan`, `modifier plan`.

La structure globale du dialogue est modélisée par une pile de plans du domaine et de méta-plans, elle-même composée d'autant de sous-piles que le plan de domaine a de pas ou d'étapes. Carberry [Carberry, 1990] précise que ce modèle, bien que restrictif, semble assez conforme à la majorité des comportements dialogaux observés.

Pourtant, il peut arriver que les participants au dialogue possèdent plusieurs solutions alternatives pour un même problème, et qu'en conséquence, ils souhaitent évaluer les différents plans à leur disposition. Pour ce faire, le modèle de Ramshaw [Ramshaw, 1989], reprit depuis par Lambert et Carberry [Lambert and Carberry, 1991] propose donc un troisième niveau (intermédiaire) de résolution de problème (ou élaboration de plan) qui permet aux interlocuteurs de considérer plusieurs possibilités de plans leur permettant de réaliser les buts du niveau domaine.

4.2.2 Approches des plans par les attitudes mentales

Une autre approche de la planification préconise la traduction des plans en terme d'états mentaux plutôt qu'en terme de structure de données. En effet, les plans vus comme de simples structures de données posent des problèmes :

– il y a redondance des liens de cause à effet, car la même action représentée différemment évoquera l'une ou l'autre ;

– la notion de pré-condition est ambiguë : est-ce que l'action est déclenchée lorsque les pré-conditions sont remplies ou est-ce que cela rend juste l'action possible ?

Ainsi, pour Pollack [Pollack, 1990], il existe deux types de relations entre les actions : « rendre possible » et « génération ». Il existe une relation « rendre possible » entre deux actions si la réalisation de la première permet la réalisation de la seconde. La relation de génération est plus forte puisqu'elle indique qu'une action sera réalisée par la réalisation d'une autre. Pollack distingue les *recettes* des plans. Formellement, une recette est une suite d'actions pour réaliser une action plus complexe. Comme certaines des actions de la recette peuvent être des actions complexes, une recette complète se représente sous la forme d'une arborescence appelée graphe de recette (Rgraph). Dans ce cadre, un agent A a un plan P pour réaliser une action b en effectuant une recette R si (cette définition simplifiée ne tient pas compte des relations « rendre possible ») :

1. A croit qu'il peut réaliser chaque action de R ;
2. A veut réaliser chaque action de R ;
3. A croit que réaliser R provoquera la réalisation de b ;
4. A veut exécuter R comme moyen de réaliser b ;
5. A croit que chaque action de R joue un rôle dans le plan P ;
6. A veut que chaque action de R joue un rôle dans le plan P.

Dans le système de Pollack, des règles d'inférence de plan permettent à l'interlocuteur de se représenter le plan du locuteur en terme de ses attitudes mentales. Ainsi, ce modèle permet de traiter les dialogues au cours desquels le plan du locuteur est reconnu comme étant invalide. C'est le cas dans le dialogue suivant :

A : Je dois aller à Paris. Puis-je avoir un billet pour le train 777 ?

B : Le train 777 ne va pas à Paris. Vous feriez mieux de prendre le train 888.

Où B a pu inférer :

1.$Bel(B, Bel(A, Exec(aller\_Paris)))$

2.$Bel(B, Bel(A, Exec(prendre\_train\_777)))$

3.$Bel(B, Bel(A, Generation(prendre\_train\_777, aller\_Paris)))$

4.$Bel(B, Int(A, (aller\_Paris))$

5.$Bel(B, Bel(A, Int(prendre\_train\_777)))$

6.$Bel(B, Bel(A, Int(en(prendre\_train\_777, aller\_Paris)))$

Or, la troisième croyance de A, inférée par B, est erronée.

Plus récemment, les formalismes de plans ont été adaptés de manière à traduire la nature intrinsèquement collective et collaborative de certaines activités. Grosz et Sidner ont ainsi proposé la notion de *plan partagé* [Grosz and Sidner, 1990] en s'appuyant sur les travaux de Pollack. Un plan partagé est une collection d'attitudes mentales faisant intervenir des désirs et des croyances des différents protagonistes (les travaux de ces auteurs, se sont limités à deux interlocuteurs). Suite à de nombreuses critiques, cette notion a été reprise et améliorée par Grosz et Kraus [Grosz and Kraus, 1996] pour obtenir la structure récursive de plan partagé suivante. Un groupe d'agents G a le plan partagé de réaliser a en utilisant la recette R_a si et seulement si :

1. Le groupe G a la croyance mutuelle des actes et des contraintes qui composent la recette R_a ;
2. Pour les actions individuelles, il existe un agent de G tel que tous les agents de G croient mutuellement que :
   (a) cet agent a l'intention de réaliser cet acte ;
   (b) cet agent est capable de réaliser cet acte en utilisant une certaine recette ;
   (c) cet agent a un plan individuel pour réaliser cet acte ;
   (d) le groupe G est engagé sur la réussite de cet agent pour l'action considérée.
3. Pour les actions non-individuelles, il existe un sous-groupe de G tel que les membres de G croient mutuellement que :
   (a) ce sous-groupe a un plan partagé pour réaliser cet acte en utilisant une certaine recette ;
   (b) ce sous-groupe est capable de réaliser cet acte en utilisant cette recette ;
   (c) le groupe G est engagé sur la réussite de ce sous-groupe concernant cette recette.

Cette définition requiert donc que les protagonistes s'accordent sur une recette, une distribution des tâches et l'engagement à la réussite des autres. Notons que cette approche ne nécessite pas d'intention collective, car la structure du plan capture la dimension collective. L'intention collective comme attitude mentale primitive doit, de toute façon, être attribuée à un individu en dernier recours.

Cette représentation permet de justifier les comportements coopératifs au cours de l'exécution du plan. Il reste toutefois à déterminer :

1. comment les individus identifient un besoin de collaborer à propos d'une action ?
2. comment les groupes se forment pour réaliser l'action en question ?
3. comment le plan partagé est élaboré ?
4. comment est-il exécuté ?

Ces quatre étapes sont ce que Wooldridge et Jennings appellent le processus de résolution coopérative de problème [Wooldridge and Jennings, 1994]. Comme le dialogue est une activité dynamique, les agents ne peuvent pas planifier de manière définitive et complète. C'est pourquoi la notion de plan partiel à été introduite.

L'application de la théorie des plans partagés au dialogue à été étudiée par Lochbaum [Lochbaum, 1994]. Selon celle-ci, si les interlocuteurs sont engagés dans un dialogue, ils ont une bonne raison pour cela : c'est soit qu'ils ne peuvent pas réaliser un certain plan seul, soit que leur plan est incomplet en l'état, ... Autrement dit, les interlocuteurs dialoguent pour compléter des plans partiels. Un plan partiel est un plan pour lequel, les agents, soit (1) n'ont qu'une recette ou un graphe de recette partiel ; soit (2) une ou plusieurs actions n'ont été attribuées à aucun agent. Pour Lochbaum, il existe deux relations entre plans (individuels ou partagés, partiels ou pas) :

– *contribution* : un plan contribue à l'établissement de l'une des croyances ou intentions requises pour le second plan. En d'autres termes, la complétion du premier plan contribue à celle du second.

– *pré-satisfaction* : un plan doit être complété avant l'autre. En d'autres termes, la complétion du premier plan conditionne la complétion du second.

Il est ainsi possible d'analyser les sous-dialogues comme l'exécution de plans contributoires à des sous-tâches. En outre, les sous-dialogues de corrections et de clarifications (qui avaient menés Litman et Allen à distinguer plan du discours de plan du domaine) visent à établir des pré-conditions de croyances liées à un plan, essentiellement pour identifier des paramètres ou des recettes à utiliser.

Selon Lochbaum, les plans partagés modélisent donc exactement les DSP de Grosz et Sidner en un seul et même formalisme qui suffit pour prendre en compte de multiples types de dialogues.

4.2.3 Approches de l'interaction rationnelle basées sur la logique

Ces approches, connues sous le nom de *théories de l'interaction rationnelle*, sont une autre application des théories philosophiques présentées ci-dessus. Elles ont, sur les approches par planification, l'avantage d'introduire des sémantiques clairement définies. Elles se basent généralement sur des logiques modales épistémiques (type KD45) pour les attitudes mentales et sur la logique dynamique pour la représentation des actions. Elles trouvent leur principale application dans les sémantiques des ACLs (voir section 4.3).

Cohen et Levesque [Cohen and Levesque, 1990c] ont bâti une théorie de l'interaction rationnelle sur leur théorie de l'action [Cohen and Levesque, 1990a]. Dans leur théorie de l'action, l'intention n'est pas une primitive, mais est exprimée en terme de but persistant. Un agent a le but persistant G relatif à la motivation M ($PGoal(A, G, M)$), si et seulement si :

1. A croit que G n'est pas vrai actuellement ;

2. A veut que G soit réalisé ;

3. cet état de fait restera vrai tant que : (i) G n'est pas vrai, (ii) G est jugé atteignable et (iii) la motivation M de l'atteindre est présente.

L'introduction de la motivation permet d'éviter que tous les buts ne soient irrévocables, i.e. poursuivis fanatiquement. Le but est un engagement individuel qui ne peut être annulé sans raison. Dans ce cadre, un agent a l'intention de réaliser un but G s'il a le but persistant d'avoir réalisé G, et de l'avoir réalisé intentionnellement (c'est à dire non accidentellement).

L'idée principale de leur théorie de l'interaction rationnelle est que les propriétés des actes illocutoires peuvent être dérivées des seuls états mentaux des interlocuteurs. Autrement dit, il n'est plus nécessaire de supposer que le locuteur a l'intention que l'auditeur reconnaisse son intention de réaliser un acte illocutoire particulier, puisque les effets liés à l'acte reconnu seront dérivés. La reconnaissance de l'acte ne sert qu'à inférer l'effet fondamental à partir duquel les autres conclusions seront déduites. Il ne s'agit en aucun cas de la reconnaissance de la force illocutoire au sens de Searle et Vanderveken. Le problème est alors de déterminer quels sont les effets fondamentaux des actes. Techniquement, pour Cohen et Levesque, ces effets sont contextualisés et décrits à l'aide d'axiomes alors que pour Sadek [Sadek, 1991a], ils sont divisés en effets perlocutoires attendus et effets fondamentaux (qui traduisent la préservation des pré-conditions de faisabilité).

Sincérité et bénévolat (la disposition d'un agent à adopter les buts d'autrui si ceux-ci ne contredisent pas les siens) sont deux caractéristiques sur lesquelles repose le comportement coopératif des agents. Ces notions dépassent le cadre du comportement rationnel puisque rien a priori ne contraint un agent rationnel à être un tant soit peu coopératif. Elles sont donc capturées par des hypothèses spécifiques, notamment l'hypothèse de sincérité.

Pour adapter leur théorie au cadre des actions collectives, Cohen et Levesque tentent de définir ce que peut être une intention conjointe [Cohen and Levesque, 1991] dans le cadre de l'interaction rationnelle. La solution retenue est intermédiaire entre la version minimale où chaque agent a la croyance mutuelle que chacun possède l'intention que l'action collective soit réalisée et la croyance mutuelle que chacun ait l'intention de tenir son rôle tant que les autres le font et la version maximale qui exige en plus que cette croyance mutuelle persiste jusqu'à ce qu'il soit mutuellement acquis que l'activité soit réalisée, irréalisable ou plus motivée. La première n'offre aucune garantie que la croyance mutuelle persiste alors que la seconde ne permet pas aux agents de porter un jugement personnel (et privé) sur le statut du but à atteindre. La notion de but affaibli (weak achievement goal) est introduite pour contraindre les agents à avertir les autres d'un éventuel changement de leur attitude vis-à-vis de l'objectif commun.

Un agent a le but affaibli de réaliser p relativement à la motivation M et à un groupe d'agents si :

- soit p est un but « normal » pour lui ;
- soit si p est déjà atteint (ou inatteignable, ou n'est plus motivé), alors l'agent a le but que ce statut soit mutuellement connu de tous les membres du groupe.

La notion de but persistant conjoint est alors introduite et définie de la manière suivante. Un groupe d'agents T a le but persistant G, si et seulement si :

1. il est mutuellement cru par les membres de T que G n'est pas vrai actuellement ;
2. il est mutuellement cru par les membres de T que chacun veut que le but G soit réalisé ;
3. il est vrai et mutuellement cru que tant qu'il n'est pas mutuellement admis que G est déjà atteint, inatteignable ou plus motivé, il sera toujours mutuellement cru que chaque agent de T possède G comme but affaibli.

Comme dans le cas de l'intention individuelle, l'intention conjointe peut alors être définie comme le but persistant conjoint d'avoir réalisé une action, avec la croyance mutuelle que cette action a été réalisée intentionnellement (consciemment et volontairement). On retrouve ici, comme chez Lochbaum, l'idée que la communication est expliquée par son rôle de contrôle dans une activité collective. La communication provient rationnellement de la nécessité de maintenir les croyances associées à l'intention conjointe.

4.3 Sémantique des ACLs

L'annexe A présente les théories des actes de langage et leurs instanciations dans le cadre des systèmes multi-agents : les langages de communications agents, notés ACLs[Agent Communication Language]. Les sémantiques des ACLs sont une application directe des approches intentionnelles logiques discutées ci-dessus (on les appelle encore approches mentalistes). Fournir une sémantique formelle aux actes de langage va permettre d'analyser rigoureusement l'utilisation des communications aussi bien dans les SMAs que dans les sociétés humaines [Chaib-draa and Vanderveken, 1998]. Avant de parler de la sémantique des ACLs, précisons que l'annexe B rappelle comment on envisage la sémantique des langages formels.

4.3.1 Sémantique de FIPA-ACL

Les deux ACLs principaux KQML et FIPA-ACL sont basés sur la théorie des actes de langage : les messages sont considérés comme des actions avec leurs conséquences sur l'environnement. Des sémantiques mentalistes ont été définies pour ces actes. Celle de FIPA-ACL s'exprime en termes de conditions de faisabilité et d'effets rationnels. Un acte est sélectionné lorsque les conditions de faisabilité sont remplies et que l'effet rationnel correspond à une intention du locuteur.

La sémantique de FIPA-ACL est basée sur celle d'ARCOL [Sadek, 1991b]. Elle définit des pré-conditions de faisabilité des messages et des post-conditions décrivant les effets rationnels attendus. Le langage sémantique de FIPA-ACL (SL) est une logique multimodale avec des opérateurs pour les croyances (B), les désirs (D), les croyances incertaines (U), les intentions (but persistant PG). Basée sur les travaux de Cohen et Levesque (voir section 4.2.3) sa forme actuelle est due à Sadek [Finin et al., 1999].

La sémantique actuelle de FIPA-ACL est alignée sur un modèle cognitif et comportemental incluant un certain nombre de primitives de type BDI [Belief, Desire and Intention]. Cette sémantique est donc basée sur des notions mentalistes telles que les croyances et les intentions. On a vu que les communications y sont envisagées comme un type d'action. Cela signifie, en pratique, que la théorie sémantique de FIPA-ACL repose sur une logique multimodale (croyance/intention) combinée à une théorie de l'action. Cela a pour conséquence que les agents qui veulent utiliser FIPA-ACL de manière sémantiquement cohérente doivent être conçus dans un style BDI, ce qui peut constituer une contrainte gênante.

4.3.2 Sémantiques de KQML

Contrairement à la sémantique de FIPA-ACL, la sémantique originelle de KQML ne présuppose pas que les agents soient de type BDI. La sémantique était juste contrainte par le fait que l'agent devait pouvoir manipuler une base de connaissances virtuelle (ajouter/extraire des assertions). Ces faibles contraintes ont laissé une grande liberté aux concepteurs et c'est

	$\texttt{tell}_{\mathrm{A},B}(X)$	proactive-tell
$Pre(A)$	$Bel_{\mathrm{A}}(X) \wedge Know_{\mathrm{A}}(Want_{\mathrm{B}}(Know_{\mathrm{B}}(S)))$	$Bel_{\mathrm{A}}(X)$
$Pre(B)$	$Int_{\mathrm{B}}(Know_{\mathrm{B}}(S))$ avec $S = Bel_{\mathrm{B}}(X)$ ou $S = \neg Bel_{\mathrm{B}}(X)$	
$Post(A)$	$Know_{\mathrm{A}}(Know_{\mathrm{B}}(Bel_{\mathrm{A}}(X)))$	$Know_{\mathrm{A}}(Know_{\mathrm{B}}(Bel_{\mathrm{A}}(X)))$
$Post(B)$	$Know_{\mathrm{B}}(Bel_{\mathrm{A}}(X))$	$Know_{\mathrm{B}}(Bel_{\mathrm{A}}(X))$
Complétude	$Know_{\mathrm{B}}(Bel_{\mathrm{A}}(X))$	$Know_{\mathrm{B}}(Bel_{\mathrm{A}}(X))$

TAB. 5 – Sémantique KQML des performatives `tell` et `proactive-tell`.

pourquoi différentes sémantiques virent le jour [Cohen and Levesque, 1990b]. Labrou et Finin [Finin and Labrou, 1998] ont fourni une sémantique pour chaque performatifs[9] en termes de pré-/post-conditions pour le locuteur et l'interlocuteur avec une condition de complétude en plus.

Les pré-conditions indiquent les états mentaux qu'un agent doit nécessairement posséder pour pouvoir utiliser un performatif et pour que le récepteur puisse l'accepter. Les post-conditions indiquent les états mentaux de l'émetteur après un énoncé réussi d'un performatif et ceux du récepteur après la réception du message. Finalement, la condition de complétude indique les états mentaux qui correspondent à la satisfaction de l'intention qui motive l'échange.

Par exemple (voir tableau 5), la pré-condition de `tell` pour le locuteur (A) impose que A croit ce qu'il dit (X) et qu'il sache que le récepteur (B) veut/désire savoir quoi croire sur X. La pré-condition de `tell` pour B est qu'il doit vouloir savoir quoi croire sur X. La post-condition d'un message `tell` pour B est qu'il peut conclure que A croit X. La condition de complétude permet de s'assurer que le performatif a réussi dans le contexte de communication de A et B. En l'occurrence, il s'agit de la post-condition de B. Certains auteurs ont néanmoins préféré redéfinir une version moins contraignante de ce performatif appelée `proactive-tell` .

La sémantique proposée par Labrou [Labrou, 1996] est basée sur une logique multi-modale sophistiquée et contraint les agents à être conçus selon une architecture de style BDI, ce qui nous ramène au cas FIPA-ACL. On pourrait appliquer la sémantique présentée par Chaib-draa et Vanderveken [Chaib-draa and Vanderveken, 1998] à KQML comme proposé dans leur article, mais il en serait de même (par contre, cette dernière prend en compte le degré d'intensité et certaines facettes des conditions préparatoires qui ont été laissées de côté par Labrou et Finin).

4.4 Limitations des sémantiques mentalistes des ACLs

Il peut sembler que pour les SMAs, les théories formelles de l'énoncé en langage naturel, telle que la théorie des actes de langage, soient pertinentes moyennant quelques extensions. L'idée d'un cadre standard a donné lieu aux ACLs. Et finalement, ce sont les sémantiques des ACLs qui dans l'esprit des approches intentionnelles sont censées assurer le passage de l'énoncé au dialogue. Pour autant, on ne peut pas dire que les agents conversent. Le simple enchaînement d'actes de langage isolés ne donne pas forcément une conversation cohérente. Si comme présenté

[9] Les primitives de KQML sont appelées performatifs même si elles ne sont pas des actes performatifs au sens linguistique du terme. Cette appellation qui prête à confusion a d'ailleurs été critiquée.

en annexe A, il n'y a pas de « consensus » autour de la forme syntaxique que les ACLs peuvent prendre, il n'y en a pas non plus pour ce qui est de leurs sémantiques [Dignum, 2000]. Il n'y a pas de modèles clairs et calculables de la sémantique des actes de langage et moins encore de la sémantique des conversations dans les SMAs.

Comme le souligne Dignum [Dignum, 2000], les problèmes soulevés par les ACLs tels qu'ils existent actuellement sont nombreux. Cela va des problèmes d'ontologie aux problèmes de complétude des ACLs. Sans oublier que les critiques générales des approches intentionnelles logiques s'appliquent ici (elles sont présentées à la section 6.2), on se concentrera ici sur les problèmes concernant spécifiquement la sémantique et la pragmatique des ACLs actuels.

4.4.1 Problèmes concernant la sémantique des ACLs

- *Problème de la minimalité sémantique* : les systèmes de pré/post-conditions habituellement utilisés dans les ACLs permettent de rendre compte du sens minimal des messages. Malheureusement, il est des situations où on a besoin d'un sens plus précis, spécifique au contexte. C'est un problème général de la théorie sémantique des ACLs. D'un côté, on veut que la sémantique soit suffisamment générique pour rendre compte de toutes les situations d'utilisation des ACLs. De l'autre, les systèmes de pré/post-conditions obtenus sont trop généraux et abstraits pour être adéquats à toutes les situations.

- *Problème de conformité sémantique* : les ACLs sont si génériques et explicites que leur pouvoir d'expression est très grand, mais les sémantiques bien définies qu'ils acceptent sont basées sur des logiques tellement puissantes que le calcul de la signification d'un message arbitraire par un agent demande tellement de déduction qu'il est typiquement formellement intractable [Dignum, 2000]. Pourtant, l'usage d'un ACL avec une sémantique complète, facilement extensible, est un énorme atout pour des SMAs hétérogènes et ouverts. Il est très difficile de vérifier si les agents sont dans des états mentaux qui vérifient les pré/post-conditions car en fait, même si le problème est formellement bien-défini, le niveau calculatoire ne suit pas.

- *Problème de l'alignement des sémantiques* : l'alignement de la sémantique d'un ACL sur le modèle cognitif et comportemental de l'agent pose problème quand ce dernier permet l'expression d'actes de communication qui ne sont pas sémantiquement définis dans le ACL. Ce type de problème peut advenir lorsque les contraintes (pourtant nécessaires) de la sémantique du ACL sont trop fortes pour l'agent. Cela arrive par exemple avec les conditions de sincérité. La plupart des sémantiques du ACL (en particulier celles de FIPA-ACL et de KQML) ne permettent pas aux agents d'affirmer quelque chose qu'eux-mêmes ne croient pas. C'est une hypothèse simplificatrice pour les ACLs qui découle de l'analogie avec l'humain, qui ne communique généralement pas en supposant que son interlocuteur ment. Les théories comportementales sophistiquées permettent à l'agent d'agir avec l'intention de tromper (manipulation, mensonge, ...) si cela l'aide à atteindre ses buts[10]. La condition de sincérité le met donc dans l'impossibilité d'exprimer ce qu'il désire exprimer. Notons que d'autres hypothèses simplificatrices (comme la joignabilité sûre qui stipule que les agents reçoivent correctement tous les messages

[10] Cas fréquents en commerce électronique et plus généralement en économie lors de stratégies de maximisation de gain.

qui leur sont envoyés, ...) pour les ACLs peuvent poser des problèmes avec les modèles comportementaux des agents.

– *Problèmes liés à la distribution et à l'autonomie des agents* : les sémantiques des ACLs doivent tenir compte du fait que les agents sont distribués et autonomes (idéalement). Pour un programme ordinaire, les post-conditions peuvent être précisément calculées, car son contexte est accessible et les différentes actions/instructions ne sont pas autonomes les unes des autres (programmation séquentielle). Cette distinction pose pour les ACLs une barrière entre l'effet espéré d'un acte de langage et son effet réel. Quand un agent A transmet l'information X à B, il a l'intention que B va au moins croire que A croit X. Mais comme les agents sont autonomes, un agent ne peut jamais changer directement les croyances d'un autre et l'effet d'un acte de langage n'est jamais garanti. Autrement dit, les agents n'ont pas le contrôle des effets perlocutoires de leurs actes de langage et de leurs communications en général.

Les communications agents souffrent donc de l'absence d'une sémantique concise, tractable et universellement acceptée. Comme l'indiquent Koné et ses collègues [Kone et al., 2000], cela confine la communication agent à des environnements restreints et cela rend difficile (voir impossible dans certains cas) la communication entre des agents hétérogènes, c'est-à-dire développés par des concepteurs différents.

4.4.2 Problèmes concernant la pragmatique

– *Problème de prise en compte du contexte social* : l'interférence du contexte social sur l'utilisation des ACLs. Le contexte social contraint les actions (y compris les actes de langage) que les agents peuvent entreprendre par le biais d'obligations, de normes et d'engagements de toutes sortes. Ce type de contraintes n'est pas pris en compte par les théories sémantiques actuelles des ACLs.

– *Problème d'expression de la pragmatique* : le comportement communicationnel d'un agent est le résultat implicite ou explicite de sa planification. Par exemple, une question est posée en attente d'une réponse et cette réponse est censée participer à la progression de l'agent vers ses buts ou l'achèvement d'une tâche. De même, un service est demandé dans le but que l'autre agisse en conséquence. Idéalement, cela devrait faire partie des pré-conditions de ce type d'actes de langage. Ces pré-conditions de type pragmatique sont pourtant très difficiles à exprimer avec les théories sémantiques actuelles des ACLs.

4.5 Application des approches intentionnelles

Les modèles issus des approches intentionnelles sont complexes à implanter. Bien qu'envisagés spécifiquement pour les agents artificiels via les sémantiques d'ACLs développés, ce sont les approches intentionnelles basées sur les plans qui ont été appliquées aux IHM [Interfaces Homme-Machine], dans un cadre mono-agent.

Par exemple, le modèle développé par Balkanski, Hurault-Plantet et leurs collègues utilise la reconnaissance et l'élaboration de plans pour comprendre et accomplir des actes communicatifs dans le contexte orienté-tâche bien délimité d'un répondeur téléphonique qui redirige les usagers vers le poste téléphonique souhaité [Balkanski and Hurault-Plantet, 2000]. Il existe des systèmes comparables également basés sur les plans comme le système TRAINS [Allen et al., 1995] dû à Allen et son équipe ou encore le raisonneur de plans de domaine de Ferguson [Ferguson, 1995]. D'autres sont plus axés sur le raisonnement et utilisent les plans via les approches logiques comme le Circuit Fixit System de Duke [Smith et al., 1995] et le système Artimis développé par l'équipe de France Télécom [Sadek et al., 1997]. Tous sont des IHM.

5 Approches conventionnelles

5.1 Fondements philosophiques

5.1.1 Convention, normes et règles

Une *convention* est une régularité qui existe au sein d'une communauté (sans pour autant avoir nécessairement fait l'objet d'accords explicites). Les *normes sociales* sont généralement définies comme des conventions accompagnées des notions de sanction, de jugement des autres, d'obligation dont le maintien ne peut pas être réduit à des considérations d'ordre rationnel. Nous ne distinguerons pas les deux notions. Il est habituel d'exprimer normes et conventions sous forme de règles. Pour ce qui est des conventions dans le dialogue, Allwood [Allwood, 1994] constate trois types de régularités :

– celles qui dépendent des relations au sein d'un énoncé ;

– celles qui dépendent des relations entre énoncés ;

– celles qui dépendent des relations entre les facteurs globaux et les énoncés.

C'est sur le second type de régularité dans le cours d'un dialogue que les approches conventionnelles se concentrent. L'idée de ces approches est donc que la cohérence conversationnelle peut être garantie (sans hypothèses de coopération ni d'intentionnalité) par les seules conventions. Le problème n'est pas de savoir comment ces régularités sont établies, on suppose les interlocuteurs prêts à s'y conformer. L'idée est alors de limiter les formes possibles d'expression de manière à ce que l'intention véhiculée soit non ambiguë et qu'il ne soit plus nécessaire d'effectuer une analyse pragmatique complexe. Parmi ces modèles, on trouve les protocoles de communications (voir section 5.2).

5.1.2 Le tableau de conversation : fond commun

Dans les approches conventionnelles et par opposition aux approches intentionnelles, l'emphase est mise sur l'aspect public de la conversation. Introduit pour rendre compte de la construction collective et publique des conversations, le *tableau de conversation* est un enregistrement (éventuellement structuré) de l'état de la conversation. Il fait partie du fond commun encore appelé arrière-plan conversationnel (voir section 3.3). Les débats pour savoir ce que doit contenir le tableau de conversation ne sont pas clos et on ne sait s'il doit se limiter à l'histoire de la conversation ou bien être plus général et capable de s'adapter de manière à rendre corrects les énoncés. L'implantation de celui-ci est aussi problématique. La question de savoir si ce tableau doit être centralisé ou bien si chaque agent doit en posséder sa propre version reste en suspend. Dans le second cas, on parle d'*agenda*.

5.1.3 Les engagements

Le comportement des agents cognitifs conventionnels est basé sur l'état interne privé, c'est-à-dire les attitudes mentales de l'agent. Or, les agents sont habituellement définis en termes d'autonomie, d'inter-opérabilité et de leur capacité à atteindre ensemble des buts communs. Le point de vue mentaliste, centré sur l'agent-individu, n'aide pas à définir une sémantique pour tout ce qui est commun et partagé dans un groupe d'agents. Dans ce cadre, la communication agent souffre d'un manque de sémantique formelle concise et universellement acceptée (voir 4.4.1), de sorte que la communication agent est restreinte à des domaines précis dans des environnements qui ne sont pas ouverts. Un certain nombre de solutions à ce problème ont été récemment proposées dans lesquelles le caractère social des agents est mis de l'avant [Singh, 1998, Colombetti, 1998, Moulin, 1997]. La notion d'*engagement* a alors rapidement émergé comme permettant de capturer un niveau public du dialogue.

Les engagements jouent un rôle central dans le dialogue, ce sont des entités complexes qui alimentent le tableau de conversation. Dans la littérature, on recense les caractéristiques suivantes :

- *Les engagements sont sociaux* : ce sont des engagements vis-à-vis des autres membres d'une communauté (différent de l'engagement individuel utilisé par Cohen et Levesque pour traduire la persistance liée à l'intention). Les engagements pris (implicitement) contraignent la poursuite du dialogue. On distingue la personne qui engage de la personne engagée (pas forcément le locuteur). Par exemple, une assertion engage le locuteur alors qu'un acte directif vise à engager l'interlocuteur.

- *Les engagements sont propositionnels ou en action* : il s'agit en fait de distinguer deux types d'objets sur lesquels peuvent porter les engagements. L'engagement propositionnel porte sur la validité d'une proposition tandis que l'engagement en action concerne l'exécution d'une action.

- *Les engagements peuvent être conditionnels* : un engagement conditionnel est un engagement qui tient si certaines conditions sont remplies. Les engagements conditionnels sont à rapprocher de l'idée de règle, comme par exemple dans : si tu me dis z, je te dirais p.

- *Les engagements sont explicites ou implicites* : les engagements explicites résultent d'actes (de langage ici) alors que les engagements implicites correspondent à des habitudes d'interaction.

- *Les engagements peuvent être établis à différents niveaux d'ancrage* : (1) au niveau linguistique, ce sont les faits qui ont été explicitement validés de manière linguistique, (2) au niveau par défaut on trouve des faits inférés mais révisables et (3) au niveau hypothèse où les faits sont simplement supposés.

- *Les engagements sont ordonnés* : certains engagements en action doivent être satisfaits avant d'autres. C'est un moyen classique pour structurer le tableau de conversation (ou les agendas) qui les contient que de les ordonner suivant la priorité des actions que de tels engagements manipulent.

5.1.4 Engagements communs sur des projets communs

Pour Searle [Searle, 1992b], toute conversation est l'expression d'une intention conjointe (*We-intention*). Ce type d'intention transcende la conjonction des états intentionnels individuels. Cette intentionnalité partagée est un facteur d'explication important pour tous les comportements sociaux. Searle considère cette intention collective comme un élément de base qui ne se réduit pas à la somme des intentions individuelles (*I-intention*) et de leurs connaissances mutuelles. Néanmoins, des intentions individuelles sont incluses dans la *We-intention*. Dans ce cadre, les conversations sont des formes de *We-intention*. Par exemple, la *We-intention* « nous parlons du prix de X » inclut la *I-intention* (pourtant très différente) « j'offre 5$ pour X » et la *I-intention* « je refuse ton offre ».

Clark, lui, considère l'usage social du langage comme un *type d'activité commune*. Il utilise le concept de *type d'activité* de Levinson [Levinson, 1979], qui est une notion plus générale que la notion de type de discours, car elle inclut des événements sociaux autres que le discours ou bien dans lesquels le discours n'a qu'une place incidentale. Clark introduit la notion de projet conjoint. Un *projet conjoint* est une action commune proposée/suggérée par un des participants et acceptée/réalisée par tous. Un *projet conjoint* pourrait être un plan, une recette ou une procédure pour accomplir une activité ensemble [Chaib-draa and Vongkasem, 2000]. Pour avoir des chances d'aboutir, un tel projet nécessite l'engagement de tous les participants. Ainsi, on peut dire que les *We-intentions* se manifestent par des *engagements conjoints* sur des *projets conjoints*.

Lors d'une conversation libre, l'*intention conjointe* doit émerger des interlocuteurs et de la conversation elle-même. Cette émergence est partie intégrante de la conversation et est l'objet d'évaluations et de négociations continues entre les interlocuteurs. Dans le cas des conversations orientées/contraintes, cette intention conjointe pré-existe souvent. En effet, les conversations contraintes ont un propos, un sujet ou un but auquel se raccrocher. On peut voir les protocoles de communication utilisés avec les ACLs comme l'expression d'autant de projets conjoints. Chaque participant doit s'engager dans le projet conjoint, après quoi il s'agit d'une intention conjointe.

Quand les agents prennent part à une *activité commune*, à l'exécution d'un plan commun ou à un *projet conjoint*, ils accomplissent des actions conjointes. Beaucoup de ces actions conjointes (ou de leurs parties) sont des actes de communication nécessaires au bon accomplissement de l'action (se faire comprendre, tester la compréhension des autres, divulguer une information, ...). Cela nous amène à considérer les actes de langage sous une nouvelle perspective. En effet, on ne peut plus étudier les actes de langage sous la seule perspective du locuteur. Il nous faut étudier *pour le locuteur comme pour l'interlocuteur* : le rôle de l'acte dans la conversation, les liens de l'acte aux autres actes de la conversation. Il convient d'étudier aussi, la façon dont le locuteur génère l'acte dans le contexte et la manière dont l'interlocuteur peut comprendre ce que veut lui dire le locuteur et qui peut relever de ces liens plus que de l'acte lui-même.

5.2 Protocoles de communication

L'idée de protocole est une réification de l'idée de projet conjoint conventionnel. Dans le champ de la communication dans les systèmes artificiels, les protocoles de communication pallient à un problème majeur des ACLs hérité de la théorie des actes de langage. En effet, la théorie des actes de langage est une théorie de l'énoncé « isolé » ; or, la communication donne lieu à des discours ou à des conversations qui sont des ensembles d'énoncés inter-dépendants.

Chaque agent doit posséder une procédure de décision qui lui permet de choisir puis de générer des actes de langage en fonction de ses propres intentions. Il ne s'agit pas simplement de trouver l'acte de l'ACL dont la sémantique s'unifie avec les intentions de l'agent. Pour être pertinent, l'agent doit au minimum prendre en compte le contexte de son acte de langage (les événements en cours et passés, y compris les actes de langages précédents). Cela pose le problème du lien entre la définition de la sémantique d'une primitive d'ACL et la conversation à laquelle elle participe. D'un côté, il semble clair que le sens général d'une conversation (les engagements/promesses/informations qui y circulent) ne peut se passer du sens des performatifs qui la constituent. De l'autre, certains chercheurs pensent que la sémantique de la conversation elle-même doit être vue comme une primitive. Dans cette voie, les éléments de base de la sémantique des conversations doivent être sociaux plutôt qu'individuels pour être compatibles avec les théories ou leurs ébauches présentées ci-dessus (voir 3.2). Les activités, actions et intentions communes sont alors les éléments de base de ces sémantiques au même titre que les croyances, désirs et intentions individuelles.

Actuellement, dans les SMAs concrètement développés, la prise en compte du contexte se fait de manière simplifiée par l'utilisation de conversations pré-planifiées, stéréotypées. Cela permet de réduire considérablement l'espace de recherche pour la réponse à donner tout en restant consistant avec la sémantique. Du fait de cet avantage computationnel, quasiment tous les SMAs utilisant un ACL sont dotés d'une couche « conversation », qu'elle soit explicite ou implicite. La spécification de ces conversations se fait à l'aide de stratégies (conversation policies) de conversation qui sont en fait de petits protocoles. Les protocoles (ou stratégies de conversations) spécifient des séquences d'actes communicatifs sans rien préciser du contenu des actes. Ils ont donc l'avantage de simplifier le calcul des réponses possibles à un message donné.

5.2.1 Règles de conversation de KQML et FIPA-ACL

KQML est muni d'une seule règle de conversation. Elle est simple, même si de nombreuses variantes sont permises. La conversation commence lorsque qu'un agent envoie un message KQML à un autre et se termine lorsque ce dernier répond. Les variantes sont obtenues par l'utilisation de performatifs de régulation de conversation. Comme leur nom l'indique, ces performatifs permettent aux agents d'intervenir dans le cours normal d'une conversation. Elles permettent notamment d'enrichir une conversation en prolongeant la règle par défaut (`standby`, `next`, `rest` ou `discard`) ou au contraire de mettre prématurément fin à une conversation (`eos`, `error` ou `sorry`). La thèse de Labrou [Labrou, 1996], présente un certain nombre de mini-conversations ainsi que leurs sémantiques et les contraintes qui lient les différents performatifs les composant. Ces mini-conversations sont conçues pour servir d'éléments de base pour des échanges plus importants.

FIPA-ACL, contrairement à KQML, fournit un certain nombre de protocoles de communication impliquant chacun plusieurs actes de communication (notés CA pour Communicative Act). Par exemple : demande d'action, réseau de contrats (contract net), différents types d'enchères, ...

5.2.2 Problèmes concernant les protocoles

Malgré leur apparente simplicité, l'utilisation des protocoles soulève un certain nombre de questions. Même si les conversations peuvent être structurées par enchaînement de protocoles, ce type d'approche semble trop rigide à la majorité des chercheurs.

En outre, il existe plusieurs candidats non-équivalents pour la spécification de protocoles : les réseaux de Pétri, les réseaux de transitions à états finis, les arbres de sous-buts, les graphes de Dooley ainsi que d'autres approches fondées sur la logique. Ces formalismes offrent des degrés de souplesse des conversations envisageables très variables.

Finalement, les protocoles sont extrêmement contraignants et les messages non attendus dans le protocole ne seront pas examinés. Reste aussi à savoir comment des agents se mettent d'accord sur l'utilisation d'un protocole. Puis, comment les implémenter dans les systèmes multi-agents ? Doivent-ils faire partie de l'axiomatique de la communication ? Comment apprend on de nouveaux protocoles ? Comment intégrer les protocoles au sein du fonctionnement de l'agent ?

C'est pourquoi de nombreuses recherches récentes considèrent une approche alternative aux protocoles centrée sur le niveau social. En effet, si les protocoles sont limités, les agents doivent néanmoins suivre des conventions sociales qui rendent possible la conversation. Ces approches sont généralement inspirées de la dialectique formelle.

5.3 Les systèmes dialectiques

De son étude des arguments formellement fallacieux, Hamblin [Hamblin, 1970] déduit que certains arguments sont inadéquats sans être formellement non valides, c'est pourquoi la logique formelle n'est pas adaptée pour rendre compte de l'argumentation. C'est la naissance de la logique informelle (ou logique dialectique). Un système dialectique est un système normatif de régulation du dialogue[11] constitué de :

1. *un ensemble de coups* : idéalement, les différents coups ou locutions définies couvrent les différents types d'actes de langage. On trouve par exemple des coups pour l'assertion, la question, le défi, ...

2. *une liste d'engagements par participant* : Hamblin suppose l'existence d'une liste d'engagements qui permet notamment de gérer la cohérence du dialogue. Pour autant, il faut qu'un joueur soit capable d'anticiper/détecter les inconsistances dans les listes d'engagements. Afin de ne pas doter les agents de capacités irréalistes (dans le genre du problème de l'omniscience) Hamblin isole un ensemble de schémas d'axiomes inconsistants.

3. *un ensemble de règles pour les coups* : les règles du système interdisent de jouer certains coups. Ces règles sont de la forme : Si *conditions* alors les actes linguistique $X_1, ..., X_n$ sont interdits. Les *conditions* pouvant faire référence au passé du dialogue.

4. *un ensemble de règles pour les engagements* : ces règles indiquent quel est l'impact des coups sur les listes d'engagements

Notons que les systèmes dialectiques ne traitent pas de la production des arguments, il ne font pas le lien entre les états mentaux de l'agent et son activité dialogique. La sous-section suivante présente brièvement un exemple de système dialectique.

[11] Initialement, Hamblin ne traitait que les dialogues de persuasion.

5.3.1 Le système DC

Ce célèbre système dû à MacKenzie [MacKenzie, 1979] modélise les interactions de type argumentatives entre deux opposants qui soutiennent des thèses contradictoires. Plus précisément, le système DC a pour but de prévenir les esquives de question comme dans le petit dialogue suivant :

1.A : Je l'ai peint en rouge
2.B : pourquoi ?
3.A : parce que je ne l'ai pas peint en bleu
4.B : pourquoi ?
5.A : parce que je l'ai peint en rouge

Dans DC, cinq types de coups sont possibles : affirmation, retrait, question, défi et résolution. Deux types de règles sont définis : les règles de mise à jour et les règles de dialogue. Ces coups peuvent changer les engagements du locuteur mais aussi ceux de l'interlocuteur. C'est une gestion de l'acceptation implicite. Les interlocuteurs respectent alors le principe de coopération suivant : « Exprimer un désaccord de croyance dès que possible ».

5.3.2 Mouvement/glissement dialectique

En dialectique, il est admis que chaque type de dialogue doit disposer de son système dialectique [Walton and Krabbe, 1995]. Or, l'analyse de conversation montre bien qu'un dialogue est rarement d'un seul et même type du début à la fin. Il est courant d'imbriquer des types de dialogues. On appelle ces changements de type de dialogue (et donc de contexte) au cours de celui-ci des *glissements dialectiques*. Pour que ces glissements soient constructifs, il faut qu'ils soient acceptés par tous les participants.

En conclusion de cette section, ajoutons que la dialectique formelle a permis le développement de nouveaux types de stratégies de conversations : les protocoles basés sur les engagements et ceux basés sur les jeux de dialogue. Les sections suivantes présentent ces nouvelles stratégies de conversation envisagées spécifiquement pour les SMAs.

5.4 Approches des protocoles basées sur les engagements

Ces approches ont en commun de reposer sur la notion d'engagement (présentée en section 5.1.3) pour capturer la dimension publique de la communication.

5.4.1 Le modèle d'agent social de Singh

Singh est le premier à identifier le besoin d'un modèle sémantique formel des ACLs en terme de « notions sociales » [Singh, 1998]. Il propose sa propre sémantique sociale pour ACL [Singh, 2000] comme une partie de sa théorie des agents sociaux. Singh introduit une sémantique basée sur l'engagement social intégrée à une logique modale temporelle avec temps ramifié (CTL[Computational Tree Logic]). Il définit trois niveaux de sémantique, qui correspondent à trois assertions valides pour chaque acte de langage :

- *assertion objective* : la communication est vraie, c'est-à-dire que le locuteur s'engage sur son acte de langage. Par exemple, si le locuteur informe le groupe que p, alors il s'engage envers le groupe sur cette croyance.

- *assertion subjective* : la communication est sincère. Par exemple, si le locuteur informe le groupe que p, alors il s'engage envers le groupe sur sa sincérité (supposée).
- *assertion pratique* : la communication est justifiée. Par exemple, si le locuteur informe le groupe que p, alors il doit avoir des raisons de penser que p est vrai.

Singh encapsule donc l'approche mentaliste dans le niveau social. Il y a bien une différence entre le fait que le locuteur soit sincère (hypothèse de sincérité dans les approches intentionnelles) et dire qu'il est socialement engagé comme étant sincère. Par contre, cette sémantique sociale ne remplace en rien les protocoles classiques, car si elle permet de connaître *a posteriori* le résultat d'une composition de communication, elle n'indique pas qu'elle est la bonne composition à utiliser.

5.4.2 Approche par la dialectique formelle de Amgoud et al.

Amgoud et al. [Amgoud et al., 2000] ont définis des règles de dialogue et des règles de mise à jour des connaissances pour les différentes locutions de leur système. Dans ces travaux, directement inspirés par les travaux de MacKenzie sur les systèmes dialectiques (voir section 5.3), les règles de dialogue indiquent les séquences de locutions autorisées (comme pour les protocoles) tandis que les règles de mise à jour capturent les effets de ceux-ci sur l'état du dialogue (représenté par les tableaux d'engagement des participants). Les locutions définies étendent celles de MacKenzie de sorte à couvrir tous les types de dialogues proposés par Walton et Krabble (présentés en section 3.4).

5.4.3 Approche par la logique modale de Colombetti

Colombetti explore l'idée d'états mentaux sociaux à l'aide de la logique modale. Il distingue différents mécanismes par lesquels les agents pourraient acquérir des croyances communes : déduction, information montrée, observation mutuelle, communication intentionnelle, ...

Colombetti a récemment proposé le langage ALBATROSS [Agent Language Based on the Treatment of Social Semantics] [Colombetti, 2000] qui - dans la même veine que les travaux de Singh - donne une sémantique (exprimée dans la logique temporelle $CTL*$) aux actes illocutoires courants. Une des particularités de ce langage est la notion de pré-engagement, qui est un type d'engagement conditionnel. Par exemple, une requête pré-engage l'interlocuteur à qui elle s'adresse, signifiant que si celui-ci accepte, il sera engagé à agir en conséquence.

5.4.4 Approche de Flores et Kremer

Récemment, Flores et Kremer ont aussi proposé une approche qui vise à définir les protocoles et la sémantique des actes de langage dans un modèle unifié basé sur la notion d'engagement social [Flores and Kremer, 2001]. Comme dans les travaux d'Hamblin, les agents maintiennent un tableau des engagements partagés qui peuvent être ajoutés ou retranchés. Un unique protocole pour la proposition définit comment ces engagements sociaux peuvent être négociés. Dans le système de Flores et Kremer, tous les actes de dialogue (proposer, accepter, rejeter, contrer) sont produits conjointement via ce protocole, ce qui est une hypothèse forte, contraignante et discutable. Un agent peut par exemple, proposer d'ajouter un engagement social. Si sa proposition est acceptée, tous les agents conversant adoptent cet engagement.

TAB. 6 – Sémantique des actes de dialogue dans le système de Flores et Kremer.

Énoncé	**But illocutoire**	**Opération**	**Créditeur**	**Débiteur**
Demande	Propose	Ajout	locuteur	interlocuteur
Offre	Propose	Ajout	interlocuteur	locuteur
Retire	Propose	Retrait	locuteur	interlocuteur
Annule	Propose	Retrait	interlocuteur	locuteur

Ainsi, les actes de dialogue comme la proposition définissent des opérations génériques sur les tableaux d'engagement, qui peuvent s'instancier différemment, comme indiqué par le tableau 6.

Par exemple, une demande est une proposition d'adopter un engagement social fourni par le locuteur, alors qu'une offre est une proposition par le locuteur d'adopter un des engagements sociaux de l'interlocuteur.

5.5 Approches des protocoles basées sur les jeux de dialogue

Les jeux de dialogue formels sont des jeux dans lesquels les coups sont des locutions régis par des règles. L'idée de jeux de dialogue remonte à Aristote, elle fut successivement exploitée par Wittgenstein (dans un sens différent de celui donné ici) puis par Hamblin, père de la dialectique formelle [Hamblin, 1970]. L'approche par les jeux de dialogue passe par la définition d'un système dialectique qui comprend :

- un ensemble de sujets de discussion ;
- la syntaxe d'un ensemble de locutions concernant ces sujets ;
- un ensemble de règles qui gouvernent l'énonciation de ces énoncés ;
- un ensemble de règles qui établissent quels engagements doivent être créés ou modifiés par les différents énoncés ;
- un ensemble de règles qui indiquent dans quelles circonstances le dialogue se termine.

Les désiratas pour ces systèmes dialectiques [Wooldridge et al., 2002], lorsqu'ils sont appliqués aux SMAs sont :

- *Énonciation des objectifs du dialogue* : un système dialectique doit avoir un ou plusieurs objectifs que les locutions et règles définies doivent faciliter ;
- *Diversité des objectifs individuels* : le système dialectique doit permettre à chacun d'atteindre ses objectifs consistants avec ceux du dialogue. Ces objectifs peuvent être conflictuels (négociation) ou coïncider (recherche d'informations) ;
- *Inclusif* : le système dialectique ne doit pas empêcher un agent de participer au dialogue s'il est qualifié et qu'il en a la volonté ;

– *Transparence* : les participants d'un dialogue doivent avoir une connaissance *a priori* des règles et de la structuration du système. En particulier, toutes les références du système dialectique vers la réalité extérieure doivent être explicites (principalement les engagements en action) ;

– *Justice* : le système doit traiter tous les participants de manière équivalente ou en cas d'asymétries (dues aux rôles), elles devront être explicites ;

– *Clarté de la théorie argumentative* : le système dialectique devra être basé sur une théorie de l'argumentation partagée, de sorte que les obligations dialectiques, les règles d'inférence, soient connues et acceptées. Ainsi, les agents peuvent raisonnablement anticiper le comportement des autres (dans une fourchette définie par le jeu). Par exemple, un agent qui conteste une assertion doit savoir s'il va recevoir une défense de l'assertion ou pas, ...

– *Séparation claire de la syntaxe et de la sémantique.*

Historiquement, l'idée de jeux de dialogue s'est concrétisée de différentes manières : scripts partagés de Levin et Moore [Levin and Moore,], recettes partagées de Mann [Mann, 1988], réseaux de transition de Lewin [Lewin, 2000] et méta-règles de conversation de Airenti et al [Airenti et al., 1993]. On présente ici les approches les plus récentes proposées dans le cadre de la communication agents.

5.5.1 Approche de Reed

Reed [Reed, 1998] a proposé la notion de cadre de dialogue comme structure d'échange abstraite. Ses travaux sont directement basés sur ceux des chercheurs en dialectique Walton et Krabbe [Walton and Krabbe, 1995].

A. Structure

Un cadre de dialogue est formellement défini comme un quadruplet :

$F =<< t, \Delta >\in D$, $\tau \in \Delta$, $\{u_{x\rightarrow y}{}^0, ..., u_{x\rightarrow y}{}^n\} >$, où :

$$D = \{< persuade, B >, < negociate, C >, < inquire, B >, < infoseek, B >, < deliber, P >\}$$

D est l'ensemble des types de dialogue et de leurs notions associées (B pour les croyances, C pour les contrats et P pour les plans). τ est le sujet du dialogue et les u^n sont les énoncés qui peuvent être produits au $n$ième tour du dialogue. Les u^n définissent donc un protocole pour ce cadre de dialogues entre les agents x et y. Ce protocole peut être vide.

B. Acceptation

Les structures proposées par Reed sont manipulées par deux méta-actes de communication : *propose/accept*. Ces énoncés ont pour but d'ouvrir les cadres de dialogues. Par exemple, l'énoncé suivant est bien formé et indique que x propose à y de négocier le prix de *Zidane*, avec 500 comme proposition initiale :

$u_{x\rightarrow y}{}^0$:$<< propose(negociate, < buy(y, Zidane), \{< price, 500 >\}), \{0\} >$

Les actes *concede* et *accept* ferment automatiquement le cadre.

C. Composition

Reed considère deux types de compositions : séquentielle et imbriquée, toutes deux capturées d'emblée par la structure définie ci-dessus. En effet, comme les propositions sont des coups comme les autres, elles peuvent être faites en cours de dialogue. Quand un nouveau cadre de dialogue est proposé au tour i et accepté au tour $i + 1$, si aucun cadre n'est ouvert, c'est un séquencement sinon c'est une imbrication. Dans ce dernier cas, le cadre courant est simplement suspendu jusqu'à ce que le nouveau cadre soit terminé. C'est au concepteur de s'assurer que ses cadres termineront tous et de réguler les possibilités d'imbrication (qui doivent rester en nombre fini).

5.5.2 Approche de Dastani et al.

Dastani et ses collègues ont proposé une méthodologie pour la construction de protocoles de négociation flexibles [Dastani et al., 2000]. Bien qu'ils négocient, les agents partagent le but commun de coordonner leurs actions. Une représentation partielle des actions coordonnées sous la forme de recette est donnée et les jeux de dialogue en sont un type particulier. Les auteurs mettent également de l'avant la notion de *cohérence*. Un dialogue est cohérent dans son contexte si : (1) il correspond à un plan qui peut permettre d'atteindre le but apparent d'un agent (2) il suit les règles d'interaction courantes. Dépendamment des attitudes des agents et de celui qui a l'initiative, (1) ou (2) prend le dessus. Même si le travail effectué concerne la négociation, le cadre proposé est censé être suffisamment générique pour permettre d'autres types de dialogues orientés tâche.

Les énoncés sont formés d'actes de dialogue qui sont composés d'un contenu sémantique et d'une fonction communicative. Un enregistrement conversationnel garde trace des actes de dialogue et des engagements associés qui circulent (représentation du contexte). La fonction communicationnelle est liée à la tâche sous-jacente et/ou à la régulation de l'interaction. La cohérence tient donc autant à la cohérence au niveau de la tâche qu'à celle du niveau interactionnel. Dans leur système, la cohérence orientée tâche est assurée par l'inférence de plans alors que celle du niveau interactionnel repose sur des recettes pré-planifiées pour l'action communicative commune telles que présentées par Hulstijn [Hulstiln, 2000].

A. Structure

Un acte de dialogue est soit initiatif, soit réactif et la structure de base des jeux développés est une unité initiation/réaction. Cependant, les échanges sont régulés par des conditions de cohérence sur le contenu sémantique.

B. Composition

Les jeux peuvent être composés de manière statique (au moment de leur conception) ou dynamique par séquencement ou chaînage.

C. Acceptation

Bien que l'utilité d'une phase de négociation du jeu courant soit mentionnée, aucune indication n'est fournie sur ce point.

5.5.3 Approche de McBurney et Parsons

McBurney et Parson [McBurney and Parson, 2000] ont récemment proposé une autre approche utilisant explicitement des structures de jeux ayant pour ambition de représenter les types de dialogues proposés par Walton et Krabbe (voir section 3.4) ainsi que certains méta-dialogues. Pour ce faire, ils proposent un modèle en trois couches : (1) une couche de sujet (topique) qui définit quels sont les sujets possibles du dialogue, (2) une couche de dialogue et (3) une couche de contrôle.

A. Structure

Les jeux de dialogue sont définis dans la couche de dialogue et consistent en un système dialectique traditionnel composé de : (1) règles d'ouverture, (2) règles de locution, (3) règles de dialogue, (4) règles de mise à jour et (5) règles de terminaison.

B. Acceptation

C'est la couche de contrôle qui assure l'acceptation via un méta-dialogue de contrôle auquel on suppose que les agents sont prêts à participer. Ce méta-dialogue de contrôle permet aux agents de décider conjointement des jeux de dialogue à jouer. Pour ce faire, des coups de méta-niveau sont définis ($begin(G(p)), end(G(p))$) et un coup spécial ($propose.return.control$) permet aux agents de remonter au niveau contrôle alors qu'ils jouent un jeu de dialogue, assurant ainsi la liaison entre les deux couches.

C. Composition

C'est aussi la couche de contrôle qui permet la composition des jeux. Différents types de composition sont permis :

– *itération* G^n : répétition de n dialogues du type G, chacun des jeux débutant après la fermeture du précédent ;

– *séquence* $G; H$: H débute après la fermeture de G ;

– *emboîtement* $G[H : I]$: H débute pendant G après la séquence de coups I ;

– *parallélisation* $G \bigcup H$: G et H débutent simultanément ;

– *test* $< p >$: $< p >$ est un dialogue de contrôle pour tester le statut de vérité de p. Le dialogue courant termine si p s'avère faux.

On note que la parallélisation est un mode de composition original, qui n'est pas présent dans la littérature classique, mais peut s'avérer utile pour la communication agent. En outre, et contrairement à la version de Reed, l'emboîtement ne suspend pas le jeu emboîtant.

5.5.4 Approche de Maudet et Chaib-draa

Le modèle de Maudet adapté aux systèmes multi-agents avec l'aide de Chaib-draa, Bourget, Labrie et Pasquier DIAGAL[DIAlogue Game based Agent communication Language] se base sur la distinction entre jeux de dialogue et jeux de communication. Un des jeux de communication, le jeu de contextualisation, joue le rôle de la couche de contrôle du modèle de McBurney et Parson (présenté en section 5.5.3) [Maudet and Chaib-draa, 2001].

A. Structure

Les jeux sont des structures bilatérales, définies par leurs conditions d'entrée (qui doivent être remplies pour qu'un jeu puisse être ouvert), leurs conditions de sortie (qui définissent le but des conversants en entrant dans le jeu) et des règles de dialogue. Toutes ces notions sont définies en termes d'engagements (éventuellement conditionnels), ce qui fait du modèle une approche par les engagements stricte à comparer à celle de Flores et Kremer (présentée dans la sous-section 5.4.4).

B. Composition

Deux types de compositions étaient initialement considérés :

– *le pré-séquencement* : il y a pré-séquencement lorsqu'un jeu est ouvert alors qu'un autre est proposé. Le pré-séquencement sert essentiellement à établir les conditions d'entrée dans le jeu pré-séquencé ;

– *l'emboîtement* : un jeu $j1$ est emboîté dans un autre $j2$ s'il est ouvert alors que $j2$ est déjà ouvert. Ainsi, les emboîtements sont gérés via une pile de jeux ouverts. Les engagements du jeu courant sont prioritaires sur les engagements des autres jeux.

C. Acceptation

L'acceptation est prise en charge via le jeu de contextualisation qui définit les coups en termes d'engagements : la proposition d'entrer dans le jeu, ($prop.in(x, j)$, propose le jeu et engage le partenaire à accepter, refuser ou proposer un pré-séquencement), celle d'en sortir (qui doit être partagée ; ce qui est une hypothèse forte et discutable). À chaque coup dans un jeu de dialogue, correspond un coup de continuation dans le jeu de contextualisation et un jeu peut être imbriqué à tout moment par simple proposition dans le jeu de contextualisation.

5.5.5 Conclusion sur les jeux de dialogue

Même s'il reste du travail de formalisation, d'implémentation et de validation, les approches par jeux de dialogue semblent être une avenue prometteuse pour pallier la rigidité des protocoles traditionnels. En effet, ces approches adoptent un formalisme plus flexible que les automates à états finis : l'utilisation des engagements garde une trace plus riche du dialogue que le simple dernier coup considéré dans les protocoles traditionnels. Ces derniers contraignent les agents à se conformer aux transitions attendues alors que les engagements motivent les agents à se conformer à un comportement attendu. Les approches par engagements sont donc capables de considérer les messages inattendus ou exceptionnels mieux que les protocoles classiques.

En outre, et même si de nombreuses clarifications seraient nécessaires de ce côté, les jeux de dialogue peuvent être composés. Sous l'impulsion de Reed, des méta-actes (jeux de contextualisation, couche de contrôle, ...) ont été définis pour négocier la stratégie de conversation (le jeu de dialogue) courante.

6 Limites et comparaisons

Cette section présente un bref bilan des avantages et des inconvénients des deux grandes familles d'approches étudiées jusqu'ici. La discussion est menée dans l'optique de l'utilisation de ces théories pour les SMAs. On présente ensemble les aspects théoriques et techniques, car les limites théoriques se retrouveront comme des limites de la puissance des modèles de communication pour les SMAs qui en seront issus. La section 2.1 justifie que des considérations issues des théories du langage naturel interviennent dans les théories de l'interaction en IAD et leurs applications aux SMAs. Ce type de réflexion intégrant différents domaines des sciences cognitives nous semble seul permettre une vision globale et une réflexion féconde.

6.1 Avantages des approches intentionnelles

Le principal avantage des approches intentionnelles est leur complétude (informelle). En effet, les approches intentionnelles couvrent trois des composantes majeures du traitement de la communication : la syntaxe est couverte par les ACL qui découlent de la théorie des actes de langage ; la sémantique des ACL est bien définie en termes des attitudes mentales des agents et la pragmatique est gérée par un système de planification collective ou par une théorie de l'interaction rationnelle (complétée par une forte hypothèse de coopération). En conséquence, un des intérêts majeur des approches basées sur la planification est qu'elles ont été appliquées et donc implémentées avec succès dans le domaine des IHM (voir section 4.5). Cependant, l'utilisation de la planification implique certaines limitations (voir section 6.2). Les approches logiques qui dépassent ces limitations, sont quant à elles trop lourdes et complexes (en terme de complexité computationnelle) pour les agents et pour leurs concepteurs de sorte qu'on ne leur connaît pas d'applications ne passant pas par la planification.

6.2 Limites des approches intentionnelles

L'approche intentionnelle logique propose une théorie complète de la communication. Cependant, sa version théorique générale est trop complexe pour être implantée telle quelle. Les modèles mentaux impliqués sont généralement exprimés dans des logiques multi-modales dont l'implémentation est encore un sujet de recherche. Les simplifications consenties pour arriver à une implantation font perdre une partie de la puissance de ces modèles, notamment leur sémantique. C'est ainsi que les approches intentionnelles basées sur la planification peuvent être vues comme une simplification des approches logiques dans le cadre computationnel plus facilement implémentable de la planification. Ce faisant, les domaines d'application ont été limités aux domaines orientés tâche, laissant de côté les autres types de dialogues. Les approches intentionnelles basées sur les plans ont été appliquées et donc implémentées avec succès. Cependant, comme le souligne Cohen [Cohen, 1996], il reste des limites théoriques inhérentes aux modèles basés sur les plans, en particulier :

- *la dépendance au domaine* : le modèle est basé sur la bonne définition des recettes qui doivent prévoir toutes les possibilités, c'est à dire couvrir tout le domaine. Cela limite d'autant la portée de ce type de modèle.

- *le manque de bases théoriques* : même si ces modèles sont computationnels et ont permis de nombreuses avancées dans la compréhension des conversations, il leur manque des fondements théoriques. En ce sens, ce sont des modèles à court terme qui sont efficaces du fait de leur nature procédurale, mais ils ne constituent pas une théorie du dialogue.

À cela, viennent s'ajouter des limitations communes à toutes les approches intentionnelles :

- *la complexité des inférences* : les algorithmes de reconnaissance de plan sur lesquels reposent les approches intentionnelles par planification sont combinatoirement intractables dans le pire des cas, et indécidables dans certains cas [Bylander, 1991]. Pour ce qui est des approches logiques, la sémantique des actes communicationnels est tellement riche qu'il est trop complexe de déterminer les réponses possibles en inférant les états mentaux des autres agents [Maudet and Chaib-draa, 2001].

- *la sémantique des messages* : Comme le rappellent Maudet et Chaib-draa, dans les approches intentionnelles logiques, la sémantique des messages est formulée en terme d'états mentaux, aspects privés aux agents [Maudet and Chaib-draa, 2001]. Cela pose le *problème de la vérification sémantique* : pour que la sémantique des messages soit vérifiable, il faudrait avoir accès aux états mentaux privés des agents ce qui n'est généralement pas possible[12]. Le second problème majeur posé par cette formulation est *le problème de l'hypothèse de sincérité* nécessaire à la définition d'une telle sémantique. Cette hypothèse est jugée trop contraignante par la communauté SMA [Dignum, 2000]. Elle interdit notamment d'envisager correctement certains types de dialogues dans des domaines où une telle hypothèse ne saurait tenir, comme c'est le cas par exemple pour les dialogues de négociation dans le commerce électronique.

- *l'hypothèse de coopération* : les approches intentionnelles reposent toutes sur une hypothèse de coopération. Cette hypothèse a d'abord été formulée comme adoption de but (voir section 4.2.1) puis de manière moins contraignante comme une participation attendue à l'action conjointe (dans les théories de l'interaction rationnelle, voir section 4.2.3). L'idée selon laquelle la communication émerge rationnellement de l'action collective sous-jacente est fondamentale dans les approches intentionnelles. Si cette idée est élégante, elle reste néanmoins liée à la considération d'une activité collaborative. En conséquence, il n'est pas possible d'expliquer par ces approches, ce qui motive la communication lors de situations non-collaboratives [Traum, 1994].

Finalement, le rôle central de l'intention dans toutes ces approches peut aussi être questionné du point de vue de son universalité et ces approches sont parfois qualifiées d'ethnocentriques [Nuyts, 1994]. En effet, l'analyse de conversation montre bien qu'une grande partie des dialogues comprennent des phases ritualisées (ouverture, fermeture, remerciements, ...)

[12]La vérification sémantique dont il s'agit ici ne doit pas être confondue avec la vérification formelle que les agents implémentent correctement une sémantique mathématique particulière. La vérification dont il est question ici est plutôt la vérification que peut entreprendre un agent sur un autre pour s'assurer qu'il agit de manière cohérente vis-à-vis des dialogues tenus.

qui sont sans rapport avec la reconnaissance d'intention et qui ne sont pas prises en compte par ces approches. On peut donc se demander si ces approches sont pleinement satisfaisantes pour la modélisation du dialogue et plus généralement : les attitudes mentales considérées sont-elles pertinentes ? Les notions classiques de croyances, désirs et intentions ne traduisent nullement l'idée d'engagement social ou d'obligation pourtant cruciale pour considérer l'action collective [Traum, 1994]. Les approches intentionnelles n'indiquent pas comment caractériser cette intuition de « liant collectif » qui explique que les interlocuteurs peuvent généralement compter sur certaines actions les uns des autres.

6.3 Avantages des approches conventionnelles

Parmi les avantages des approches conventionnelles sur lesquels il est bon d'insister, car ils sont pour une bonne part de l'intérêt porté à ces approches par la communauté SMA, il faut relever que les approches conventionnelles résolvent le problème de l'hypothèse de sincérité et le problème de vérification sémantique qui étaient attachés aux approches intentionnelles. Le problème de l'hypothèse de sincérité est résolu puisque les engagements ne sont pas nécessairement sincères. Par contre, un engagement doit être tenu et s'il ne l'est pas son créditeur s'expose à des sanctions sociales ou autres punitions. En cela les engagements sont des attitudes sociales indépendantes mais pas indifférentes des spécificités internes aux agents. Quant au problème de la vérification, il est simplifié grâce au caractère public des engagements qui les rend accessibles à tous, en particulier pour vérification. En outre, les approches conventionnelles parviennent, grâce à la notion d'engagement, à éviter toute spécification mentaliste dans la sémantique des langages de communication utilisés. Le traitement des aspects sociaux du dialogue en est simplifié.

Aspects positifs spécifiques aux jeux de dialogue

L'approche par les jeux de dialogue est déclarative (les règles sont explicitées) : cela accroît leur clarté et rend possible leur définition dans un langage explicite comme XML (comme c'est le cas par exemple dans le simulateur de jeux de dialogue développé avec Chaib-draa et ses collègues [Chaib-draa et al., 2002]).

Comme le signale Maudet [Maudet, 2001], les jeux sont utiles au niveau du dialogue aussi bien dans les phases d'interprétation que dans les phases de production. De plus, ils sont une contribution intéressante pour ce qui est de la structuration du dialogue. La structure locale est donnée par la structure intra-jeu (ou par les obligations langagières correspondantes). La structuration globale est donnée par la structure inter-jeux. Mais, le niveau jeu est-il réellement nécessaire ? Qu'est-ce que la notion de jeu apporte de plus qu'un modèle intentionnel augmenté d'obligations ? En guise de réponse à ces questions, Maudet indique que :

1. Les jeux sont empiriquement fondés, comme le montrent les expériences menées sur les annotations de dialogues [Kowtko et al., 1991] ;

2. Les jeux sont des structures prédictives : Poesio [Poesio and Mikheev, 1998] montre que sur le corpus de $MAPTASK$ les prédictions sont de 50% avec les jeux contre 38% si l'on considère uniquement le coup (l'énoncé) précédent ;

3. Les jeux raffinent et concrétisent la notion de coopération dialogique. Ils réalisent concrètement les notions d'acceptation (grounding) et de projet conjoint discutées en sec-

tions 3.2, 3.3 et 5.1.4.

6.4 Limites des approches conventionnelles

6.4.1 Limitations des approches par engagement

On peut s'étonner du peu de concepts communs aux approches intentionnelles et aux approches conventionnelles. En fait, en introduisant une couche publique via les engagements, normes et conventions (qui se déclinent en obligations, permissions, interdictions), les approches conventionnelles séparent bien le niveau public du niveau privé, mais au prix d'une perte de complétude. En effet, l'introduction d'un niveau public ne dispense pas de la définition d'un niveau privé et il reste alors à articuler ces deux niveaux. Il s'agit donc de définir un système de gestion du niveau public cohérent sur lequel les agents savent raisonner. Comme c'est là notre problématique, nous reviendrons sur ce problème plus longuement à la section 7.

6.4.2 Limitations des approches par les jeux de dialogue

L'approche par les jeux de dialogue nous semble, de par sa simplicité, la plus prometteuse des approches conventionnelles. Cependant, ses développements multi-agents sont récents et donc encore incomplets. Parmi les éléments qui restent à définir pour en faire un cadre interactionnel complet, on note les besoins suivants :

- Clarifier les notions d'engagements, notamment en différenciant les engagements ayant trait à différents niveaux de dialogue. Il y a les engagements véhiculés par les jeux de dialogue eux-mêmes pour la structuration de celui-ci et ceux se rapportant au sujet du dialogue (extra-linguistique).
- Régler les problèmes d'implémentation : implantation distribuée du gestionnaire de dialogue et des agendas (dans un cadre multi-agent) ;
- Étendre les cadres existants (dialogique) aux conversations (plus de deux interlocuteurs) ;
- Clarifier les garanties de terminaison des dialogues (nombre d'imbrications possibles, ...) ;
- Optimiser les jeux de dialogue, par exemple en identifiant les chemins inutiles qui peuvent être éliminés sans préjudice sémantique (en terme d'engagement) ;

Il reste, en outre, un certain nombre de points non expliqués par les approches conventionnelles en général. En effet, ces approches se concentrent sur les aspects externes - sociaux - de la communication sans définir comment ces derniers seront pris en charge par les agents cognitifs tels qu'ils sont définis actuellement. C'est-à-dire que rien n'est dit sur la manière dont les agents devraient utiliser ces langages et structures afin de gérer leurs engagements d'une manière utile à leurs objectifs individuels et collectifs. Maudet [Maudet, 2001] indique simplement que les agents devront être normatifs et délibératifs. Normatifs pour suivre les règles de dialogue, se conformer aux jeux de dialogue et agir en fonction de leurs engagements et délibératifs pour prendre en compte leurs propres besoins/intentions. Reste donc à savoir comment ces différents niveaux - privé et public - peuvent être combinés. C'est d'ailleurs la problématique que l'on se propose d'attaquer dans notre travail de thèse et que l'on présente plus en détail à la section 7.

6.5 Conclusion

Il y a au moins quatre aspects pertinents dans l'étude de la communication inter-agent qui correspondent à quatre facettes générales de l'étude du langage : syntaxe, structure, sémantique et pragmatique. Pour ce qui est de la syntaxe des énoncés, les ACLs fournissent un bon outil puisqu'ils ont la puissance expressive suffisante pour rendre compte de tous les énoncés. Ce résultat découle du fait qu'ils sont l'application de la théorie des actes de langage qui couvre l'ensemble des énoncés en langage naturel[13]. On renvoie à l'annexe A pour une présentation détaillée de ces aspects syntaxiques.

Au niveau structurel, une alternative semble émerger entre les approches strictement cognitives comme les approches intentionnelles ou les approches strictement sociales comme celles basées sur les engagements qui ne spécifient rien de la structure (vue comme émergeant des enchaînements des énoncés des agents) et les approches par protocoles qui réduisent l'espace de recherche des continuations possibles au strict minimum, mais font perdre souplesse et adaptativité aux conversations. Cette alternative est proposée par les jeux de dialogue.

Pour ce qui est de la sémantique des unités conversationnelles, la communauté scientifique spécifiquement multi-agent s'est déplacée d'une sémantique mentaliste vers une sémantique sociale exprimée en termes d'engagements permettant de résoudre le problème de la vérifiabilité, lever l'hypothèse de sincérité et faciliter le traitement des aspects sociaux de la communication. Il reste cependant de nombreux débats pour savoir à quel niveau on place la sémantique (au niveau de l'énoncé ou au niveau de la conversation) et déterminer quelle forme peut prendre cette sémantique[14].

Concernant la pragmatique, ces déplacements n'invalident sans doute pas complètement la pragmatique Gricéenne et les considérations sur la coopération restent valables, mais l'introduction de la couche publique des engagements mérite que celle-ci soit ré-examinée. Dans les approches conventionnelles, les concepts de base fournissant la pragmatique des approches intentionnelles sont complétés d'une couche sociale constituée d'engagements. L'introduction de ce niveau social et public nécessite de repenser une pragmatique étendue à ce cadre plus général.

En effet, l'agent ne doit plus raisonner directement sur les intentions (communicationnelles ou pas) des autres, mais sur les engagements pris et à prendre. Ces engagements sont autant les engagements posés par l'agent lui-même, les engagements posés par les autres agents, les engagements que le cadre interactionnel l'ont contraint à accepter que les engagements négociés, imposés ou concédés lors de conversations avec d'autres agents. C'est pourquoi la proposition d'une nouvelle théorie de l'utilisation du cadre interactionnel nous semble être un objectif intéressant pour re-visiter la pragmatique des conversations dans le cadre des SMAs à base d'agents cognitifs et rationnels. La section 10 présente l'approche que nous proposons dans cette optique. En attendant, voyons dans la section suivante, comment notre problématique se dégage de cet état de l'art.

[13]En réalité, ce résultat n'est que potentiel, car les ACLs développés pour l'heure ne couvrent pas tous les types d'actes de langage définis dans la théorie.

[14]L'annexe B présente une typologie sommaire des sémantiques mathématiques envisagées pour les ACLs et souligne l'ambiguïté qui réside entre sémantique linguistique et sémantique mathématique.

Deuxième partie

Problématique

7 Introduction

Puisque les approches intentionnelles ne sont pas complètement satisfaisantes et que les chercheurs de la communauté SMA (et ils ne sont pas les seuls) s'accordent à penser que le niveau des engagements publics est une nécessité, on souhaite contribuer à compléter les approches conventionnelles, notamment en indiquant comment les cadres interactionnels qui y sont définis peuvent être utilisés par les agents. Il est acquis que le cadre interactionnel est en définitive un ensemble de contraintes visant à simplifier la structuration des conversations grâce à une couche conventionnelle qui contraint les enchaînements conversationnels. Il est également acquis que le modèle cognitif d'un agent (ou d'un groupe d'agents) fournit et produit un ensemble de contraintes qui devront être satisfaites au mieux pour assurer la satisfaction de l'agent (ou du groupe). Reste donc à établir les liens entre le modèle cognitif et le cadre interactionnel, c'est-à-dire à définir comment l'agent va utiliser le cadre interactionnel (en en respectant les contraintes) de façon à satisfaire ses contraintes cognitives (c'est-à-dire à maximiser sa satisfaction).

Les cadres interactionnels proposés par les approches conventionnelles ne fournissent pas les éléments nécessaires à leur utilisation automatique par des agents cognitifs (comme nous l'avons vu à la section 6.4). De la même manière, ils ne fournissent aucune garantie quant à l'utilité des conversations tenues et tel n'est pas leur objectif. Pourtant, ce n'est pas l'habileté des agents à structurer leurs dialogues qui nous intéresse mais leur habileté à communiquer de manière utile à leurs objectifs individuels et collectifs. On pourrait donner une première formulation de notre problématique sous la forme des deux questions suivantes : de quelle manière pourrait-on automatiser la communication agent (au sens de l'utilisation automatique du cadre interactionnel) ? Comment un agent pourrait-il procéder pour déterminer si une conversation lui a été profitable ou pas et agir en conséquence ?

C'est à ces questions que l'on souhaite essayer de donner une réponse. Il y a là un certain nombre d'enjeux dont le principal pour l'informatique concerne la conception des SMAs : pour l'heure, le comportement communicationnel n'est pas automatisé. C'est au concepteur de définir quels dialogues seront tenus dans telle ou telle circonstance. Bien souvent, seul le concepteur connaît les raisons de ces dialogues et lui seul est garant de leur utilité. C'est également au concepteur de déterminer quelles seront les poursuites du dialogue, si poursuite il y a. Certains concepteurs ont évidemment imaginé certains algorithmes ad hoc pour automatiser certains enchaînements conversationnels, mais il n'y a pas de théorie générale de la communication agent par conversation (autre que les approches intentionnelles, dont on a vu les limitations à la section 6.2). Ainsi, du point de vue informatique, notre problématique est de chercher à fournir des éléments théoriques et pratiques pour permettre aux agents de :

- manier automatiquement le cadre interactionnel conventionnel pour répondre à leurs besoins ;
- juger par eux-mêmes de l'utilité de leurs actions et en particulier des conversations tenues ;
- décider de manière automatique des suites à donner à une conversation.

Nous pensons que ces éléments pourraient être utiles à la conception des SMAs et donc susceptibles d'alléger le travail du concepteur. La réponse aux deux questions évoquées dans le paragraphe précédent sera d'autant plus satisfaisante qu'elle sera générique et que ce soit bien un mécanisme général qui soit exhibé et non une simple réponse ad hoc. Pour parvenir à cette généricité, nous devrons prendre garde de fonder notre approche sur des bases théoriques solides et ainsi contribuer, en retour de cet apport à l'informatique, aux sciences cognitives en général et à la pragmatique en particulier.

Commençons par reformuler cette problématique en introduisant une nouvelle distinction dans les communications agents. Nous indiquerons ensuite nos objectifs.

8 Cohérence structurale et cohérence cognitive

Notre problématique peut être reformulée autour de la notion de cohérence. Dans les théories de la communication, on distingue les théories cognitives des théories interactionnelles [Littlejohn, 1992]. Les théories interactionnelles traitent de la forme de la communication : comment modéliser un énoncé, un discours, une conversation ? Quelles sont les régularités structurelles des conversations ou quelles sont les contraintes conventionnelles qui pèsent sur la forme du dialogue et sur les enchaînements permis (enchaînements d'énoncés, d'actes de langage) ? Il existe de nombreuses théories interactionnelles : analyse de conversations, théories des actes de langage, ...

Les théories cognitives s'intéressent, quant à elles, à la production des messages (quoi dire, quand le dire et à qui le dire) ainsi qu'à la réception et au traitement cognitif des messages (quoi comprendre, comment le comprendre et comment réagir). Elles adressent l'aspect fonctionnel de la communication aux niveaux interne et externe. Quels sont les éléments qui poussent un agent à former tel énoncé plutôt que tel autre ? Comment un agent réagit-il à un énoncé au niveau interne ainsi qu'en terme de mise à jour du modèle d'autrui et de ses propres croyances ? Au niveau externe et public (c'est-à-dire, vis-à-vis de son environnement), quels sont les engagements que l'agent veut obtenir ? Pourquoi ? Quelle est l'utilité de la conversation, quelle est son importance ? L'agent et plus généralement le groupe d'agents conversant est-il satisfait par la conversation ? Il y a bien une différence entre la satisfaction des conventions qui pèsent sur le dialogue (par exemple, satisfaire un jeu de dialogue) et la satisfaction des agents.

Si dans les deux approches la cohérence est une notion centrale, il faut se garder de confondre la cohérence structurale du dialogue (souvent appelée cohérence conversationnelle [Craig, 1983]) - est-il permis de poursuivre le dialogue de cette façon ? - de sa cohérence de fond encore appelée cohérence cognitive. Le contenu du message est-il approprié à la vue des messages précédents et des états mentaux de l'agent ? Est-ce que le contenu du message est cohérent avec l'état interne de l'agent ? Est-ce que les agents ont tenu des propos pertinents quant à leurs objectifs ? Est-ce que la conversation leur est profitable ? Évidemment, ces deux dimensions de la cohérence sont souvent liées. Et, travailler sur une théorie cognitive comme le prescrit notre problématique ne signifie pas nier le besoin d'une théorie interactionnelle. En effet, lorsque l'on a déterminé quoi dire, quand le dire et à qui le dire, reste à savoir comment le dire. Par contre, cela permet de dépasser ce niveau et les idées avancées dans les sections suivantes sont valables pour tout cadre de communication conventionnel suffisamment riche au niveau interactionnel.

Pour ce qui est des communications, la communauté SMA se concentre depuis quelques années sur l'élaboration d'un cadre interactionnel standard. Les principaux langages de communication agent actuels, KQML [Finin and Fritzson, 1994b] et FIPA-ACL [FIPA, 2002], sont basés sur la théorie des actes de langage [Searle, 1969] augmentée d'une sémantique mentaliste [Finin and Labrou, 1998]. Le dialogue est censé émerger de l'enchaînement des productions d'actes issus des intentions de chaque agent via la reconnaissance et le raisonnement sur les intentions des autres. Cette approche, dite « mentaliste », a été critiquée [Moulin, 1997, Singh, 1998] et une redéfinition de la sémantique en termes plus sociaux ainsi que la construction d'une surcouche conversationnelle ont été rendues nécessaires à différents égards. Les protocoles se sont fait reprocher leur manque de souplesse [Dignum, 2000] et les polices de conversation puis plus récemment les jeux de dialogue sont proposés pour pallier les défauts de ceux-ci [Chaib-draa and Dignum, 2002, Maudet, 2001]. Dans ces derniers, l'emphase est mise sur la cohérence structurelle.

À ce niveau, il est important de noter qu'une sémantique du cadre interactionnel, quelle qu'en soit la forme, ne garantit pas l'utilité des conversations tenues dans ce cadre. Pour nous, comme pour d'autres : philosophes du langage (Searle, Witgenstein [Wittgenstein, 1953]), épistémologues [Barreau, 1995] ou encore chercheurs en sciences cognitives [Vignaux, 1991], la possibilité d'aboutir à un modèle complet qui soit purement structurel/interactionnel est hypothétique, et ce du fait même de l'existence et de la prédominance de la cohérence cognitive du dialogue, c'est-à-dire des aspects cognitifs, psychologiques et des conditions de satisfaction qui s'y rapportent. Dans de nombreux cas, c'est cet aspect sémantique, fonctionnel et pragmatique qui domine. L'important est ce qui est dit, pas comment cela est dit ni même dans quel ordre. Examinons un exercice classiquement réalisé en cours de linguistique pour s'en convaincre : à la suite d'une intervention radio d'une dizaine de minutes, force est de constater qu'aucun des élèves n'est capable de se souvenir ou de reproduire les cinq premiers énoncés de la conversation, ni de savoir s'ils constituaient un enchaînement structurellement cohérent alors que tous savent ce qui s'est dit à ce moment. La cohérence et la pertinence du dialogue sont jugées au niveau des idées qui transitent dans le dialogue, de leur impact cognitif[15], pas au niveau de sa forme qui n'en est que le médium. Une étude approfondie serait nécessaire pour étudier les rapports entre ces deux types de cohérence. Prenons deux exemples pour bien les différencier :

Exemple 1

1.A *: Est-ce que je peux te poser une question ?*

2.B *: Oui, vas-y.*

3.A *: Est-ce que tu as l'heure ?*

4.B *: Non, j'aimerais bien la connaître...*

L'exemple précédent montre une conversation dans laquelle tous les aspects interactionnels conventionnels sont remplis. Le dialogue est bien formé et les engagements pris par les agents sont respectés (notamment l'engagement de répondre à une question pris par B en 2.B). Pourtant, cette conversation a en définitive une assez faible utilité si elle s'arrête à ce moment-là. En effet, bien que les intervenants finissent par partager leur problème (ils désirent connaître l'heure, mais ne disposent pas de l'information), ce problème n'est pas résolu. De manière un peu simpliste, on pourrait dire que les agents communiquent pour résoudre ou éviter des

[15]Les approches intentionnelles rendent compte de cette dimension cognitive, mais sans offrir la souplesse des approches conventionnelles en terme de structuration des dialogues.

problèmes (pas de problème, pas de communication, l'inverse étant faux). Dans ce cadre, les communications ne sont que des tentatives pour résoudre ou éviter ces problèmes et comme telles, elles peuvent échouer. C'est ce qui se passe dans cet exemple. Cette vision des choses permet de bien comprendre pourquoi les cadres interactionnels actuels (protocoles, jeux de dialogue, ...) ne garantissent pas l'utilité des conversations et qu'un mécanisme est requis pour la mesurer et permettre aux agents d'agir en conséquence.

Exemple 2

1.*A : Est-ce que tu peux me prêter ta règle ?*

2.*B : Si on ne part pas tout de suite, nous serons en retard.*

À l'inverse, dans l'exemple précédent, la cohérence structurale n'est pas respectée, le dialogue est mal formé du point de vue de la dialectique. Les règles conventionnelles de gestion du sujet dans les paires d'adjacences ne sont pas respectées, mais l'échange est utile, car A et B vont peut-être pouvoir éviter un problème (on suppose que A et B tiennent à être ponctuels). En effet, la cohérence de fond est indéniable dès lors que B fait allusion à une urgence et se permet donc de violer les règles structurelles les plus élémentaires pour expliciter un problème commun qu'il juge plus important que le problème pour lequel A réclame son aide. C'est un type de « court-circuit » utilitariste et opportuniste très courant dans l'utilisation du langage naturel.

Ainsi, il est probable qu'une théorie purement structurale des conversations énoncerait des règles qui, si elles sont souvent respectées, sont également très largement enfreintes pour des raisons de cohérence cognitive. Comme le montre l'exemple 2 : pour augmenter l'utilité du dialogue, un agent peut être amené à ne pas respecter les règles interactionnelles standards. Ces infractions sont comparables aux violations des maximes de Grice [Grice, 1957] (ici, les maximes de qualité et de manière sont violées). En effet, comme elles, les théories interactionnelles définissent généralement un cadre normatif qui peut (et parfois doit) être enfreint. Cependant, pour des agents rationnels, ces infractions (si elles sont permises) tout comme les dialogues bien formés doivent supporter une caractéristique plus générale d'utilité. Notre problématique inclut donc naturellement la question de la définition d'une métrique pour juger de cette utilité.

9 Objectifs

Niveau d'analyse du langage appelé de ses voeux par Morris en 1938, la pragmatique rassemble tous les éléments de sens du langage. Elle prend en compte la théorie de la référence qui caractérise la sémantique, mais aussi une théorie de l'usage et du contexte que la langue anime pour faire sens. *Notre objectif principal est de fournir une telle théorie de l'usage qui soit adaptée aux cadres interactionnels agent que proposent les approches conventionnelles et en particulier les jeux de dialogue.*

Il s'agit donc de fournir des outils de méta-niveau pour guider l'agent dans son comportement communicationnel. On devra donc définir des outils de gestion de la dynamique du dialogue pour guider l'agent dans son comportement communicationnel, autant au niveau de l'utilisation du cadre interactionnel (quel type de dialogue choisir ?, quelle structuration choisir ?. . .) qu'au niveau cognitif (avec qui communiquer, quand et quoi ? pourquoi ?). Ce faisant, on souhaite contribuer à établir un lien entre les aspects privés et les aspects publics dans les communications agents. En effet, on a vu à la section 6 que pour l'heure, les approches conventionnelles échouent à :

- indiquer comment et pourquoi le dialogue est initié ?
- établir les liens entre les aspects privés (états mentaux) et la couche publique (les engagements). Il s'agit essentiellement de définir un système de gestion des engagements et une théorie de l'utilisation du cadre interactionnel par les agents ;
- indiquer comment gérer la structuration dynamique du dialogue et ce aussi bien au niveau (1) local (comment un agent choisit dynamiquement parmi les différentes structurations possibles dans une conversation ?) que (2) global (comment un agent choisit son prochain interlocuteur et le type de dialogue dans lequel il souhaite s'engager et pourquoi ?) ;

De manière plus synthétique, on peut reprendre les différents aspects informatiques de notre problématique en indiquant les objectifs qui en découlent :

- *l'utilité dans les communications agent* : les objectifs sont de la définir, de fournir des moyens de la calculer et de montrer en quoi cette notion peut être utile pour la structuration des communications aussi bien au niveau du choix des locutions/coups dans un dialogue qu'au niveau du choix d'un interlocuteur et d'un type de dialogue (capturés par des jeux de dialogue ou des protocoles) sur le moyen terme ;
- *l'automatisation des communications agents* : puisqu'il s'agit pour un agent de maximiser la satisfaction de ses contraintes cognitives, tout en respectant les contraintes normatives imposées par les conventions ayant cours dans le système multi-agent, ainsi que par le cadre interactionnel utilisé. Un autre objectif est de fournir des éléments d'automatisation des communications agents via l'utilisation des outils de satisfaction de contraintes. Cet aspect devra découler de l'étude entreprise sur l'utilité dans les communications et être en accord avec cette dernière.

Ces objectifs découlent de notre problématique, elle-même dérivée de l'état de l'art présenté dans la partie précédente. La prochaine partie présente notre contribution, notamment quant à l'atteinte de ces objectifs.

Troisième partie

Solution proposée

10 Théorie de la cohérence cognitive

10.1 Définitions préliminaires : intentionnalité, cognitions et attitudes

En sciences cognitives (et donc en intelligence artificielle), l'intentionnalité est la propriété qu'ont les états mentaux de représenter des états de choses du monde :

– réalisées : c'est le cas des croyances, dont le contenu fixe l'état du monde représenté ;

– à réaliser : c'est le cas des désirs, buts et aspirations dont le contenu fixe l'état[16] que le monde devrait atteindre au sens de l'agent.

Les cas paradigmatiques d'états intentionnels sont les attitudes propositionnelles (notées AP dans la suite) : « On regroupe sous le nom d'attitudes propositionnelles les croyances, buts, désirs, intentions, obligations, craintes, espoirs, souhaits, attentes, etc., qui ont en commun d'être identifiées par leur contenu propositionnel. » [Houdé et al., 1998].

Les cognitions regroupent tous les éléments cognitifs : les perceptions, les attitudes propositionnelles (composante cognitive), les émotions (composante affective), les comportements (internalisation des actions, volitions). De l'ensemble des cognitions résultent les attitudes qui sont des dispositions psychologiques positives ou négatives en rapport à un objet concret, abstrait ou à un comportement. Toutes les théories des attitudes, également appelées théories de la cohérence cognitive, font appel au concept d'*homéostasie*, c'est-à-dire à la faculté qu'ont les êtres vivants de maintenir ou de rétablir certaines constantes physiologiques ou psychologiques qu'elles que soient les variations du milieu extérieur. Ces théories partagent en prémisse le *principe de cohérence* qui pose la cohérence comme mécanisme organisateur premier :

L'individu est plus satisfait avec la cohérence qu'avec l'incohérence.

L'individu forme un système ouvert dont le but est de maintenir la cohérence autant que possible (on parle aussi de balance ou d'équilibre). Les changements d'attitude résultent de ce principe dans les cas d'incohérence.

Nous avons retenu la théorie de la dissonance cognitive pour fonder nos travaux, autant pour son aspect complet et général que pour sa proéminence historique et l'abondance des travaux et des formalisations qu'elle a engendrés[17].

[16] Notons que les états représentés ne sont pas nécessairement des états existants ni même des états possibles du monde. Par exemple, on peut désirer rencontrer le père Noël et on peut croire que $4 + 3 = 9$.

[17] Avant de faire ce choix, on a étudié de nombreuses théories du changement d'attitude, parmi lesquelles : la théorie du renforcement [Miller and Dollar, 1941], la théorie du traitement de l'information [Hovland et al., 1953], la théorie du jugement social [Sherif and Hovland, 1961], la théorie de la balance [Newcomb, 1953], la théorie de la congruence [Osgood, 1963], et la théorie de la consistance cognitive [McGuire, 1960].

10.2 Généralités sur la théorie de la dissonance cognitive

La théorie de la dissonance cognitive, initialement présentée en 1957 par Festinger est l'une des plus importantes théories de psychologie sociale [Festinger, 1957]. Elle a généré des centaines d'études et d'extrapolations sur les attitudes, les comportements et les croyances humaines, l'internalisation (sic) des valeurs, les motivations et les conséquences des prises de décisions, les désaccords inter-personnels, la persuasion et autres phénomènes psychologiques importants [Harmon-Jones and Mills, 1999]. Ceci s'explique en partie par la formulation très générale et abstraite de cette théorie qui la rend facile à manipuler. Dans les théories de la communication [Littlejohn, 1992], elle apparaît comme l'une des principales théories de réception et de traitement des messages.

Dans sa version originale, cette théorie considère que deux éléments de cognition (perceptions, attitudes propositionnelles ou comportements) sont en rapport ou pas (un lien pertinent les relie ou non). Deux cognitions liées sont soit consonantes (ou cohérentes) soit dissonantes (incohérentes). Elles sont dites consonantes si l'une entraîne ou supporte l'autre. A l'inverse, deux cognitions sont dites dissonantes si l'une entraîne ou supporte le contraire de l'autre.

L'hypothèse de base de cette théorie est que la dissonance produit chez le sujet une *tension* qui l'incite au changement. L'existence d'une dissonance plonge le sujet dans un état inconfortable de sorte que cela le *motive* à réduire cette dissonance. Plus la dissonance est intense, plus ce « malaise » psychologique est fort et plus la pression pour réduire la dissonance l'est aussi. Une dissonance peut être réduite en : (1) supprimant ou réduisant l'importance des cognitions dissonantes, (2) ajoutant ou augmentant l'importance des cognitions consonantes.

La seconde hypothèse de Festinger est qu'en cas de dissonance, l'individu ne va pas seulement changer ses cognitions ou essayer de changer celles des autres pour essayer de la réduire, il va aussi éviter toutes les situations qui risquent de l'accroître. Ces deux hypothèses ont été vérifiées par de nombreuses expériences de psychologie [Wickland and Brehm, 1976].

Un des intérêts majeurs de la théorie de la dissonance cognitive est de fournir une mesure de la dissonance, c'est-à-dire une métrique de la cohérence cognitive. Initialement, Festinger définissait l'intensité (la magnitude) de la dissonance introduite par une cognition X avec une mesure du taux de dissonance[18] définie de manière informelle comme suit :

$$dissonance_X = \frac{produit\ des\ importances\ des\ cognitions\ dissonantes\ avec\ X}{produit\ des\ importances\ de\ toutes\ les\ cognitions\ en\ rapport\ avec\ X}$$

La dissonance globale d'un agent peut, elle aussi, être calculée à partir des dissonances introduites par ses différentes cognitions. Toutes ces mesures ont été raffinées par la suite, donnant lieu à de nombreuses formalisations de la dissonance cognitive [Shultz and Lepper, 1999, Sakai, 2001]. On peut se demander dans quelles circonstances la dissonance survient. En fait, il y a différentes situations dans lesquelles la dissonance est presque inévitable :

1. *contact direct initial avec une situation* : une situation entièrement nouvelle est susceptible d'introduire un certain nombre de nouveaux éléments de cognition dissonants avec ceux qui pré-existent ;

2. *un changement dans la situation* : de la même façon, un changement dans la situation peut amener des éléments de cognition jusqu'alors consonants à devenir dissonants ;

[18] On peut tout aussi bien mesurer la notion duale de la dissonance : la consonance.

3. *communication* : la communication avec les autres est susceptible d'introduire de nouveaux éléments qui sont dissonants avec ceux de l'agent ;
4. *existence simultanée de différentes cognitions dont certaines sont consonantes et d'autres dissonantes* : dans le cas général, une cognition est liée à plusieurs autres dont certaines sont consonantes et d'autres dissonantes.

Un état de consonance est un état d'équilibre et aucune force n'agit pour changer les relations entre les différentes cognitions de cet état. Au contraire, un état dissonant fait naître une pression, une motivation, pour l'éliminer et rétablir la consonance. Pour éliminer cette tension, l'agent peut agir de différentes façons :

- changer la cognition afin qu'elle soit consonante avec la situation : c'est le principe de base de l'apprentissage, ou adaptation, dans la théorie de Festinger.
- changer la situation afin qu'elle soit consonante avec la cognition : un agent peut agir sur l'environnement pour l'amener dans un état où la cognition est de nouveau consonante avec la situation en question.

Si une consonance existe, l'agent va éviter les changements d'attitudes ou de comportements susceptibles d'introduire de la dissonance. De même, si une dissonance est présente, l'agent va éviter les changements de cognition susceptibles d'augmenter la magnitude de la dissonance et s'orienter vers des changements susceptibles de la diminuer. En l'absence d'équilibre, ces tendances se manifestent donc par des changements d'attitudes ou de comportements. Pour rendre la théorie plus utilisable, il faut pouvoir déterminer quelles sont les conditions qui décident si ce sont les attitudes ou les comportements qu'il faut changer. Cela dépend de rapports d'importance entre ces tendances et de la résistance au changement des différentes cognitions et comportements en jeu.

En effet, les cognitions ne sont pas toutes également manipulables. La probabilité qu'une cognition soit modifiée pour réduire la dissonance dépend de ce que Festinger nomme sa *résistance au changement*. La résistance au changement d'une cognition est directement fonction du nombre et de l'importance des éléments avec lesquels elle est consonante, de son ancienneté ainsi que de la manière dont elle a été acquise : perception, cognition ou communication.

Festinger a également écrit sur le lien entre dissonance cognitive et communication en se restreignant au problème de l'acquisition des cognitions. Pour Festinger [Festinger, 1954], un individu a deux sources majeures d'informations : sa propre expérience et la communication avec les autres. L'impact de l'expérience directe est plus grand en ce qu'elle exerce une forte pression cognitive pour s'y conformer. En effet, la communication peut être vue comme une source d'expérience indirecte. L'intensité de l'impact des communications dépend de la relation entre ceux qui communiquent. Cette relation peut être analysée en termes de rôle, de confiance/réputation, d'attraction et de passif. Plus cette relation est « forte », plus la communication aura de l'impact sur la cognition des agents communicants. Souvent, ces deux sources, directe et indirecte, sont utilisées simultanément : un enfant peut apprendre de sa mère que le feu est dangereux et également se brûler au toucher.

10.3 Formalisation de la dissonance cognitive en terme d'éléments et de contraintes

Notre reformulation de la théorie de la dissonance cognitive emprunte à la théorie de la cohérence du philosophe computationnel Thagard [Thagard and Verbeurgt, 1998] qui nous permet de faire directement le lien entre la théorie de la dissonance cognitive et les notions, plus communes en informatique, d'éléments et de contraintes. Dans notre théorie, les *éléments* sont les cognitions privées et publiques des agents : croyances, désirs, intentions et engagements. Les éléments sont partitionnés en deux ensembles : l'ensemble A des éléments acceptés (qui sont interprétés comme crus vrai, activé ou valide selon le type des éléments) et l'ensemble R des éléments rejetés (qui sont interprétés comme crus faux, inactivés ou non valides selon le type des éléments). Tous les éléments qui ne sont pas explicitement acceptés sont rejetés. Les *contraintes* binaires sur ces éléments sont induites des relations qui existent entre ces éléments dans le modèle cognitif de l'agent :

– *Les contraintes positives* : des contraintes positives correspondent aux relations de cohérence ou de consonance que sont : les relations d'explication, les relations de déduction, les relations de facilitation et toutes les associations jugées positives.

– *Les contraintes négatives* : des contraintes négatives sont induites des relations d'incohérence ou de dissonance : exclusion mutuelle, incompatibilité, inconsistance et toutes les relations jugées négatives.

À chacune de ces contraintes est attribué un poids reflétant le degré de validité de la relation sous-jacente. Ces contraintes peuvent être satisfaites ou pas : une contrainte positive est satisfaite, si et seulement si, les deux éléments qu'elle lie sont soit tous les deux acceptés soit tous les deux rejetés. À l'inverse, une contrainte négative est satisfaite si et seulement si un des deux éléments qu'elle lie est accepté et l'autre rejeté. Ainsi, deux éléments sont dits *cohérents* s'ils sont liés par une relation à laquelle correspond une contrainte satisfaite. Et inversement, deux éléments sont dits *incohérents* si et seulement si ils sont liés par une relation à laquelle correspond une contrainte non-satisfaite.

Étant donné une partition des éléments entre A et R, on peut mesurer le *degré de cohérence* d'un élément en calculant la somme des poids des contraintes afférentes à celui-ci qui sont satisfaites divisée par le nombre des contraintes afférentes. Et symétriquement, on peut mesurer le *degré d'incohérence* d'un élément comme la somme des poids des contraintes non satisfaites divisée par le nombre total de contraintes afférentes. De la même façon, on peut mesurer la cohérence d'un ensemble d'éléments comme la somme des poids des contraintes afférentes à cet ensemble (les contraintes dont au moins un pôle est un élément de l'ensemble considéré) qui sont satisfaites divisée par le nombre des contraintes afférentes total. Symétriquement, l'incohérence d'un ensemble de cognitions peut être mesurée comme étant la somme des poids des contraintes non-satisfaites afférentes à cet ensemble divisée par le nombre de contraintes afférentes total.

Dans ce cadre et en complet accord avec le principe de cohérence, l'hypothèse de base de la théorie de la dissonance cognitive est que l'incohérence (ce que Festinger nomme la dissonance) produit chez le sujet une tension qui l'incite au changement. Plus l'incohérence est intense, plus l'insatisfaction est forte et plus la motivation pour la réduire l'est aussi. L'incohérence d'une cognition peut être réduite en (1) supprimant ou réduisant l'importance des cognitions incohérentes ou en (2) ajoutant ou augmentant l'importance des cognitions cohérentes.

La seconde hypothèse de Festinger reste qu'en cas d'incohérence, l'individu ne va pas seulement changer ses cognitions ou essayer de changer celles des autres pour essayer de la réduire, il va aussi éviter toutes les situations qui risquent de l'accroître.

Un des intérêts majeurs de la théorie de la dissonance cognitive capturé par notre formulation est de fournir une mesure de l'incohérence (la dissonance), c'est-à-dire une métrique de la cohérence cognitive. Les mesures utilisées par Festinger (définies à la section 10.2) correspondent précisément aux mesures d'incohérence définies précédemment. Dans notre formalisation, une contrainte non-satisfaite est une relation de dissonance tandis qu'une contrainte satisfaite est une relation de consonance.

10.4 Dissonance, changement d'attitude et influence sociale

Dans les SMAs, la question de savoir quand l'agent doit essayer de modifier l'environnement (entre autres, la couche des engagements sociaux publics) pour satisfaire ses intentions et quand l'agent doit modifier ses états mentaux pour être cohérent avec son environnement est cruciale. Dans notre modèle, l'agent cherche à maximiser sa cohérence, c'est-à-dire qu'il cherche à réduire ses incohérences en commençant par la plus intense. Pour réduire une incohérence, l'agent doit accepter ou rejeter certaines cognitions de manière à satisfaire au mieux les contraintes qui les lient. Ces cognitions peuvent être privées (les états mentaux) ou publiques (les engagements). Mais toutes les cognitions ne sont pas également modifiables : c'est la notion de résistance au changement d'une cognition définie à la section 10.2. Pour pouvoir intégrer la communication dans notre modèle, il faut maintenant introduire le lien fondamental qui existe entre notre formulation de la théorie de la dissonance cognitive et la notion d'engagement.

Les engagements sociaux sont des cognitions particulières qui ne sont pas modifiables individuellement, mais doivent être socialement établis (c'est l'objet des jeux de dialogue que de fournir des outils pour établir des engagements conjointement acceptés). C'est-à-dire que pour modifier, faire abandonner ou faire accepter un nouvel engagement afin de réduire une incohérence, un agent doit dialoguer. C'est par le dialogue que les agents vont essayer de faire établir les engagements sociaux cohérents avec leurs autres cognitions. Cependant, à l'issue de ces dialogues, certains engagements peuvent rester incohérents tout en n'étant plus modifiables. Ce sont alors des « obligations sociales » et cela fixe un des pôles des contraintes qui leur sont liées. Pour réduire d'éventuelles incohérences tout en se conformant aux engagements pris, c'est alors ses états mentaux que l'agent devra changer pour rétablir la cohérence. C'est le ressort du changement d'attitude dans notre système et cela formalise la vision des psychologues Brehm et Cohen à ce sujet [Brehm and Cohen, 1962], soutenue par des myriades d'expériences.

Pour ces psychologues, l'individu va chercher (par le dialogue) les engagements sociaux qui sont cohérents avec ses attitudes. Dans un débat d'idées, par exemple, chacun va poser des engagements propositionnels correspondant à ses croyances et être prêt à s'en justifier. Néanmoins, il n'est pas rare que l'agent accepte (ou se retrouve engagé sur) des engagements qui sont incohérents avec ses attitudes. Prenons l'exemple de l'argumentation : si l'individu est confronté à des éléments incohérents, il peut argumenter pour défendre la cohérence de son point de vue. Supposons que de cette argumentation ressorte que le point de vue de l'interlocuteur est plus convainquant, plus acceptable que le sien (au sens du système d'argumentation choisi), il va alors devoir accepter/adopter les engagements posés par son interlocuteur. Ces engagements étant incohérents avec ses attitudes et les tentatives de modifier ces engagements ayant déjà échouées, il lui faut effectuer les changements d'attitudes rétablissant la cohérence.

En effet, lorsque des engagements non modifiables sont incohérents avec ses attitudes, un des pôles de l'incohérence est fixé. Pour rétablir la cohérence, l'agent ne peut alors qu'essayer de revenir sur ces engagements dissonants (ce qui n'est pas toujours possible et porte généralement à conséquence : sanction sociale, ...) ou bien changer ses attitudes.

De nombreuses expériences, détaillées dans la littérature [Brehm and Cohen, 1962], viennent préciser ce point. Plus une attitude est importante, moins elle est sujette à changement. Néanmoins, plus l'incohérence révélée par la communication est importante, plus les possibilités de changement sont grandes. L'incohérence entre engagements et attitudes mentales privées est révélée par les communications (par exemple, je reçois un ordre d'un supérieur qui m'amène à m'engager à réaliser ce qui est demandé). La théorie de la dissonance cognitive ne fait pas de prédiction concernant le fait que le locuteur s'engage ou non sur des éléments dissonants, c.-à-d. incohérents. Cependant, dans notre théorie de la cohérence cognitive, les outils fournis peuvent être utilisés à cette fin. Lorsque cela arrive, c.-à-d. un agent s'engage sur une voie incohérente avec ses attitudes (généralement par obligation sociale), il se peut alors que ces attitudes soient modifiées pour réduire l'incohérence. Des expériences montrent cela [Brehm and Cohen, 1962]. En effet, Brehm a montré au travers des goûts culinaires que moins une personne aime un met, plus le fait de s'être engagé à en manger va susciter un changement d'attitude envers ce produit. Évidemment, ladite personne aura a priori tendance à éviter ce type d'engagements. C'est-à-dire que ce phénomène ne survient que si l'individu est obligé de maintenir cet engagement. Sinon, toutes les autres tentatives de réduction sont observées : tentative de décrédibiliser ce pour quoi ou celui pour qui il s'est engagé, tentative de désengagement, rejet de la communication, rejet de l'information, ... Il faut donc que toutes ces possibilités soient rendues difficiles par la situation (par exemple, du fait des sanctions sociales ou autres qui pèsent sur le désengagement).

Néanmoins, et pour faire le lien avec la théorie du jugement de l'école de Yale, il n'est pas rare que les engagements pris en définitive dans un dialogue soient des compromis entre les attitudes défendues par les différents interlocuteurs (c'est par exemple le cas typique dans les dialogues de négociation). De sorte que, pour réduire l'incohérence externe qui les occupait, chacun des agents accepte de réduire une incohérence interne plus faible en changeant ses attitudes.

La section 12.3 présente un exemple de ce mécanisme de changement d'attitude dans une application de la théorie de la cohérence cognitive à l'automatisation de l'utilisation des jeux de dialogue dans les SMAs.

11 Extension à la communication agent

11.1 Application aux SMAs

Le corps de notre théorie de la cohérence cognitive est ainsi construit en deux étapes : (1) une unification de la théorie de la dissonance cognitive de Festinger avec celle de la cohérence cognitive de Thagard[19] et (2) une extension fondée sur les mécanismes de changement d'attitude et la notion de résistance au changement pour y inclure le traitement de la communication[20]. Cela en fait une théorie cognitive générale et complète. Les mesures qu'elle

[19]Cette unification permet de garder les propriétés de la théorie de la cohérence cognitive de Thagard en capturant celles de la théorie de la dissonance cognitive

[20]Cette extension embrasse la théorie de Festinger et celle de Thagard.

propose définissent une métrique de la cohérence cognitive de l'agent et du groupe. Le modèle de cohérence cognitive que nous proposons s'articule autour de deux axes :

- *moteur cognitif et comportemental* : un agent travaille à maintenir une cohérence interne la plus élevée possible. Une incohérence qui ne peut être réduite par l'agent seul fera l'objet de communications ou de changements d'attitudes.
- *moteur social* : dans un cadre coopératif (au sens de Camps [Camps, 1998]), les agents travaillent à maintenir la cohérence du groupe la plus élevée possible.

Ces deux dimensions[21] répondent au besoin d'un système de régulation (une homéostasie) de la cohérence cognitive aussi bien au niveau de l'agent qu'au niveau du groupe, au moins dans les situations coopératives. En effet, les agents cognitifs sont des systèmes complexes et cette complexité est accrue par leur multiplicité dans le groupe. Ainsi, se pose pour les systèmes multi-agents les mêmes questions que pour les systèmes complexes : garantir la convergence, dans notre cas la cohérence individuelle ou collective. La théorie de la cohérence cognitive fournit un tel système de régulation de la cohérence de méta-niveau. Il nous semble que cette théorie fournit de nombreuses indications pour résoudre des problèmes de conception des systèmes multi-agents. En outre, elle permet de capturer un grand nombre de dimensions de la communication inter-agents. Dans les sections suivantes nous indiquons différents problèmes théoriques, spécifiques ou pas aux communications dans les systèmes multi-agents, pour lesquels cette théorie peut être utile.

11.2 Typologie des incohérences

Cette sous-section propose une typologie des incohérences qui adapte notre formulation de la théorie de la dissonance cognitive à un cadre explicitement distribué comme celui des SMAs. La typologie que nous avançons a pour but d'introduire un vocabulaire simple mais utile pour traiter des problèmes de cohérence. L'incohérence étant conceptuellement très proche de la notion de conflit, la typologie suivante est empruntée aux travaux sur le concept de conflit [Pasquier and Dehais, 2000] :

- *incohérences internes et externes* : une incohérence est *interne* quand toutes les cognitions impliquées sont relatives à un même agent et *externe* quand les cognitions incohérentes impliquent au moins deux agents. Plus concrètement, une incohérence est externe pour un agent si c'est une incohérence entre des éléments de son modèle du monde et des éléments de ce qu'il connaît des modèles des autres.
- *incohérences explicites et implicites* : nous définissons *explicite* par le fait d'être dans « l'état d'avoir connaissance de » et *implicite* par celui d'être dans « l'état de ne pas avoir connaissance de ». On peut avoir connaissance de quelque chose sans être dans « l'état d'avoir connaissance ». C'est le cas de l'oubli par exemple (on a la connaissance que pour conduire la nuit, il faut allumer les phares, mais il peut arriver que l'on oublie). Une incohérence est *explicite* pour un agent si toutes les cognitions qui y participent

[21] Pour des raisons de temps, lors de cette recherche, nous n'explorerons en fait qu'une partie du premier axe, c'est-à-dire que nous considérerons la théorie comme un moteur comportemental pour la cognition liée à la communication. Une étude séparée serait requise pour montrer que la dimension sociale émerge de ce premier postulat.

sont explicites pour l'agent. Une dissonance est *implicite* si au moins une des cognitions incohérentes est implicite pour au moins une des parties concernées. Il est à noter qu'une incohérence implicite est une incohérence explicite potentielle. Notons également que dans un cadre multi-agent l'incohérence interne sera sans doute toujours explicite. La figure de la section 11.4 détaille cette typologie et la section 11.8 fournit un exemple de chaque type de d'incohérence.

11.3 Lien cohérence - initiative, sujet et pertinence

Un problème particulièrement délicat est celui de l'ouverture du dialogue. À quel moment, et pourquoi, un agent doit-il prendre l'initiative d'un dialogue ? La réponse fournie par notre cadre de cohérence est qu'un agent prend l'initiative s'il a une incohérence à réduire et qu'il ne peut la réduire seul. Soit parce qu'il sait que c'est une incohérence externe qui implique d'autres agents, soit parce qu'il s'agit d'une incohérence interne et qu'il n'a pas les capacités de la résoudre seul (il doit alors compter sur la coopération des autres agents). On peut régler le seuil de l'intensité de l'incohérence à partir duquel un agent prend l'initiative du dialogue, cela définit une partie de son caractère. C'est un équilibre à régler de la même façon que celui des priorités entre attitudes propositionnelles qui définissent le caractère d'un agent, comme dans l'extension BOID [Broersen et al., 2001] du cadre BDI. Cette incohérence qui initie le dialogue en donne également le sujet.

Avec la théorie de la pertinence, Sperber et Wilson [Sperber and Wilson, 1986] avançaient l'idée que le locuteur choisit ce qu'il va dire en évaluant dynamiquement la pertinence de ses idées. Cette pertinence varie pour chaque élément de cognition au cours de la conversation. Le locuteur ne s'engage dans un acte de langage que lorsque la pertinence en est maximale. Avec notre approche, un agent qui prend l'initiative va s'attaquer à l'incohérence qui a la plus grande magnitude, car c'est ce qui est cognitivement pertinent pour lui. La section suivante indique comment le cadre de cohérence permet à l'agent de choisir quel type de dialogue engager.

11.4 Lien avec les types de dialogues

Dans cette sous-section, nous analysons comment les types de dialogues observés en dialectique peuvent être liés à la cohérence cognitive. Un certain nombre de travaux récents utilisent pour les dialogues la typologie de Walton et Krabbe [Walton and Krabbe, 1995]. Ils y distinguent six types de dialogues définis par leur but premier (auquel les interlocuteurs souscrivent) et les buts/intentions propres à chacun (qui peuvent être incompatibles, c.-à-d. incohérents).

1. *la persuasion* : la situation initiale est une incohérence externe de point de vue et le but global est de la résoudre. Chaque participant essaye de ne pas changer ses croyances (c'est la résistance au changement) et de faire changer celles des autres. Pour ce faire, les agents ont typiquement recours à l'argumentation [Keefe, 1991]. La persuasion est donc une technique de réduction d'incohérence externe. Ceci étant dit, il s'agit en fait pour un agent qui souhaite persuader, de montrer aux autres que son point de vue est plus cohérent que le leur. Alors qu'argumenter se résume, pour un agent, à exhiber des éléments cohérents avec sa proposition (c'est-à-dire liés à elle par des contraintes satisfaites). À l'inverse, attaquer/réfuter l'autre, consiste à exhiber des éléments incohérents avec sa proposition (c'est-à-dire, liés à elle par des contraintes non satisfaites).

2. *la négociation* : à partir d'un conflit d'intérêts (un type d'incohérence externe), le but global est de conclure un contrat, d'arriver à un accord. Chaque agent a son propre but et veut maximiser son profit, ses intérêts. La résolution du conflit se fait habituellement par un échange d'offres et de contre-offres. Il est fréquent que des dialogues de persuasion soient imbriqués dans une négociation, les offres étant ainsi argumentées. C'est une technique de réduction d'incohérence externe.

3. *l'investigation* : les participants de ce type de dialogues sont dans une situation d'incohérence interne partagée (ils souffrent tous d'une incohérence interne qui a été reconnue comme commune). Le but commun coïncide avec les buts individuels. Il s'agit de prouver un fait pour renforcer la cohérence. C'est une réduction d'incohérence interne partagée (c'est-à-dire que les agents impliqués dans le dialogue ont reconnu qu'ils ont la même incohérence à résoudre et ils s'entraident)

4. *la délibération* : chaque agent a ses préférences et tous doivent choisir parmi les offres de chacun. Les participants ont pour but commun de prendre une décision (choisir un plan ou une action). Leur but individuel est d'influencer la décision dans leur intérêt (les préférences des uns et des autres ayant généralement été reconnues différentes, pour qu'il y ait délibération)ou dans ce qu'ils considèrent être l'intérêt commun (afin d'accroître leurs cohérences). C'est une technique de réduction d'incohérence externe explicite.

5. *la recherche d'informations* : c'est le seul type de dialogue qui est toujours asymétrique. Un agent cherche à obtenir de l'information des autres. Il s'agit d'une technique de réduction d'incohérence interne. Dans ce type de dialogue, seul l'agent demandeur d'information est dans un état d'incohérence. La réduction est asymétrique, mais pour la faciliter, il est fréquent que le demandeur explicite son incohérence aux autres agents (en indiquant pourquoi il cherche de l'information, c'est-à-dire en explicitant son incohérence). Cette technique de réduction peut prendre la forme d'un dialogue, mais peut aussi prendre la forme d'autres actions (par exemple : lire un livre de référence, chercher sur Internet, ...) pour autant que l'incohérence soit réduite.

6. *la dispute* : la situation initiale est conflictuelle et incohérente. Contrairement aux autres types de dialogues, l'engueulade n'est pas rationnelle, elle fait généralement plus appel aux sentiments, émotions et pulsions qu'à la raison et à la cohérence. C'est pourquoi, nous ne la détaillerons pas ici.

Comme on peut le constater, tous les dialogues naissent d'une situation d'incohérence. Notons que la réciproque n'est pas vraie, c.-à-d. toutes les incohérences ne sont pas traitées par le dialogue.

> *Notre conjecture est donc que les agents communiquent si une incohérence les y contraint.*

Dès lors, la conversation est vue, outre ses caractéristiques déjà connues, comme une procédure générique d'explicitation et de tentative de réduction ou d'évitement d'incohérence. De par la proximité conceptuelle entre les notions de dissonance, d'incohérence et celle de conflit, cette conjecture est proche de la position classique de la dialectique : tout dialogue naît d'un conflit [Hamblin, 1970]. Depuis quelques années, plusieurs auteurs insistent de nouveau sur le rôle des conflits dans le processus conversationnel :

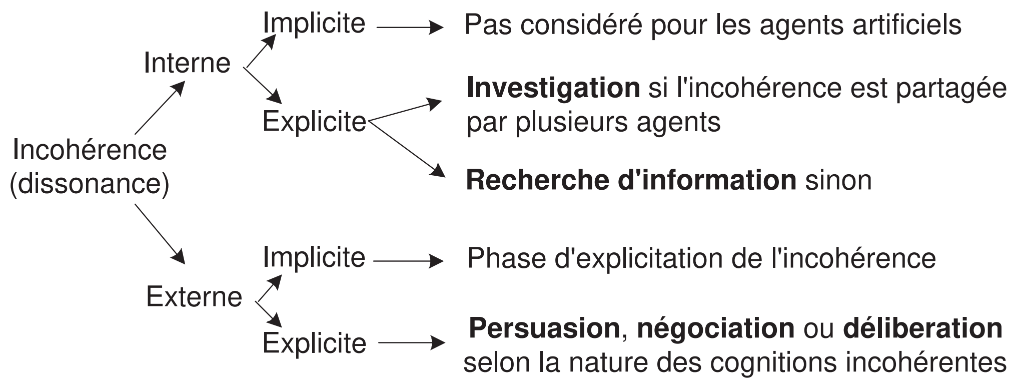

FIG. 3 – Typologie des incohérences/dissonances cognitives et lien avec les types de dialogue.

– pour Walton et Krabbe [Walton and Krabbe, 1995], la question « Y a-t-il un conflit ? » est à la base de l'analyse des situations initiales du dialogue ;

– pour Reeds et Long [Reed and Long, 1997],« nombre de dialogues ont pour situation initiale un conflit » ;

– pour Dessalles [Dessalles, 1998], « un grand nombre de dialogues trouvent leurs origines dans les conflits cognitifs entre désirs et/ou croyances » ;

– pour Baker [Baker, 1991], « les dialogues résultent d'une opposition entre des buts conflictuels ».

Finalement, la figure 3 résume les différents types d'incohérence ainsi que les types de dialogue qu'elles peuvent engendrer dans le cas général. L'exemple présenté à la section 12.3 présente le cas de ce choix dans le cadre plus spécifique des jeux de dialogue tels que définis dans le langage DIAGAL [Chaib-draa et al., 2002].

11.5 Lien cohérence - projet conjoint

La communication entre agents se fait par conversation, séquence d'actes de langage dont la somme des sens isolés ne rend pas compte de la signification. C'est pourquoi, de nombreux chercheurs rendent aux conversations leur dimension sociale et tentent d'analyser le dialogue du niveau conversationnel vers le niveau des actes de langage plutôt que le contraire. Cohen et Levesque ont proposé leur point de vue avec les buts persistants et les croyances mutuelles [Cohen and Levesque, 1990b]. Grosz [Grosz and Kraus, 1996] a proposé les plans partagés, Singh l'engagement commun [Singh, 1994] et Traum [Traum, 1997] étudie les croyances mutuelles. Chaib-draa et Vongkasem [Chaib-draa and Vongkasem, 2000] partent dans la même voie avec les idées de Searle et Clark [Clark, 1996] qui voient la conversation comme une activité conjointe. Nous proposons de voir le dialogue comme *un projet conjoint de réduction d'incohérence*.

A notre sens, le modèle de cohérence cognitive peut fournir des éléments de compréhension sur la nature du caractère conjoint des conversations. Selon notre hypothèse et en accord avec le principe de cohérence, les agents communiquent pour réduire, éviter ou empêcher l'accroissement d'une incohérence cognitive.

Dès lors que le dialogue est entamé parce qu'une incohérence ne peut être réduite par un agent seul, cette réduction (le dialogue) est un projet commun, une activité conjointe. On y reconnaît bien une intention commune quant à un but commun, même si cette intention reste généralement implicite dans les dialogues entre humains[22]. Ajoutons que les étapes de mise en commun (établissement, acceptation et validation[23]) introduites dans le cadre interactionnel sont indispensables pour que cette réduction commune soit un succès. Sans mise en commun, la réduction risquerait d'être unilatérale c'est-à-dire un leurre, un malentendu, une incompréhension, un quiproquo, ...

Nous différencions deux aspects dans cette mise en commun. Il y a une mise en commun structurelle capturée par les jeux de dialogue. Par exemple, il faut s'assurer que lorsqu'un joueur propose une négociation, l'autre accepte de négocier avant que celle-ci ne débute. Les agents doivent être d'accord sur le type de dialogue dans lequel ils s'engagent. Mais comme les dialogues sont des tentatives de réduction de dissonance, une mise en commun cognitive de cette dissonance est également nécessaire à leur bon déroulement. Cela signifie que le sujet du dialogue, la dissonance à laquelle le dialogue s'attaque doit être explicité pour tous les agents qui participent au dialogue. C'est-à-dire qu'en plus d'être prêts à négocier, il faut que les agents soient d'accord sur le problème, c.-à-d. la dissonance à régler par la négociation.

En différenciant les incohérences externes implicites des dissonances externes explicites, nous avons souhaité montrer en quoi la phase d'explicitation de la dissonance est cruciale. En effet, pourquoi un agent accepterait-il de s'engager dans un type de dialogue si cela ne règle pas un problème dans lequel il se sait impliqué ? Plus précisément, un agent peut être amené à communiquer à cause d'une incohérence interne ou externe (incohérence entre son modèle du monde et ce qu'il connaît des modèles d'autrui). Dans les deux cas, le dialogue devra passer par une phase d'explicitation de la dissonance. L'incohérence devra être mise en commun, l'exemple 3 illustre cette nécessité.

En effet, les dialogues de type négociation, persuasion et délibération sont tous initiés dans des cas d'incohérences externes (voir section 11.4). Les agents qui participent au dialogue sont généralement les agents impliqués dans ladite incohérence et sa résolution constitue leur projet conjoint. Dès lors, il est indispensable que chaque étape soit validée, acceptée par chacun des agents (établissement) ce sans quoi la résolution ne sera pas commune et la dissonance risque de subsister.

Dans le cas de l'investigation, l'incohérence n'est pas externe mais interne partagée. Les agents doivent alors expliciter cette incohérence pour s'assurer qu'elle est bien commune (et que les incohérences de chacun ne diffèrent pas trop, de sorte que le problème à résoudre soit adopté comme étant commun). Ils vont ensuite entamer la résolution/réduction comme projet conjoint.

Finalement, dans le cas de la recherche d'informations, le projet qui doit être adopté comme commun vise à réduire l'incohérence interne du demandeur d'information. Généralement, celui-ci partage son incohérence avec un agent choisi pour sa bonne volonté coopérative, ses compé-

[22]Les expériences montrent bien que les humains agissent sans connaître ni reconnaître la théorie de la dissonance cognitive. Il y a néanmoins des domaines comme celui de la recherche scientifique où les individus explicitent une problématique (une dissonance, un conflit, une incohérence) avant d'en débattre [Dessalles and Ghadakpour, 1999].

[23]Notions abordées en section 3.3.

tences ou la qualité de leur relation dans l'espoir que ce dernier l'aide à résoudre son incohérence.

Exemple 3

L'agent A prévoit d'utiliser la ressource non-partageable R pendant un intervalle de temps futur T et l'agent B aussi. Tant que A et B n'ont pas communiqué, l'incohérence reste implicite : seul un observateur averti peut avoir conscience du conflit de ressource qui attend nos deux agents. En l'absence d'un système de gestion de la ressource R, pour qu'un agent divulgue spontanément son intention quant à la ressource R, il faut :

- qu'il sache que R est une ressource non partageable et qu'il y a donc un risque de conflit potentiel ;
- qu'il sache que l'incohérence (ou le conflit) sera d'autant plus intense qu'elle adviendra au dernier moment (le moment de l'utilisation de la ressource), ce qui l'amène à désirer anticiper et à expliciter l'incohérence potentielle sur le champ ;
- qu'il sache qui pourrait vouloir utiliser la ressource (ici B) ;

Muni de ces connaissances contextuelles, A sait quoi communiquer et à qui le communiquer tout comme il sait pourquoi cela peut être utile. Il saura a posteriori si la conversation aura été utile ou pas selon qu'une incohérence externe aura été explicitée ou pas, puis réduite ou pas.

11.6 Lien dissonance - utilité et dynamique du dialogue

11.6.1 Utilité des dialogues

Dans la théorie de cohérence cognitive définie ci-dessus, les agents peuvent calculer l'*espérance d'utilité* des dialogues qu'ils envisagent. Cette espérance d'utilité est égale à l'intensité de l'incohérence de l'élément auquel le dialogue s'attaque à laquelle on retranche l'intensité de l'incohérence de cet élément après le dialogue, si celui-ci réussit en faveur de l'agent. Les agents peuvent également calculer l'*utilité* d'une conversation dynamiquement en recalculant l'incohérence de l'élément en question au cours du dialogue. Lorsqu'une unité de dialogue est terminée, soit l'incohérence est réduite, soit l'agent peut continuer (par un enchaînement dialogique) à essayer de la réduire (dans la mesure du possible, car certains dialogues ne peuvent pas être remis en cause lorsqu'ils ont échoués).

11.6.2 Dynamique intra-dialogue

Un agent sélectionne un type de dialogue en fonction du type de l'incohérence qu'il souhaite réduire, c.-à-d. du type de problème qu'il souhaite régler. Mais au cours de cette résolution, d'autres incohérences peuvent apparaître. Celles-ci peuvent parfois devoir être réduites pour que la réduction principale puisse se poursuivre. C'est ce qui amène les agents à imbriquer un sous-dialogue pour réduire la nouvelle incohérence avant de reprendre le dialogue principal concernant l'incohérence première. Dans d'autres cas, l'incohérence peut se déplacer, amenant les agents à enchaîner deux types de dialogues. On a donc également un outil pour gérer l'imbrication et l'enchaînement des unités de dialogue au cours d'une conversation.

Par exemple, il est habituel de délibérer sur les modalités de paiement ainsi que des conditions de réception du produit lors de la négociation d'un produit. Les agents souhaitent que cet aspect du contrat recherché soit cohérent avec leurs attentes (probablement divergentes) à ce sujet. Le fait que cela ne soit pas le cas est une incohérence (au sens de notre théorie). C'est l'irruption de cette nouvelle incohérence liée au reste de la négociation du prix qui va amener les agents à imbriquer une délibération. En effet, les modalités de paiement sont un élément important de la négociation d'un produit et son imbrication dans cette dernière garantit de pouvoir tenir compte de cette sous-résolution pour la suite de la négociation.

11.6.3 Dynamique inter-dialogues

Dans notre théorie, les dialogues ne sont que des tentatives de réduction d'incohérence ; de telles tentatives peuvent échouer. La mesure d'utilité des dialogues définie dans le cadre de consonance permet de guider l'agent dans sa conduite communicationnelle. En effet, suite à un dialogue pas ou peu utile, i.e. la dissonance n'est pas réduite, l'agent doit décider comment réagir. L'agent va probablement persévérer dans sa tentative de réduction en prenant en compte cet échec : il proposera un type de dialogue différent ou une proposition différente du même type de dialogue, il prendra note de cet échec qui pourra être utile pour le guider lors des dialogues/tentatives suivants. C'est, par exemple, cette mesure d'utilité qui peut lui permettre de ne pas tenir plusieurs fois de suite le même dialogue infructueux avec le même agent.

En particulier, dans un cadre ouvert et hétérogène, un agent est amené à communiquer avec des agents inconnus. Il lui faut alors se faire une idée des dialogues tenus avec ceux-ci. L'agent pourra tenir compte de l'utilité des dialogues tenus pour sélectionner ses interlocuteurs. Un agent aura intérêt à renforcer ses échanges avec les agents avec lesquels les dialogues sont utiles et de nombreuses incohérences (c.à.d. problèmes) ont été résolues. À l'inverse, il pourra tenir compte des dialogues inutiles en affaiblissant ses liens sociaux avec les interlocuteurs concernés. C'est-à-dire que la mesure de l'utilité des dialogues fournit une information précieuse qui pourra être utilisée par un outil de gestion des accointances (qui pourrait être opérationnalisé via l'apprentissage par renforcement, ce qui n'est pas notre objet ici).

Par exemple, en commerce électronique, un agent n'a pas intérêt à continuer de tenter des négociations qui n'aboutissent jamais (peut-être ses conditions de paiement et celles de l'interlocuteur sont incompatibles et les sous-dialogues de négociation de celles-ci n'aboutissent pas non-plus). Au contraire, il aura intérêt à procéder avec des agents avec qui il tient des dialogues utiles. Dans ce cadre de commerce électronique, le cadre de dissonance (on suppose la modélisation réussie) garantit que les dialogues les plus utiles correspondront aux contrats les plus intéressants, les plus satisfaisants.

11.7 Lien cohérence - humeur, intensité

Ces dernières années, le besoin d'intégrer les émotions aux agents artificiels s'est fait sentir [Bates, 1994, Velasquez, 1997]. Le modèle de consonance proposé permet de faire un lien direct entre les mesures de cohérence et l'humeur de l'agent. Notre théorie fournit un système de valeurs dans lequel un gain de cohérence est un soulagement, une joie, un réconfort. Un état de cohérence est un état de bien-être (sourire, aspect relaxé, ...). À l'inverse, un agent peut avoir peur d'une incohérence perçue comme potentiellement future, être préoccupé ou malheureux d'un état d'incohérence ou encore déçu d'une tentative de réduction échouée (par exemple une conversation). On pense aux interfaces hommes/machines et aux systèmes de tutoriels intelligents, entre autres. En outre, certains cadres interactionnels autorisent l'utilisation de différents degrés d'intensité des forces illocutoires des actes de langage[24]. Pourtant, aucune théorie d'agents n'indique comment la sélection de ce degré d'intensité s'effectue. Les mesures quantitatives définies par la théorie de la cohérence cognitive fournissent selon nous le moyen de guider l'agent dans le choix du degré d'intensité approprié.

Dès lors qu'une conversation est engagée pour réduire une incohérence, il semble légitime que la magnitude de celle-ci influence de manière directe le choix des degrés d'intensité des actes à utiliser. Par exemple, un agent qui a besoin d'une information pour réduire une incohérence interne va se lancer dans un dialogue de recherche d'informations et produire un acte directif pour essayer d'obtenir cette information. Le degré d'intensité de la force illocutoire va dépendre du degré d'intensité de ladite incohérence :

– une invitation ou un conseil, si l'incohérence est très légère ;

– une recommandation ou une demande, si elle est un peu plus intense ;

– une supplication, une imploration ou un ordre, si la magnitude de l'incohérence est très intense et donc sa réduction cruciale.

Si ces paramètres d'émotions, d'humeur et d'intensité semblent moins importants pour les SMAs entièrement artificiels, cette piste est intéressante pour les systèmes homme-machine. Évidemment, ce facteur de sélection du degré d'intensité n'est pas unique, car il y a un certain nombre d'autres facteurs susceptibles d'intervenir dans ce choix : les conventions sociales (il est généralement interdit de donner un ordre à un supérieur hiérarchique. . .) et les relations entre les agents (proximité, confiance, passif de la relation, . . .) sont aussi importantes pour sélectionner les degrés d'intensité des forces illocutoires des actes posés.

11.8 Exemples supplémentaires

Cette sous-section présente deux exemples informels qui couvrent tous les types d'incohérence que l'on peut rencontrer dans les SMAs ainsi que tous les types de dialogues tels que définis par Walton et Krabbe [Walton and Krabbe, 1995]. On suppose que les agents ont le modèle de consonance intégré à leur fonction d'utilité pour le choix des actes de langages, des actes conversationnels ou des coups du jeu selon le cadre interactionnel utilisé.

Incohérence externe implicite et explicite (persuasion, argumentation, délibération)

Pour illustrer ce type d'incohérence, considérons deux agents A et B qui voyagent ensemble en voiture vers un lieu L. A est le chauffeur et B le passager. À un moment donné, ils s'arrêtent à un carrefour où ils peuvent tourner à droite ou à gauche. L'agent A a l'intention de passer par la droite alors que B a l'intention de passer par la gauche. Tant que ces informations n'ont pas été communiquées, on est en présence d'une incohérence externe implicite (en effet, les deux intentions sont incompatibles).

[24]Cette idée est valable pour les actes de langage classiques comme pour les autres actes de conversation : actes d'établissement, actes de dialogue, . . .

1.A : *On prend à droite ?* (l'agent A cherche un support pour renforcer son état de cohérence)

L'impact de cette information sur l'agent B dépend de la relation qu'entretiennent A et B. Pour l'agent B qui a l'intention contraire, cette intervention de A explicite une incohérence externe. B peut calculer sa réponse selon la résistance au changement de son intention de passer par la gauche et l'impact de l'information donnée par A qui influence la magnitude de l'incohérence, c.-à-d. l'importance qu'il lui accorde.

Si la résistance au changement de cette croyance est faible et/ou l'impact des communications avec A fort, il peut changer ses croyances pour réduire l'incohérence (qui est maintenant explicite pour lui et implicite pour A) et répondre sans plus discuter :

2.B : *OK.* < fin de la conversation, incohérence réduite >

Si la résistance au changement de cette attitude est suffisamment élevée et/ou l'impact des communications avec A faible, il peut expliciter l'incohérence pour ouvrir une conversation plus complète :

2′.B : *Moi, je prendrais plutôt à gauche.* < B explicite l'incohérence pour A >

Les agents ayant mis leur modèle d'autrui à jour, l'incohérence est maintenant une incohérence externe explicite. A ce moment, si la résistance au changement de l'intention de A de tourner à droite est faible et/ou que sa relation avec B est forte, celui-ci peut abandonner son intention et réduire l'incohérence en concluant :

3.A : *OK, on passe par la gauche.* < fin de la conversation, incohérence réduite >

Si au contraire la résistance au changement de son intention est forte et/ou l'impact de la communication de B, il peut se lancer dans un dialogue de persuasion en réaffirmant ce qui est cohérent avec son intention. Cela peut être :

3′.A : *Par la droite, c'est plus court et nous sommes pressés.*

Selon ses états mentaux, B peut s'incliner ou essayer de persuader A à son tour :

4.B : *D'accord, allons y.* <fin de la conversation, incohérence réduite>

4′.B : *Oui, mais il y a des embouteillages à cette heure-ci, ce sera plus rapide par la gauche.* < etc., jusqu'à ce que l'un des points de vue l'emporte ou que d'autres impératifs l'emportent sur la conversation>

Ce faisant, la magnitude de l'incohérence augmente. Finalement, A peut s'incliner ou bien surenchérir selon la résistance au changement de son intention de passer par la droite et l'impact des communications avec B...

On note qu'il est difficile de savoir si les dialogues ci-dessus relèvent de la persuasion (chaque agent persuade l'autre d'adopter son point de vue), de la négociation (les deux agents négocient un chemin) ou de la délibération (chacun exprime ses préférences et le groupe doit choisir). Ce qui est sûr, c'est qu'il s'agit d'une réduction d'incohérence externe comme c'est le cas dans chacun de ces trois types de dialogues (voir 11.4). En outre, on peut constater que lorsque l'incohérence est réduite, le dialogue s'arrête. Le contraire n'est pas vrai car il se peut que l'incohérence persiste, mais que le dialogue cesse, car le feu passe au vert et A devant agir, tourne à droite.

Incohérence interne (explicite) et incohérence interne partagée (recherche d'informations et investigation)

On rappelle l'exemple 1 de la section 8 :

1.A : *Est-ce que je peux te poser une question ?*

2.B : *oui, vas-y.*

3.A : *Est-ce que tu as l'heure ?*

4.B : *Non, j'aimerais bien la connaître.*

Dans cet exemple, A commence un dialogue de type recherche d'information pour obtenir l'heure (énoncé 1 à 3) mais B explicite une incohérence interne commune en répondant qu'il a la même incohérence que A à résoudre, car il n'a pas l'heure et il souhaiterait la connaître aussi. Dans cette situation d'incohérence interne commune explicite, le seul type de dialogue (si les agents décident de poursuivre le dialogue pour réduire l'incohérence plutôt que toute autre action) possible est l'investigation (voir 11.4).

12 Application à l'utilisation automatique des jeux de dialogue dans les SMAs

12.1 Lier les cognitions privées aux cognitions publiques

Puisque les agents raisonnent avec leurs états mentaux privés, mais à propos des engagements publics, il est nécessaire d'établir un lien entre ces types de cognitions. C'est une condition nécessaire pour pouvoir utiliser les cadres interactionnels conventionnels décrits en section 5. Si pour l'heure cette liaison entre cognitions privées et cognitions publiques n'est proposée dans aucun de ces cadres interactionnels conventionnels (approches basées sur les engagements ou jeux de dialogue), c'est sans doute qu'elle relève plus de notre problématique, c'est-à-dire de leur utilisation, que de leur définition.

Pour établir ce lien, on ne considère que les états mentaux, cognitions privées des agents, qui font le « lien » avec le niveau extérieur - social - et qui sont ceux issus du raisonnement pratique (practical reasoning), de la délibération, à savoir les intentions (voir section 4.1.1). Nous faisons la distinction classique entre *intention de* (faire quelque chose ou faire faire quelque chose) et *intention que* (une certaine proposition tienne, c'est-à-dire, soit vraie ou considérée comme vraie) [Bratman, 1990]. Les cognitions sociales - publiques - considérées sont les engagements sociaux et on distingue les *engagements propositionnels* des *engagements en action* (voir section 5.1.3) [Bell, 1995, Walton and Krabbe, 1995]. Pour ce faire, on considère qu'*un engagement en action est le pendant socialement accepté d'une « intention de » et qu'un engagement propositionnel est le pendant socialement accepté d'une « intention que »*.

12.2 DIAGAL[DIAlogue Game based Agent communication Language]

Pour notre exemple, on suppose que dans le système considéré, ce sont les jeux de dialogue tels que définis avec Chaib-draa et Maudet [Chaib-draa et al., 2002] qui sont utilisés comme cadre interactionnel (la section 5.5.4 propose une introduction générale de cette approche au sein des différentes propositions basées sur les jeux de dialogue). Dans ce cadre interactionnel conventionnel, les engagements ne sont pas modifiables, ils constituent des « obligations sociales » qui amènent des sanctions si elles ne sont pas respectées. Un engagement accepté[25] de x envers y pris au temps t sur la proposition p (ou l'action α) qui entraîne les sanctions s_x s'il n'est pas respecté par x et s_y s'il est annulé par y, est noté : $C(x, y, p, t, s_x, s_y)$. Un engagement rejeté[26] est noté : $\neg C(x, y, p, t, s_x, s_y)$. Par souci de simplification, nous laisserons les sanctions et le temps de côté, pour l'instant.

Quatre jeux de dialogue sont définis dans DIAGAL : $offer$, $request$, $inform$ et ask. Les jeux sont en définitive des structures qui peuvent être utilisées par les agents dans l'espoir d'ajouter un engagement dans la couche publique. C'est-à-dire qu'utiliser un jeu de dialogue, c'est réaliser une tentative pour poser un engagement social. Il y a quatre types d'engagements qui peuvent être obtenus par l'agent initiateur d'un jeu :

– un engagement en action de lui envers l'interlocuteur ($offer$) ;

– un engagement en action de l'interlocuteur envers lui ($request$) ;

– un engagement propositionnel de lui envers son interlocuteur ($inform$) ;

– un engagement propositionnel de l'interlocuteur envers lui (ask).

Un jeu de dialogue est défini pour chacun de ces types d'engagement (le nom en est indiqué entre parenthèses dans la liste ci-dessus). Ces quatre jeux de base forment un système d'opérateurs « complet » pour l'ajout des engagements dans la couche publique[27]. Il est important de le noter, car ce n'est pas sans conséquence. Cela dit, cette monotonie[28] n'en est pas une véritable, car les engagements peuvent être retirés en étant : respectés, périmés, annulés, reniés, . . . avec les conséquences que l'on imagine en termes de sanctions individuelles et sociales. Les possibilités de structuration des jeux (séquencement, imbrication, pré-séquencement et itération) sont prises en charge via un méta-jeu de contrôle nommé jeu de contextualisation et qui permet de proposer l'entrée ou la sortie d'un jeu à n'importe quel moment. Notons que l'on distingue deux sous-types d'engagements : les engagements concernant la forme et la structuration du dialogue (indicés d'un petit j) et ceux concernant le fond et les aspects extra-dialogiques. Dans le même esprit que la distinction entre le plan du dialogue et le plan de la tâche dans les approches intentionnelles basées sur la planification, cela permet par exemple de distinguer un engagement de répondre à une question de l'engagement qui résultera éventuellement de la réponse. Lors de la sortie d'un jeu, tous les engagements résultants du jeu sont supprimés sauf ceux issus des éventuels sous-jeux et ceux correspondant à la condition de succès ou d'échec éventuellement atteinte.

[25] C'est à dire qui appartient à l'ensemble des éléments acceptés de tous les agents, car les jeux de dialogue et les conventions du système font en sorte que les engagements y soient socialement acceptés, ce qui n'est pas le cas des attitudes propositionnelles privées (croyances, désirs, intentions, ...).

[26] Même remarque concernant l'ensemble des éléments rejetés. Notons, qu'un engagement social rejeté est un engagement social possible, potentiel. Par convention, tous les engagements sociaux non explicitement acceptés sont rejetés.

[27] Cela signifie que toutes les modifications de la couche publique qui ont la forme d'un ajout d'engagement peuvent être obtenues par utilisation et/ou composition de ces quatre jeux. Ce résultat tient sous les hypothèses simplificatrices suivantes : (1) la couche publique ne contient pas d'engagement conditionnel et (2) les engagements sont posés un par un.

[28] Qui n'est qu'une première simplification, car les systèmes monotones sont toujours plus simples à étudier que leurs homologues non-monotones et qu'il n'a pas été posé explicitement comme hypothèse de modélisation que les jeux de dialogue ne devraient être conçus que pour l'ajout d'engagements.

Concrètement, chaque jeu est doté :

1. de *conditions d'entrée*, E, qui doivent être respectées pour que l'entrée dans le jeu soit possible ;
2. de *conditions de succès* pour l'initiateur, SI, qui sont remplies si le jeu est un succès en faveur de l'initiateur, c'est-à-dire que l'engagement recherché est accepté ;
3. de *conditions d'échec* pour l'initiateur, FI, qui sont remplies si l'engagement recherché est rejeté à l'issue du jeu ;
4. d'un *ensemble de règles de dialogue* R.

Les figures 4 et 5 rappellent ce quadruplet de définitions pour les jeux $request$ et $offer$ sus-nommés, à savoir $< E_r, SI_r, FI_r, R_r >$ et $< E_o, SI_o, FI_o, R_o >$ respectivement.

$$
\begin{array}{l|l}
E_r & \neg C(y,x,\alpha,t_0) \\
SI_r & C(y,x,\alpha,t_f) \\
FI_r & C(y,x,\neg\alpha,t_f) \\
R_r & C_j(x,y,\mathit{request}(x,y,\alpha),t_0) \\
 & C_j(y,x,\mathit{request}(x,y,\alpha) \Rightarrow \\
 & C_j(y,x,\mathit{promise}(y,x,\alpha)|\mathit{refuse}(y,x,\alpha),t_1),t_0) \\
 & C_j(y,x,\mathit{promise}(y,x,\alpha) \Rightarrow C(y,x,\alpha,t_2),t_0) \\
 & C_j(y,x,\mathit{refuse}(y,x,\alpha) \Rightarrow C(y,x,\neg\alpha,t_2),t_0)
\end{array}
$$

Fig. 4 – Conditions et règles du jeu $request$.

$$
\begin{array}{l|l}
E_o & \neg C(x,y,\alpha,t_0) \\
SI_o & C(x,y,\alpha,t_f) \\
FI_o & C(x,y,\neg\alpha,t_f) \\
R_o & C_j(x,y,\mathit{offer}(x,y,\alpha),t_0) \\
 & C_j(y,x,\mathit{offer}(x,y,\alpha) \Rightarrow \\
 & C_j(y,x,\mathit{accept}(y,x,\alpha)|\mathit{refuse}(y,x,\alpha),t_1),t_0) \\
 & C_j(x,y,\mathit{accept}(y,x,\alpha) \Rightarrow C(x,y,\alpha,t_2),t_0) \\
 & C_j(x,y,\mathit{refuse}(y,x,\alpha) \Rightarrow C(x,y,\neg\alpha,t_2),t_0)
\end{array}
$$

Fig. 5 – Conditions et règles du jeu $offer$

12.3 Exemple détaillé

La théorie de la cohérence cognitive présentée dans les sections précédentes doit être envisagée comme une sur-couche des architectures agents existantes. Son intégration sera réalisée via la reformulation des réseaux de cognitions en termes d'éléments et de contraintes de sorte que les différentes mesures de cohérence et d'utilité définies ci-dessus s'appliquent.

Dans notre exemple, un agent *Bob* travaille à planifier sa journée du lendemain. Il a le choix entre deux intentions : aller à la conférence ou rester travailler. À ces deux intentions

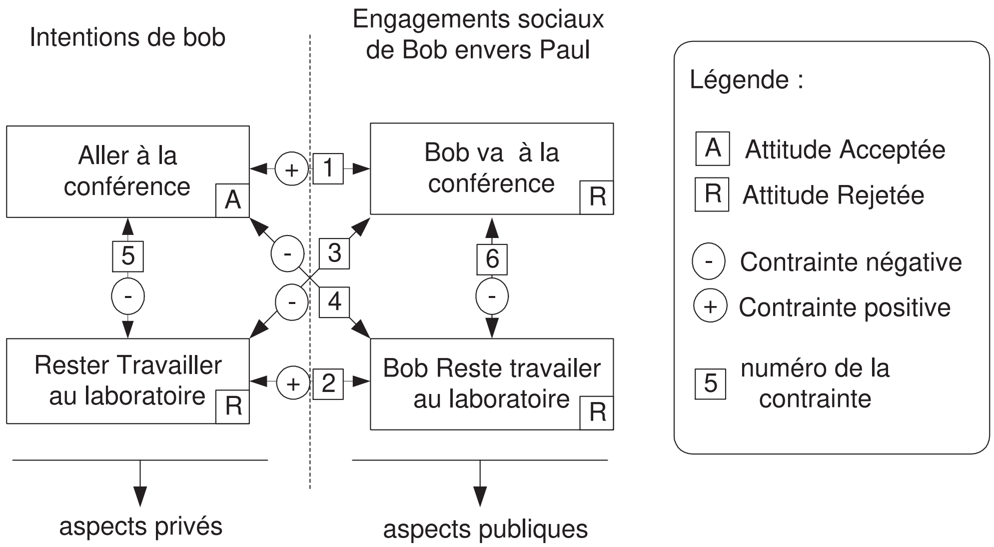

FIG. 6 – Réseau de contraintes dans la situation initiale pour l'agent *Bob*

peuvent correspondre les deux engagements sociaux suivants : l'engagement de *Bob* envers son patron *Paul* d'aller à la conférence et l'engagement de *Bob* envers *Paul* de rester travailler. Ces engagements sont liés aux intentions correspondantes par des relations de correspondance (ces liens ont été établis à la section 12.1) et sont liés entre eux par une relation d'exclusivité mutuelle. En outre, l'intention de rester travailler et l'engagement d'aller à la conférence ainsi que l'intention d'aller à la conférence et l'engagement de rester travailler sont liés par des relations d'incompatibilité.

On peut induire les contraintes correspondantes à toutes ces relations : aux relations de permission on fait correspondre des contraintes positives, aux relations d'exclusion mutuelle ou d'incompatibilité des contraintes négatives. En outre, *Bob* a initialement l'intention d'aller à la conférence, une intention qu'il a inférée de ses croyances et de ses désirs à la façon des systèmes BDI classiques. Cette intention appartient donc à l'ensemble des éléments acceptés A et tous les autres éléments à l'ensemble des éléments rejetés R (c'est-à-dire que *Bob* n'est engagé ni à aller à la plage ni à rester travailler et qu'il n'a pas l'intention de rester travailler). On utilise ce faisant, comme indiqué en section 10.3, une variante de l'hypothèse du monde clos qui consiste en : tout ce qui n'est pas accepté est rejeté. La figure 6 montre le réseau de contraintes à cet instant[29].

En affectant un poids unitaire à chacune des contraintes définies ci-dessus, on peut calculer la cohérence du réseau ainsi défini en additionnant les poids des contraintes satisfaites, cette somme étant pondérée par leur nombre : on obtient une cohérence de 3/6 (les contraintes numérotées 5, 3 et 2 sont satisfaites). De même, on peut calculer la cohérence de chacun des éléments de ce réseau. L'intention acceptée d'aller à la conférence a une cohérence de 2/3 (parmi les trois contraintes afférentes à cette cognition, seules la 4 et la 5 sont satisfaites).

[29]Notons que toutes les contraintes indiquées ne sont pas nécessaires et que le réseau décrit est redondant, mais ces considérations de représentation des connaissances n'altèrent en rien la validité du traitement qui s'en suit.

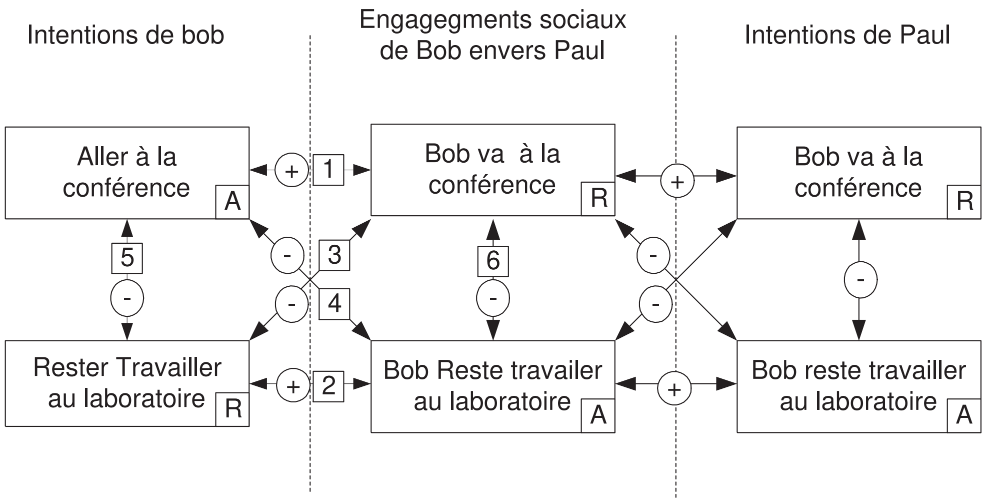

Fig. 7 – Réseau de contraintes dans la situation finale pour *Bob* et *Paul*

L'intention, rejetée, de rester travailler a une cohérence de 2/3. L'engagement rejeté d'aller à la conférence a une cohérence de 0/3 et celui, également rejeté, de rester travailler à une cohérence de 2/3.

Pour maximiser sa cohérence, *Bob* devrait donc modifier l'état d'acceptation de l'élément le moins cohérent, c'est à dire s'attaquer à réduire l'incohérence de la magnitude la plus importante. Il s'agirait donc pour lui de parvenir à faire accepter l'engagement d'aller à la conférence. Comme il s'agit d'un engagement social qui n'a de valeur que s'il est socialement acquis, *Bob* ne peut pas réduire cette dissonance seul. Cet engagement dissonant implique le patron de *Bob*, c'est donc une incohérence externe explicite pour *Bob*. *Bob* va donc entamer un dialogue avec son patron *Paul* au sujet de son intention d'aller à la conférence et plus précisément afin de faire accepter l'engagement correspondant. Pour s'en convaincre, on peut calculer l'espérance d'utilité des différents types de dialogues que l'agent peut entreprendre. Si l'agent *Bob* cherche par le dialogue à rester travailler et qu'il parvient à négocier cet engagement auprès de son patron, la cohérence de son réseau sera 2/6. S'il essaye de s'engager à aller à la conférence et qu'il y parvient, elle sera de 6/6 et finalement, si l'agent *Bob* se tait elle restera à 3/6.

Parmi les différents jeux de dialogue disponibles, *Bob* va choisir celui qui lui permet de faire accepter un engagement en action du locuteur envers son interlocuteur. Ce jeu de dialogue devra donc avoir une condition d'entrée qui peut s'unifier avec $\neg C(Bob, Paul, AllerConference)$ signifiant que l'engagement considéré n'est pas déjà établi et une condition de succès qui peut s'unifier avec $C(Bob, Paul, AllerConference)$ signifiant qu'en cas de succès du dialogue, l'engagement correspondant sera accepté. Seul le jeu d'offre satisfait ces contraintes. *Bob* va donc proposer à son patron *Paul* d'entrer dans un jeu de ce type. Sous l'hypothèse de coopération dialogique, on suppose que *Paul* accepte d'entrer dans ce jeu. Le jeu DIAGAL $offer$ va alors être joué et *Bob* va respecter son engagement de faire une offre à *Paul*, en lui proposant l'engagement d'aller à la conférence. C'est ensuite à *Paul* de réagir.

Pour éviter de perdre de la cohérence, *Paul* devra refuser l'engagement proposé par *Bob* (si cet engagement était accepté, la cohérence de *Paul* passe de 3/6 à 1/6 comme on peut le calculer grâce à la figure 7 dans laquelle l'engagement de *Bob* envers *Paul* ne serait pas accepté). Mais pour maximiser sa cohérence, il a mieux à faire avant. En effet, l'élément le plus dissonant pour *Paul* est l'engagement de *Bob* envers lui de rester au laboratoire travailler qui est rejeté (cohérence de 0/3). Pour le faire accepter, *Paul* engage un sous-dialogue en proposant le jeu qui permet d'établir un type d'engagement unifiable avec ce dernier soit le jeu de requête. *Bob* est dialogiquement coopératif, il accepte de jouer ce sous-jeu. *Paul* demande alors à *Bob* d'accepter l'engagement $C(Bob, Paul, ResterTravailler)$. Le fait que *Paul* soit un supérieur hiérarchique de *Bob* empêche ce dernier de refuser (ce sans quoi, on serait entré dans une argumentation et notre exemple aurait été plus volumineux). La condition de succès du jeu emboîté est donc remplie et sa fermeture est proposée par *Paul*. *Bob* accepte. *Paul*, dont la cohérence est maintenant maximale, c'est à dire de 6/6 comme indiqué par la figure 6, ne souhaite pas changer quoi que ce soit, il refuse donc l'offre issue du jeu d'offre ouvert par *Bob*. Les deux jeux de dialogue sont fermés.

Nous sommes alors dans la position indiquée par la figure 7. La cohérence de *Bob* a finalement chuté par ce dialogue (2/6). Pour rétablir cette cohérence, *Bob* n'a plus la possibilité d'essayer de négocier l'acceptation ou le rejet des engagements envers *Paul*, puisque ceux-ci ont déjà été discutés et qu'on ne peut pas revenir dessus (dans le système de Maudet et Chaib-draa). Ainsi, la seule solution pour *Bob*, lui permettant de rétablir la cohérence est de changer ses cognitions. *Bob* va donc rejeter son intention première d'aller à la conférence et accepter l'intention de rester travailler.

C'est ainsi que le changement d'attitude est envisagé dans notre théorie, conformément aux considérations des psychologues. Tant qu'un agent peut changer le monde extérieur pour maximiser sa cohérence, il essaye et dès lors que le monde extérieur n'est plus modifiable, il change ses attitudes pour s'y conformer.

Finalement, nos deux agents ont tenu une conversation à laquelle pourrait correspondre le dialogue suivant (même si les phases d'établissement, d'acceptation et de mise en commun sont généralement plus implicites dans le langage naturel) :

> *Bob : Je peux te faire une offre ?* <proposition d'entrer dans un jeu d'offre (le jeu DIAGAL $offer$)>
>
> *Paul : Oui, vas-y.* <proposition acceptée>
>
> *Bob : Je peux aller à la conférence ?* <offre>
>
> *Paul : Justement, est-ce que je peux te faire une requête ?* <proposition d'imbriquer un jeu de requête (le jeu DIAGAL $request$)>
>
> *Bob : Oui.* <proposition acceptée>
>
> *Paul : Demain, tu restes travailler au laboratoire. Tu n'iras donc pas à la conférence.* <requête et refus de l'offre>
>
> *Bob : OK (soupir).* <requête acceptée>
>
> *Paul : C'est bon ?* <proposition de fermer les jeux ouverts>
>
> *Bob : Oui, oui (long soupir).* <proposition acceptée>

Quatrième partie

Plan de travail

13 Tâches à completer

13.1 Tâche 0 : Formalisation

Cette étape est déjà bien entamée au moment de la rédaction de cette proposition, cependant elle sera probablement raffinée dans les mois à venir. Elle consiste à préciser de manière non-ambiguë et en la présentant formellement la théorie proposée. On pense pouvoir arriver à des schémas d'algorithmes relativement précis.

13.2 Tâche 1 : Validation théorique

Cette étape consiste à demander des retours des communautés scientifique intéressées, en l'occurrence :

- *la psychologie sociale* : on devrait consulter Shultz de l'université McGill et Mark Lepper de l'université Stanford ainsi que les psychologues de l'université Laval susceptibles d'être concernés par les extensions proposées à la théorie de la dissonance cognitive.
- *la philosophie computationnelle, la philosophie du langage et la philosophie de l'esprit* : le minimum serait de contacter Paul Thagard qui dirige le département de sciences cognitives de l'université de Waterloo pour l'informer de l'usage que l'on fait de sa théorie et des extensions que l'on propose. Sinon, certains philosophes du langage comme Daniel Vanderveken de l'université du Québec à Trois-Rivières pourront nous donner des retours lors des interactions au sein de l'équipe de travail FQRNT.
- *la linguistique* : même si, malgré l'utilisation de nombreux résultats de dialectique, on ne fait aucun emprunt technique précis à la linguistique, le terrain sur lequel est bâtie notre théorie en est un de linguiste. Il serait dommage et peu sérieux d'ignorer ce « background » fondateur du domaine. On devrait donc essayer d'interagir avec la communauté de linguistique et ce, dés les prochains mois. On pense, par exemple, soumettre un papier pour une conférence de l'ACL[Association for Computational Linguistic]. Comme on sait que les propositions concernant la pragmatique sont relativement rares, l'approche que nous développons pourrait intéresser les chercheurs de cette communauté.
- *l'informatique* : les retours de la communauté SMA sont également importants à considérer. Ces retours seront obtenus via les évaluations des différentes publications soumises et les différents contacts que nous avons dans ce domaine. Notons que les idées proposées ont d'ores et déjà été récompensées du prix de la meilleure contribution du collège SMA de l'AFIA [Association Française pour l'Intelligence Artificielle] pour l'année 2002 lors de l'édition des Journées Francophones d'Intelligence Artificielle Distribuée et Systèmes Multi-Agents de cette même année [Pasquier and Chaib-draa, 2002b]. Nous avons fait d'autres soumissions pour lesquelles nous sommes actuellement en attente de réponse [Pasquier and Chaib-draa, 2002a].

13.3 Tache 2 : Validation informatique

Même si nous avons déjà implémenté une version préliminaire de la théorie dans le cadre des travaux pratiques du cours de SMA IFT-64881A, il faudra valider la théorie dans une application ou un système de plus grande envergure. Les outils utilisés pour cette validations seront des formalismes hybrides symbolique-connexioniste comme ceux proposés par Sun [Sun and Browne, 2001, Sun and Browne, 1999]. Il y a là un enjeu supplémentaire car ce type de formalisme n'est pas encore utilisé dans la communauté SMA. Ils ont pourtant de nombreux avantages et semblent permettre l'implémentation simple de formalismes dont la puissance expressive n'a rien à envier aux logiques modales les plus complexes.

Pour mener à bien cette tâche de validation informatique, il s'agit de trouver un domaine d'instanciation ancré dans la réalité des SMAs (qui ne soit pas trop simple, pour ne pas faire jouet). Valider le modèle dans ce domaine et comparer les résultats obtenus à ceux d'autres approches. Par exemple, on pourrait instancier notre prototype sur des exemples de dialogue dans un des domaines suivants : partage de ressources entre agents, commerce électronique, prise de rendez-vous, ... Cette phase comportera une étape d'implémentation pour laquelle on pourra faire appel à une collaboration, car le développement d'un système cognitif complet est une tâche fastidieuse. N. Andrillon, un étudiant français qui effectuera son stage de dernière année d'école d'ingénieur (ENSEIHT, Toulouse) au laboratoire DAMAS à la session d'hiver 2003 manifeste son intérêt en ce sens.

Notons que les tâches 0 et 2 ne sont pas indépendantes car la validation informatique se fera sur la base de la formalisation proposée.

13.4 Tâche 3 : Rédaction

En parallèle des soumissions d'articles pour des conférences ou des journaux, il faudra rédiger la thèse, la déposer, la soutenir, puis la valoriser.

14 Collaborations

Nous avons également un certains nombre d'idées de collaborations :

- Avec Brahim Chaib-draa, Nicolas Maudet et Marc-Andrée Labrie, nous continuerons le développement de DIAGAL, un langage de communication agent basé sur les jeux de dialogue déjà mentionné dans cette proposition. C'est avec ce cadre interactionnel que nous souhaitons réaliser notre validation. Une implémentation de DIAGAL ainsi qu'un simulateur de dialogues sont déjà opérationnels. Les avancées en ce domaine donneront lieu à des publications ;

- Dans un avenir proche (en fait dès que cette proposition sera achevée), nous souhaiterions entrer en contact avec des chercheurs en psychologie (certains comme Manuel Zacklad de l'université de Troye qui a une formation initiale de psychologue ont manifesté leur grand intérêt). En particulier, des spécialistes de la théorie de la dissonance cognitive et de sa formalisation avec qui nous souhaiterions interagir sur les extensions de la dissonance cognitive proposées (comme indiqué dans le détail de la tâche 2) ;

- Nous sommes resté en contact avec l'équipe SMA du DCSD[Département Conduite, Systèmes et Décision] de l'ONERA[Office National d'Étude et de Recherche en Aérospa-

Session	Tâches à réaliser	Durée
Automne 2002	Tâches 0,1	4 mois
Hiver 2003	Tâches 0,1,2	4 mois
Eté 2003	Tâches 1,2	4 mois
Autonne 2003	Tâche 2,3	4 mois
Hiver 2004	Tâche 3	4 mois

FIG. 8 – Répartition des tâches à réaliser dans les prochaines sessions.

tiale] où nous avons travaillé avec Catherine Tessier, Laurent Chaudron, François Legras et Frédéric Dehais sur la notion de conflit (un concept voisin de ceux de dissonance et d'incohérence).

- Certains chercheurs, intéressés par notre approche, pourront susciter des collaborations : peut-être un projet de publication commune avec Humbert Fiorino de l'équipe MAGMA du laboratoire LEIBNIZ de l'institut IMAG[institut d'Informatique et de Mathématiques Appliqués de Grenoble] de Grenoble ;
- Nous ferons aussi lire notre travail à différents chercheurs qui ont des approches compatibles, on pense à Jean-Louis Dessales de l'ENST à Paris ;
- On compte également sur le groupe de travail FQRNT avec Bernard Moulin, Daniel Vanderveken et Brahim Chaib-draa ;
- Parmi les étudiants du département d'informatique de l'université Laval, Jamal Benthahar qui travaille sur l'argumentation dans les systèmes multi-agents nous semble être le plus proche de notre problématique.

15 Calendrier/échéancier

Le tableau 8 donne la répartition prévisionnelle des différentes tâches envisagées dans le temps.

16 Schéma de la thèse

Le plan de la thèse pourrait comporter les aspects suivants :

1. Introduction ;
2. État de l'art et contexte des communications agents ;
3. Problématique, motivation et objectifs ;
4. Proposition d'une approche par les théories de la cohérence et état de l'art en ce domaine ;
5. Présentation de notre approche basée sur la cohérence cognitive ;

6. Application à la communication automatique avec les jeux de dialogue de DIAGAL[DIAlogue Game based Agent communication Language] ;
7. Validation théorique et informatique (à l'aide d'exemples empruntés à la réalité) ;
8. Résultats, comparaisons, limites et discussions ;
9. Conclusion et travaux futurs.

17 Conclusion générale

La communication entre agents est un point crucial des SMAs, dès lors que ceux-ci ne sont composés que d'agents et de leurs interactions. Dans cette proposition de thèse, nous avons délimité, grâce à une revue de la littérature fournie (partie I), un certain nombre de problématiques concernant les conversations entre agents. Trop occupée à définir un cadre interactionnel standard, la communauté SMA a quelque peu délaissé les aspects cognitifs liés à la communication, laissant ainsi au concepteur la majeure partie du travail. En conséquence, notre objectif principal est de fournir une pragmatique, au sens d'une théorie de l'usage, qui soit adaptée aux cadres interactionnels agent que proposent les approches conventionnelles et en particulier les jeux de dialogue (cet objectif est raffiné en partie II).

En effet, la pragmatique des cadres interactionnels des SMAs est un aspect souvent négligé. Soulignons que ce n'est pas complètement un hasard ou une négligence puisque cet aspect est réputé difficile, et ce, dans tous les domaines d'étude de la communication (linguistique, philosophie du langage, théories de la communication, intelligence artificielle, ...).

Nous avons ensuite proposé une solution simple et élégante qui nous semble remplir correctement l'objectif sus-nommé : la théorie de la cohérence cognitive (partie III). Cette théorie, que nous proposons pour la pragmatique des communications agents, est épistémologiquement *fondée* sur les travaux de psychologie cognitive, de psychologie sociale et de philosophie computationnelle de l'esprit. Elle arbore la forme, classique en informatique, de raisonnements et de calculs sur des éléments et leurs contraintes associées. Ce faisant, elle constitue une théorie de l'utilisation des cadres interactionnels conventionnels[30] autant qu'une théorie des effets des communications (changement d'attitude). Elle est envisagée comme une sur-couche des architectures d'agents cognitifs classiques. Les mesures d'incohérence et d'utilité définies dans le cadre de cohérence proposé fournissent les mécanismes nécessaires pour répondre même partiellement aux problèmes évoqués ci-dessus et qui sont autant de sujets peu traités dans la littérature SMA :

- Quand un agent prend-il l'initiative d'une conversation, à quel sujet et pourquoi (voir section 11.3) ?
- Avec qui (voir section 11.3) ?
- Par quel type de dialogue (voir section 11.4 et plus spécifiquement, pour les jeux de dialogue, la section 12) ?
- Quelle intensité donner aux forces illocutoires des actes de langage utilisés (voir section 11.7) ?
- Comment définir et mesurer l'utilité d'une conversation (voir section 11.6.1) ?
- Quand arrêter le dialogue ou le cas échéant comment le poursuivre (voir section 11.6.2) ?
- Quels sont les impacts du dialogue sur les attitudes de l'agent (voir section 10.4) ?
- Quels sont les impacts du dialogue sur l'humeur de l'agent (voir section 11.7) ?

[30] Les cadres interactionnels considérés sont (1) le cadre interactionnel théorique défini par les types de dialogues isolés par Walton et Krabbe [Walton and Krabbe, 1995] et (2) les jeux de dialogue : en particulier, mais de manière non restrictive, ceux du langage DIAGAL [Chaib-draa et al., 2002].

– Quelles sont les conséquences du dialogue sur les accointances de l'agent (voir section 11.6.3) ?

On a aussi montré comment le cadre de consonance cognitive proposé capture naturellement la dimension commune des conversations vues comme des projets communs de réduction de dissonance (voir section 11.5).

Évidemment, chacune de ces problématiques n'a pu être approfondie ici, mais le but de cette proposition était plutôt de donner une idée d'ensemble de cette nouvelle approche de la communication agent. En effet, l'un des intérêts de notre apport est la large couverture de la théorie proposée pour l'automatisation de la communication agent. Dans le cheminement de notre thèse, de futures publications et travaux de validation répartis selon un échéancier précis (partie IV) seront plus spécifiquement dédiées à certains aspects de la théorie ébauchée ici.

Selon Lochbaum [Lochbaum, 1994] : « les interlocuteurs s'engagent dans un dialogue parce qu'ils ont une raison pour cela. ». Notre approche propose un fondement pour les communications agents en identifiant de manière générique qu'elle pourrait être cette raison : une incohérence cognitive. C'est de ce fondement que l'on développe notre théorie de l'usage des cadres interactionnels agents. À notre problématique, on répond donc par une solution théorique générique, simple et élégante, reposant sur des bases théoriques et empiriques solides (philosophie de l'esprit, psychologie sociale, psychologie cognitive, socio-linguistique et intelligence artificielle). C'est donc une *contribution aux sciences cognitives.*

Finalement, le cadre de cohérence proposé fournit un système de valeur aussi bien individuel que collectif qui nous semble être une étape nécessaire pour développer une plus grande autonomie des agents : le concepteur aura moins à s'occuper de l'utilité et de l'efficacité des comportements des agents dès lors qu'ils disposeront d'outils pour la mesurer eux-mêmes. C'est donc aussi un premier pas vers l'automatisation des communications agents et en cela une *contribution à l'informatique* en général et aux aspects théorique et pratiques des SMAs en particulier.

Cinquième partie
Annexes

A L'énoncé

A.1 théories des actes de langages : « Quand dire c'est faire. »

Historiquement, les unités conversationnelles ont été étudiées avant le cadre conversationnel dans lequel elles doivent prendre place. La théorie des actes de langage (et ses nombreuses variantes) est en ce domaine la référence incontournable. La théorie des actes de langage [Austin, 1962a, Grice, 1957, Searle, 1969, Wittgenstein, 1953] est issue de la philosophie du langage. Cette théorie rend compte du processus de communication entre agents cognitifs au sein d'une société : le langage par l'entremise de la langue. Initialement pensée pour le langage naturel, sa nature formelle la rend utilisable pour les modèles computationnels. L'idée maîtresse de cette théorie est qu'une instance d'utilisation de la langue est une action comme les autres : « dire c'est faire »[31]. Pour chaque acte de langage « primitif », on distingue quatre composantes qui peuvent être vues comme quatre actes :

- *énonciation* : le locuteur fournit l'énoncé dans le contexte par transmission ou prononciation du message ; c'est le niveau physique.
- *acte locutoire ou locution* : l'interlocuteur (ou les interlocuteurs, le cas échéant) a perçu l'énonciation. Il lui faut interpréter le sens de l'énoncé en terme de contenu propositionnel. Si le contenu propositionnel interprété est celui que le locuteur voulait transmettre, on dira que l'aspect locutoire de l'acte est accompli avec succès. Par exemple : « il pleut », « it's raining » et « es regnet » sont trois énoncés différents qui correspondent à un seul acte locutoire de contenu propositionnel : il pleut. Dans la suite, on notera p ce contenu propositionnel qui correspond à ce qui est dit.
- *acte illocutoire ou illocution* : cet acte traduit les intentions du locuteur envers son (ses) interlocuteur(s). Une fois que l'interlocuteur a perçu et interprété le sens propositionnel de l'énoncé, il doit inférer ce que le locuteur a voulu exprimer par cet énoncé. « On va être riche » peut être interprétée selon le contexte comme une information, une prédiction, une promesse ou une blague ironique et cela change tout. Si l'interlocuteur saisit le sens que le locuteur a voulu donner à son énoncé on dira que l'acte illocutoire a réussi.
- *acte perlocutoire ou perlocution* : un tel acte porte sur les effets du message sur le destinataire : action, modification de croyance, modification de ses attitudes propositionnelles (AP). L'effet perlocutoire concerne la réaction du destinataire, les effets de son interprétation sémantique du message.

Le terme « acte de langage » est souvent employé pour désigner un acte illocutoire. Notons qu'il existe différentes variantes de cette théorie certaines étant « inférentielles » [Bach and Harnish, 1979] et d'autres plus « analytiques » [Searle, 1979].

[31] Selon le titre de l'édition française [Austin, 1962b] de l'ouvrage de J.L. Austin [Austin, 1962a].

A.1.1 Force illocutoire et contenu propositionnel

Dans leurs analyses des différents types syntaxiques de phrases du langage naturel, certains linguistes ont précisé les aspects illocutoires. Les actes illocutoires (qui sont la transmission du sens d'une phrase dans un contexte donné) consistent en l'application d'une force illocutoire F sur un contenu propositionnel p, ce que l'on notera $F(p)$ dans le reste de ce texte. Par exemple, les énoncés « allons-nous en ! » et « nous allons partir » ont le même contenu propositionnel (le locuteur et son entourage vont s'en aller) mais des forces illocutoires différentes (respectivement une force illocutoire d'ordre et une force illocutoire d'assertion pour le futur). D'après Searle et Vanderveken [Searle and Vanderveken, 1985], chaque force illocutoire peut être divisée en six composantes : un but illocutoire, un mode d'accomplissement de ce but, des contraintes sur le contenu propositionnel, des conditions préparatoires, des conditions de sincérité et un degré d'intensité de ces conditions de sincérité. Détaillons cela.

Le but illocutoire

Le but illocutoire relie la proposition énoncée au monde réel. Pour Searle, il existe cinq utilisations du langage possible qui sont caractérisées par les cinq buts illocutoires (on parle aussi des cinq types d'actes de langage) :

1. *assertif / représentatif* : le locuteur exprime un contenu propositionnel qui se réfère au monde passé, actuel ou futur tel qu'il se le représente. Exemples d'actes illocutoires assertifs : affirmation, assertion, conjecture, rappel, accusation, témoignage, prédiction, ...

2. *directif* : le locuteur donne une directive représentée par le contenu propositionnel au(x) destinataire(s). Exemples d'actes illocutoires directifs : ordre, demande, prière, invitation, conseil, recommandation, ...

3. *comissif / promissif / engageant* : le locuteur s'engage (vis-à-vis du destinataire) à accomplir l'action représentée par le contenu propositionnel. Exemples d'actes illocutoires comitifs : promesse, menace, renonciation, acceptation, voeu, serment, ...

4. *expressif* : le contenu propositionnel concerne l'humeur mentale et l'affect du locuteur. Exemples d'actes illocutoires expressifs : déclarations d'amour, félicitations, remerciement, insulte, ...

5. *déclaratif* : le locuteur accomplit l'action représentée par le contenu propositionnel du simple fait de sa locution. Exemples d'actes illocutoires déclaratifs : excommunication, nomination, ratification, leg, ajournement, bénédiction, ...

Le but illocutoire est la principale composante de la force illocutoire car il indique le lien du contenu propositionnel avec le monde. Un locuteur qui accomplit un acte illocutoire peut avoir toutes sortes de buts perlocutoires et d'autres intentions. Par exemple, il peut vouloir amuser, convaincre, embarrasser ou choquer. Mais dans tous les cas, il a au moins l'intention d'accomplir le but illocutoire de son acte concernant son contenu propositionnel. Une des justifications de la complétude de cette classification est que les cinq buts illocutoires couvrent les différentes directions d'ajustement possibles entre l'utilisation de la langue et le monde. En effet, d'un point de vue logique, il n'y a que quatre directions d'ajustement possibles pour un acte de langage :

1. *la direction d'ajustement des mots aux choses* : l'énoncé offre une représentation des choses (peut-être fausse). En cas de satisfaction d'un acte de langage ayant cette direction d'ajustement, le contenu propositionnel correspond à un état de choses existant indépendamment de l'énonciation dans le monde. Par exemple, les actes de langage ayant un but illocutoire assertif ont la direction d'ajustement des mots aux choses. En effet, ils ont pour but de représenter comment les choses sont dans le monde.

2. *la direction d'ajustement des choses aux mots* : l'énoncé propose un processus de changement des choses. En cas de satisfaction d'un acte de langage ayant cette direction d'ajustement, le monde est transformé de façon à satisfaire son contenu propositionnel. Les actes de langage ayant un but illocutoire directif ou comitif ont la direction d'ajustement des choses aux mots. Leur but est que le monde soit transformé (respectivement par l'interlocuteur ou le locuteur) de sorte qu'ils correspondent à leur contenu propositionnel.

3. *la double direction d'ajustement* : l'énoncé est un changement des choses. En cas de satisfaction de l'acte illocutoire ayant la double direction d'ajustement, le monde s'ajuste au contenu propositionnel et cet ajustement consiste en l'énonciation elle-même. Les actes de langage dont le but illocutoire est déclaratif ont la double direction d'ajustement.

4. *la direction d'ajustement vide* : l'énoncé et les choses sont indépendants. Les actes de langage dont le but illocutoire est expressif ont la direction d'ajustement vide. Ce type d'acte de langage est toujours satisfait puisqu'il n'a pas affaire aux choses représentées mais à l'état mental du locuteur.

Au niveau linguistique, la théorie des actes de langages classique prévoit qu'il est possible de déterminer le type d'un acte à partir de sa forme linguistique. Les marqueurs lexicaux ou syntaxiques doivent permettre cette opération. C'est ce que l'on appelle « l'hypothèse de force littérale ». La réalité n'est pas si simple. Seule l'utilisation « d'expressions performatives explicites » de la forme « je te <verbe performatif> que … » donne le type de l'acte réalisé sans ambiguïté, encore que l'on puisse trouver des cas discutables. Dans les dialogues réels, de nombreux actes de langage ne sont pas du type indiqué par leur forme linguistique (que nous nommerons forme littérale ou acte de surface). C'est ce que l'on appelle les actes de langages indirects. Par exemple, l'énoncé S_1 suivant s'interprète en situation comme une requête à laquelle on répond en donnant une salière et non comme une question fermée (à laquelle on répond par oui ou non).

S_1 *: Peux-tu me passer le sel ?*

Bien des difficultés subsistent quant à l'étude de ce type d'actes, et ce, même si les travaux de Grice [Grice, 1969] sur la notion d'implicature fournissent des indications sur le type d'indirection réalisée pour passer de la question littérale à la requête indirecte.

Le mode d'accomplissement

La plupart des buts illocutoires peuvent être atteints de différentes façons. Ainsi, un acte de langage directif peut être réalisé de différentes manières : autoritaire, douce, supplicative, ... Le mode d'accomplissement spécifie comment le but illocutoire doit être atteint. En français, ceci est exprimé par des adverbes comme : obligatoirement, éventuellement, peut-être, ...

Les contraintes sur le contenu propositionnel

Certaines forces illocutoires imposent des conditions sur l'ensemble des contenus propositionnels qui pourraient leur être associés. Ainsi, les actes de langage dont le but illocutoire est comissif ou directif doivent avoir un contenu propositionnel représentant des actions futures respectivement du locuteur ou des interlocuteurs. Par exemple, lorsque le locuteur énonce une promesse, celle-ci doit avoir pour contenu propositionnel une action future du locuteur : « je ne le ferai plus, c'est promis ! ».

Les conditions préparatoires

Lors de la performance d'un acte illocutoire, le locuteur a généralement des croyances sur le contexte de son énonciation. Par exemple, un locuteur qui donne un conseil croit généralement que cela peut aider l'autre partie. Les conditions préparatoires d'une force illocutoire déterminent quelles doivent être les croyances du locuteur pour qu'il puisse accomplir un acte de langage ayant cette force.

Les conditions de sincérité

Dans l'accomplissement d'un acte illocutoire, le locuteur transmet un contenu propositionnel avec un but illocutoire, mais aussi des informations concernant ses états mentaux. Cela signifie que si la communication est une « extérieurisation », elle renseigne aussi sur l'état intérieur du locuteur. Par exemple, lorsqu'un locuteur énonce une demande, il a un but illocutoire directif auquel peut être associé un désir, un regret, une inquiétude, ... De tels états mentaux sont des attitudes propositionnelles $m(p)$ où :

- p est le contenu propositionnel
- m est un mode psychologique (croire, espérer, désirer, regretter, ...)

Comme tout locuteur peut mentir et transmettre des états mentaux qui ne sont pas réellement les siens, on peut distinguer les actes illocutoires sincères (le locuteur a les états mentaux qu'il exprime) des insincères. Les conditions de sincérité indiquent quels états mentaux devraient être présents chez le locuteur lorsqu'il produit un acte illocutoire sincère. Par exemple, une condition de sincérité des actes illocutoires assertifs est que le locuteur doit croire le contenu propositionnel. De même, une condition de sincérité des actes illocutoires directifs est que le locuteur doit désirer que le contenu propositionnel soit accompli. Ou encore, une condition de sincérité des actes illocutoires comissifs est que le locuteur doit avoir l'intention de réaliser le contenu propositionnel.

Le degré d'intensité

Évidemment, les attitudes propositionnelles associées aux conditions de sincérité le sont avec une intensité qui dépend de la force illocutoire. Ainsi, une supplication dénote un désir plus grand qu'une simple demande. L'intensité d'une force illocutoire est donc indiquée par un degré.

Définition récursive de la force illocutoire

On distingue cinq forces illocutoires primitives qui sont les plus simples possible :

1. *La force illocutoire primitive d'assertion* : cette force possède un but assertif, un mode d'accomplissement, un degré d'intensité et des conditions sur le contenu propositionnel neutres avec comme condition préparatoire que le locuteur ait des justifications ou des raisons de croire en la vérité du contenu propositionnel et comme condition de sincérité que le locuteur croit le contenu propositionnel.

2. *La force illocutoire primitive d'engagement* : cette force a un but comissif, un mode d'accomplissement et un degré d'intensité neutres avec comme condition sur le contenu propositionnel qu'il représente une action future du locuteur, la condition préparatoire que le locuteur soit capable d'accomplir cette tâche et la condition de sincérité qu'il ait l'intention de l'accomplir.

3. *La force illocutoire primitive directive* : cette force est définie par un but directif, un mode d'accomplissement et un degré d'intensité neutres avec la condition que le contenu propositionnel représente une action future de l'interlocuteur, la condition préparatoire que l'interlocuteur soit capable d'accomplir cette action et la condition de sincérité que le locuteur désire que l'interlocuteur accomplisse cette action.

4. *La force illocutoire primitive de déclaration* : cette force possède un but déclaratif, un mode d'accomplissement et un degré de puissance neutres avec la condition que le contenu propositionnel représente une action présente du locuteur, la condition préparatoire qu'il soit capable de l'accomplir (avec preuves) et la condition de sincérité qu'il en ait l'intention.

5. *La force illocutoire primitive expressive*[32] : cette force a tous ses attributs (but expressif, mode d'accomplissement, condition sur le contenu propositionnel, condition préparatoire, condition de sincérité et degré de puissance) neutres.

Les forces illocutoires non primitives sont alors obtenues par application des opérations suivantes : ajout de conditions sur le contenu propositionnel, ajout de conditions préparatoires, ajout de conditions de sincérité, restriction du mode d'accomplissement et finalement modulation du degré d'intensité.

[32] Elle est réalisée syntaxiquement par des énoncés exclamatifs en français. Comme il n'existe pas de verbe performatif simple nommant cette force, il n'y a pas d'énoncé exclamatif exprimant seulement la force expressive primitive

Actes illocutoires complexes

Il est important de noter que la notion de force illocutoire est itérable. On peut par exemple affirmer qu'une demande a été faite, proposer une demande ou demander une proposition, etc. Tous ces actes prennent la forme $F(F(p))$ ou plus généralement $F(...F(p)...)$ et il n'y a pas de limitations à ce type d'imbrication. En outre, certains actes illocutoires utilisant des connecteurs comme « et » ou « mais » ne sont pas de la forme $F(p)$. Il s'agit d'actes illocutoires dits complexes qui sont généralement de la forme $F_1(p_1) \wedge F_2(p_2)$. Par exemple, A est à la fois une assertion et une question.

A : Il est six heures, non ?

Un certain nombre de connecteurs illocutoires ont été proposés par Searle et Vanderveken dans le cadre de leur logique illocutoire [Searle and Vanderveken, 1985].

A.1.2 Sémantique des actes de langage

Les conditions de succès des actes illocutoires

Les conditions de succès d'un acte illocutoire sont l'ensemble des conditions qui doivent être réunies dans le contexte de l'énonciation pour que le locuteur réussisse à accomplir cet acte. Les conditions de succès d'un acte de langage sont déterminées de façon univoque par la force illocutoire et le contenu propositionnel de l'acte de langage. Un acte illocutoire $F(p)$ est accompli avec succès si et seulement si le locuteur réalise le but illocutoire de la force F sur la proposition p avec le mode d'accomplissement, les conditions préparatoires, les conditions de sincérité, le degré de puissance de F et que p vérifie les conditions sur le contenu propositionnel de F. Par exemple, la condition de succès de la requête S_2 est que le locuteur doit espérer que le récepteur accomplisse les actions représentées par le contenu propositionnel à savoir manger sa soupe.

S_2 : Mange ta soupe !

De même, la condition de succès d'une promesse est que le locuteur doit être prêt à s'engager à accomplir les actions représentées par le contenu propositionnel dans le futur.

Les conditions de satisfaction des actes illocutoires

La plupart des actes de langage concernent le monde via leur contenu propositionnel. Même si un acte illocutoire est accompli avec ses conditions de succès remplies, il peut ne pas être satisfait quant à son rapport au monde. Un acte de langage est satisfait si son contenu propositionnel est rendu vrai selon la direction d'ajustement au monde propre à son but illocutoire. Les conditions de satisfaction d'un acte illocutoire sont les conditions qui doivent être réunies pour que l'acte de langage soit entièrement satisfait (c'est-à-dire que chacune de ses trois composantes soit satisfaite). Par exemple, la condition de satisfaction d'une promesse est que le locuteur accomplisse ce qu'il a promis. De même, une demande (comme S_2) n'est satisfaite que si l'interlocuteur accomplit les actions qui sont représentées par le contenu propositionnel[33].

[33] Dans la théorie des actes de langage la notion de satisfaction est une généralisation de la notion de vérité.

Les conditions de succès et de satisfaction font partie de la sémantique des actes de langages. Pour analyser la forme logique des actes illocutoires, l'idée est de définir récursivement à la fois leurs conditions de succès et de satisfaction et les conditions de vérité de leur contenu propositionnel.

La logique illocutoire

Développée par Vanderveken et Searle [Searle and Vanderveken, 1985] dans les années 80, cette logique rend compte des relations d'implications entre actes de langage. Les quatre relations linguistiquement significatives sont :

1. $si\ success(act1)\ alors\ success(act2)\ et\ success(act3)\dots$
   Certaines conditions de succès d'actes sont plus fortes que d'autres. Par exemple, on ne peut demander à quelqu'un d'apporter un papier et un stylo sans lui demander d'apporter un stylo.

2. $si\ satisfaction(act1)\ alors\ satisfaction(act2)\ et\ satisfaction(act3)\dots$
   Certaines conditions de satisfaction d'actes sont plus fortes que d'autres. Par exemple, si on a obtenu le papier et le stylo, alors on a le stylo.

3. $si\ satisfaction(act1)\ alors\ success(act2)\ et\ success(act3)\dots$
   Certaines conditions de satisfaction d'actes sont plus fortes que les conditions de succès d'autres actes.

4. $si\ success(act1)\ alors\ satisfaction(act2)\ et\ satisfaction(act3)\dots$
   Certaines conditions de succès d'actes sont plus fortes que les conditions de satisfaction d'autres actes. Par exemple, si une déclaration est réussie, l'assertion de son contenu propositionnel est vraie : le succès de la déclaration par un maire « Je déclare Manu et Sophie mariés. » entraîne la satisfaction de l'assertion « Manu et Sophie sont mariés. »

La logique illocutoire précise qu'un acte de langage satisfait est toujours réussi. Le succès des actes illocutoires complexes de la forme $F_1(p_1) \wedge F_2(p_2)$ dépend pour sa part du succès des sous-actes qui les composent.

Sémantique formelle, basée sur le calcul des situations

Le calcul des situations est une des plus simples solutions au problème de la représentation logique des actions et de leurs effets sur le monde. Son intérêt est de pouvoir représenter les changements d'état tout en gardant une mémoire des états passés. C'est donc un formalisme de description des changements en logique du premier ordre dans lequel les termes représentent des états du monde associés à des situations. On passe d'une situation à une autre par l'entremise d'actions. Concrètement, on ajoute simplement à toutes les relations ou propriétés susceptibles de varier dans le temps un argument supplémentaire de type situation. Beaucoup de travaux ont étendu et réglé un certain nombre de problèmes de ce formalisme [Reiter, 1996]. Le calcul des situations a été utilisé par Chaib-draa et Vanderveken [Chaib-draa and Vanderveken, 1998] comme base pour définir une sémantique des actes de langage. Il y est étendu pour pouvoir représenter les états mentaux des agents. Initialement, le calcul des situations était un formalisme logique du premier ordre, cette propriété est perdue dans ce type d'extension.

Pour chaque but illocutoire, les degrés d'intensité des actes illocutoires sont quantifiés avec des entiers relatifs. Par exemple, les différents degrés pour des messages du type $< directive, p >$ pourraient être [Vanderveken, 1990] :

$$degree(order) = degree(command) = degree(require) = +1$$
$$degree(ask) = degree(request) = degree(tell.to) = 0$$
$$degree(suggest) = degree(advice) = degree(recommend) = -1$$

Les conditions de succès et de satisfactions sont définies (via leurs prédicats respectifs) pour chaque force illocutoire primitive. Les actes illocutoires sont définis récursivement à partir des cinq actes illocutoires de base : $\|assert\|$, $\|commit\|$, $\|direct\|$, $\|declare\|$ et $\|!\|$[34]. $\|\|$ est la fonction qui associe à chaque verbe illocutoire la force primitive qui lui correspond. Elle peut s'appliquer sur un contenu propositionnel p où p peut par exemple être associé à une assertion : $\|assert\|(p)$. Les forces illocutoires non primitives sont alors obtenues par application des opérations suivantes : ajout de conditions sur le contenu propositionnel, ajout de conditions préparatoires, ajout de conditions de sincérité, restriction du mode d'accomplissement et finalement en modulant les degrés d'intensité. Par exemple $\|promise\|$ est obtenu à partir de $\|commit\|$ en ajoutant des obligations au mode d'accomplissement du but commissif et en accroissant le degré d'intensité de la force illocutoire.

A.2 Les ACLs [Agent Communication Language] : instanciation de la théorie des actes de langage

Dans les SMAs cognitifs, l'hypothèse la plus répandue est que la communication inter-agent sera plus fructueuse si elle se fait par l'intermédiaire d'un langage de communication explicite [Dignum, 2000][35]. La propriété essentielle qui rend le langage utile, c'est que le sens de ses signes soit partagé. Ceci est vrai pour les langues vivantes, mais aussi pour tous les signes codés comme les coups de sifflet d'un arbitre de football. Ce que l'évolution a permis pour les langages humains, la standardisation le tente pour les agents artificiels [Finin et al., 1999]. En effet, par analogie simplifiée avec les sociétés humaines, les actes de langage peuvent servir pour réaliser la communication entre agents cognitifs[36] hétérogènes dans des systèmes artificiels. On utilise alors :

- *un dictionnaire de vocabulaire/signes commun(s)* : l'ontologie des services (dont le sens est au moins partiellement commun aux agents) qui permet une interprétation commune des contenus propositionnels des actes de langage ;
- *des actes de langage* utilisant cette ontologie qui servent de briques de base de la communication et correspondent aux attitudes propositionnelles transmises entre les agents ;

[34] Ce dernier marqueur est utilisé pour désigner l'acte illocutoire expressif de base.

[35] Les agents pourraient aussi communiquer, et c'est le cas dans certains systèmes réactifs, autrement que par des productions linguistiques classiques.

[36] Il est important de noter que les agents considérés ici sont des agents cognitifs. L'analogie avec les sociétés humaines et leurs usages du langage naturel est forte, car c'est ce qui différencie les ACLs des autres moyens d'échanges d'information et de savoir entre applications tels que les technologies RPC[Remote Procedure Call] : RMI, CORBA, ...

– *des protocoles ou règles de conversation communes pour structurer le dialogue entre les agents* : en effet, chaque acte de langage est à interpréter selon son contexte et plus particulièrement en fonction de son rapport avec les autres actes de langages qui l'entourent temporellement. L'acte de langage prend son sens comme élément d'un contexte et plus spécifiquement d'une conversation (ce point est développé au chapitre 3.2).

Les ACLs prennent donc place dans une couche logiquement supérieure à celle des protocoles de transfert (TCP/IP, HTTP, IIOP, ...) et adressent le niveau intentionnel et social des agents. Ils se différencient donc des mécanismes de la théorie de l'information[37] non seulement par leur structure et leur syntaxe plus complexe, faisant appel à des interprétations sémantiques « mentalistes », mais aussi par leurs spécifications génériques et précises qui permettent aux systèmes développés d'être théoriquement ouverts[38]. Les ACLs sont conçus pour permettre tout type de travaux théoriques ou d'applications pratiques. Au fil des années, de nombreux ACLs se sont développés, répondant à autant de besoins, mais ils ont presque tous en commun d'instancier tout ou partie de la théorie des actes de langage. Les sections suivantes présentent la palette des principaux ACLs, à commencer par KQML et FIPA-ACL qui sont les plus « standard ».

KQML [Knowledge Query and Manipulation Language]

Premier apparu, KQML fourni un ensemble d'actes de langage standards et utiles. Le langage est structuré selon trois niveaux enchâssés [Finin and Fritzson, 1994a, Finin et al., 1999] :

1. *la couche de communication* : renseigne sur la communication en indiquant l'identité du récepteur, de l'émetteur et la nature de la communication. Elle est minimale, car KQML ne prend pas en charge le transport lui-même (TCP/IP, SMTP, IIOP ou d'autres couches de niveau inférieur le font pour lui).

2. *la couche message* : donne des indications sur le contenu du message à savoir, le langage et l'ontologie utilisés pour le contenu ainsi que le type d'acte de langage attaché au contenu. C'est la couche centrale de KQML qui définit le type d'interaction que des agents-KQML pourront avoir.

3. *la couche de contenu* : contenu du message exprimé en KIF [Knowledge Interchange Format], Prolog, KQML ou autre. Notons que KQML ne traite pas cette couche si ce n'est pour savoir où le contenu commence et se termine. Comme le contenu du message est opaque, c'est à la couche message de le renseigner.

KQML prévoit deux modes de communication : la communication directe et la communication médiatisée. En communication directe un agent communique directement avec des agents connus capables de traiter ses messages. Dans le second mode de communication, un agent fait une demande de service à un agent particulier appelé facilitateur. Le facilitateur coordonne les

[37] La théorie de l'information, développée pas Shannon [Shannon and Weaver, 1975], sert de base théorique à la communication dans les réseaux informatiques.

[38] Historiquement, de nombreux SMAs ont été développés avec des mécanismes de communication ad hoc autres que les actes de langage de la linguistique classique. Mais sans théorie sémantique générique et partagée, il est impossible aux agents d'autres systèmes de comprendre leurs messages.

interactions en diffusant les demandes aux agents appropriés. Ce dernier mode suppose bien sûr que les agents s'enregistrent auprès du facilitateur. L'interopérabilité des agents logiciels peut être garantie à l'aide d'agents facilitateurs comme dans ABSE [Agent Based Software Engeenering] [Genesereth and Ketchpel, 1994].

KQML est muni d'une seule règle de conversation. Elle est simple, même si de nombreuses variantes sont permises. La conversation commence lorsque qu'un agent envoie un message KQML à un autre et se termine lorsque ce dernier répond. Les actes de langage autorisés par KQML sont restreints aux performatives. Ces performatives sont divisées en trois catégories :

1. 7 performatives de régulation de conversation traitent quelques cas particuliers (`sorry`, `error`) et permettent quelques variantes de la règle de conversation (`standby`, `ready`, `next`, `rest` et `discard`) ;

2. 17 performatives de discours permettent l'échange d'informations et de connaissances (`ask-if`, `tell`, ...) ;

3. 11 performatives d'assistance et de réseau pour étendre la conversation à plus de deux agents (`forward`, `broker-all`, ...).

KQML est issu d'un projet de la DARPA [Finin and Fritzson, 1994b], le KSE [Knowledge Sharing Initiative], initialement prévu comme moyen d'échange d'informations entre programmes à base de connaissances. Cependant, sa structure orientée message et la généricité de ses primitives lui permettent d'être utilisé comme ACL. C'est l'ACL le plus utilisé et implémenté dans la communauté SMA. Il en existe différentes variantes.

FIPA[39]-ACL

FIPA-ACL est le seul effort réellement organisé pour créer un ACL standard. Il a été conçu pour palier aux faiblesses des différentes versions de KQML. Le noyau de FIPA-ACL reprend le modèle formel et le langage de contenu d'ARCOL (voir ci-dessous) ainsi que la structure des messages de KQML. FIPA-ACL[40] diffère de KQML car :

– FIPA-ACL inclut une sémantique précise basée sur une logique multimodale quantifiée. Par contre, ce modèle sémantique est tellement complexe qu'un agent FIPA ne peut généralement pas déduire l'intention de l'agent émetteur. Pour les agents FIPA, on doit donc mettre en oeuvre un certain nombre de simplifications sémantiques (comme les stratégies de conversation) pour aider à l'interprétation des messages. KQML, lui, a des sémantiques relativement simples qui peuvent être directement utilisées lorsque le domaine est restreint.

– KQML inclut un certain nombre de messages d'administration de réseaux qui simplifient la partie du développement concernant les flux de données, la registration, le courtage, les services de répertoires et autres. Dans FIPA-ACL, ces fonctions administratives sont assurées par l'usage de messages `REQUEST` et `INFORM` normaux aux contenus spécialisés.

[39]La Foundation for Inteligent Physical Agents est un organisme international réunissant de nombreux acteurs industriels et académiques ayant pour but de développer un ensemble de standards génériques pour agents, voir [FIPA, 2002] pour de plus amples renseignements.

[40]L'ACL de FIPA s'appelle en fait ACL, mais pour éviter toute confusion nous le nommerons FIPA-ACL dans le reste du texte.

– Comme FIPA-ACL est défini par une corporation de chercheurs, sa mise à jour est lente, mais chaque version est scrupuleusement vérifiée. Le développement de KQML n'est pas centralisé et ainsi plusieurs variantes incompatibles sont nées.

ARCOL [ARtimis Communication Language]

ARTIMIS est une plate-forme générique pour agents communicants développée par France Télécom [Sadek et al., 1997]. Dans ARTIMIS, un agent peut communiquer avec un humain aussi bien qu'avec un autre agent. Les faits de communication des agents sont modélisés comme des actions rationnelles via ARCOL. Une expression ARCOL est écrite en SL[Semantic Language] pour le message et utilise SCL[Semantic Content Language] pour le contenu.

ARCOL définit quatre primitives mutuellement exclusives et composables : `Inform, Request, Confirmation` et `Inform Referent`. Une des limites d'ARCOL est qu'il présuppose la sincérité des agents, ce qui est un obstacle à l'objectif d'ouverture des ACLs.

AOP [Agent Oriented Programming]

C'est en 1989, que Shoham introduit le paradigme de programmation orientée agent (AOP) [Shoham, 1993] en empruntant à l'intelligence artificielle, à la théorie des actes de langages et à la programmation orientée objets. AOP ne dispose que de trois primitives communicationnelles composables : `Inform, Request, Unrequest.`

ICL [InterAgent Communication Language]

OAA [Open Agent Architecture] est une plate-forme multi-agent pour environnements distribués développée par SRI Internationnal [SRI Internationnal, 1999]. ICL est un langage déclaratif logique vu comme middle-ware (intergiciel) pour l'implémentation de mécanismes de coopérations entre les serveurs de services, les clients et les agents facilitateurs du cadre OAA. ICL utilise trois types de primitives : `Solve, Do` et `Post`. Le contenu des messages est exprimé en Prolog. ICL est cependant restreint à l'architecture OAA et à sa nature procédurale qui interdit les échanges bidirectionnels.

MAC [Mobile Agent Communication]

Un agent mobile est un programme qui peut migrer de façon autonome dans un réseau de machines hétérogènes. Du fait de son jeune âge, ce type d'agents relève plus du génie logiciel que de l'intelligence artificielle. De tels agents sont considérés comme de simples processus et non comme des agents cognitifs disposant d'attitudes mentales et évoluant dans une organisation sociale. Ainsi, les mécanismes de communications pour ces agents sont généralement de simples passages de messages. D'autres alternatives encore plus primitives sont utilisées : RMI[Remote Method Invocation] , CORBA[Common Object Request Broker Architecture]. Un exemple d'agents mobiles communicants par RPC[Remote Procedure Call] ou passage de message sont les Aglets[AGile appLETS] : des objets Java développés par le centre de recherche IBM de Tokyo.

Autres ACLs

On compte un certain nombre d'autres ACLs :

- COOL : langage SMA conçu par l'Université Technique de Berlin et l'Enterprise Integration Laboratory at the University of Toronto.
- LOGOS : un ACL développé par la NASA avec des éléments spécifiques au contrôle aérien.
- PLACA : langage dérivé de AGENT-0 et KQML pour intégrer les intentions et des éléments de planification.
- APRIL et MAIL : langages développés dans le cadre du projet ESPRIT : IMAGINE.

Problème de complétude des ACLs

De nombreuses discussions sont encore en cours sur la complétude des types d'actes de communication définis dans l'un ou l'autre des ACLs précédents. En philosophie du langage, on trouve les catégories suivantes, dont les cinq premières sont celles définies dans la théorie des actes de langage[41] : représentatif ou assertif, directif (ordre ou demande), commissif (promesse), expressif, déclaratif, permissif (autorisation) et prohibitif (interdiction). Or, force est de constater que les ACLs ne couvrent généralement pas toute cette panoplie. Par exemple, FIPA-ACL et KQML ne définissent que des primitives assertives et directives. Cela a pour conséquence que certaines attitudes des agents cognitifs ne sont pas exprimables dans ces ACLs populaires [Chaib-draa and Dignum, 2002].

Problèmes concernant les ontologies

Le langage de la couche de contenu, ainsi que le vocabulaire qui y est utilisé doivent être compris et donc partagé entre les agents. Pour ce faire, un champ indiquant l'ontologie à utiliser est prévu dans KQML (originellement pour une ontologie d'Otonlingua) comme dans FIPA-ACL. Pour être utile, une ontologie relative à un domaine doit être complète et donc relativement volumineuse. Le problème de savoir quand et comment un agent doit intégrer une (nouvelle) ontologie reste ouvert.

A.3 Théorie contextuelle des actes de langage

La théorie contextuelle des actes de langage est une variante célèbre de cette dernière qui nous permet d'introduire et de mettre en relief les limitations de celle-ci. Cette version contextuelle de la théorie des actes de langage est issue d'une remise en cause de l'hypothèse de force littérale qui affirme que la force illocutoire d'un acte est essentiellement pragmatique. Par exemple, l'énoncé S_3 suivant sera interprété différemment selon que l'on se trouve dans une école, au commissariat de police, dans la rue avec des amis ou dans le bureau de son directeur :

S_3 *: Qu'avez-vous fait hier ?*

L'idée est donc de se désintéresser de l'énoncé de l'acte dans sa forme pour se focaliser sur le contexte dans lequel il est survenu et sur lequel il va agir en retour. Bunt [Bunt, 1996] définit un acte de langage comme l'application d'une fonction communicative à un contenu propositionnel. Cette fonction communicative spécifie quelle mise à jour du contexte doit être

[41] Les partisans de la théorie des actes de langage réduisent les autres catégories à des sous types de celles-ci

effectuée pour prendre en compte l'information transmise. L'idée est qu'un acte de langage est dans ses effets une fonction du contexte vers le contexte. Tout dépend donc de la définition que l'on donne à la notion de contexte. Pour Bunt, le contexte n'est que la somme des contextes cognitifs des interlocuteurs, ce qui nous amène aux approches intentionnelles (voir section 4).

Pour d'autres, il s'agit d'exclure les états mentaux des interlocuteurs du contexte (Stalnaker et Levinson). Le contexte est alors directement lié à l'activité dialoguale. Il est constitué de ce qui a été accepté lors du dialogue. Un acte assertif de contenu propositionnel p ajoute l'engagement propositionnel correspondant au contexte, d'autres actes permettent de retirer certains éléments du contexte. On se dirige alors vers les approches dites conventionnelles (voir section 5). Finalement, d'autres éléments influant sur la production des actes de langages relèvent du contexte. Le contexte incluant entre autres :

1. *certaines conventions et hypothèses* : on parle en français et pas en anglais, on parle travail et on ne joue pas aux charades, ...

2. *certains gestes et inflections de voie* : froncer les sourcils, montrer du doigt ou faire des mimiques expressives, ...

3. *des éléments pertinents de l'histoire de la conversation* : pour fixer les références pronominales, ...

4. *des faits ambiants pertinents* : « on se voit dans une heure. », signifie qu'il faut avoir une idée de l'heure qu'il est au moment de l'énonciation, ...

Dans tous les cas, ces aspects contextuels ne sont pas pris en charge par la version analytique searlienne de la théorie[42].

[42]Encore une fois, ses défenseurs, desquels nous sommes, diront que ce n'est pas son objectif et repousseront cet aspect vers la théorie de l'action ou la pragmatique.

B Sémantique linguistique et sémantique mathématique

Dans la modélisation du dialogue, il y a deux notions de sémantique en jeux. Ces deux notions entrent en conflit et il est prudent de bien indiquer de laquelle on traite. En linguistique, la sémantique est la relation entre les signes et les objets auxquels ils s'appliquent. En logique mathématique, la sémantique d'un langage formel est une relation entre le langage et un espace de structures mathématiques appelées modèles. Il est commun de définir plusieurs sémantiques mathématiques pour un langage, cela permet de démontrer plus de propriétés et les comparaisons de ces sémantiques peuvent également être enrichissantes ; on parle alors de théorie des modèles ou de méta-mathématique.

À l'inverse, la sémantique linguistique devra être la plus générique et universelle possible dans l'esprit du projet linguistique lui-même. Dans le cadre des ACLs, ces deux notions de sémantique entrent en conflit et faute de précision, les discours sur la sémantique des ACLs sont souvent confus. En effet, les ACLs étant des langages formels, l'étude de leurs propriétés peut être réalisée par la définition de sémantiques mathématiques les concernant. En outre, comme ils sont un média de communication, de la même façon que le langage naturel, on aimerait s'assurer qu'ils ont une sémantique précise, garantissant à ceux qui l'utilisent de partager le sens des énoncés. Notons au passage que la vérification que des agents partagent la même compréhension d'un énoncé, c'est-à-dire qu'ils lui attribuent le même sens linguistique, passe par l'examination de leurs états mentaux. Pour McBurney [Mac Burney, 2002], cela pose un problème conceptuel puisqu'un agent pourra toujours simuler de tels états mentaux de manière insincère. Contentons-nous pour l'heure de dresser une brève typologie des sémantiques mathématiques.

Une première manière de définir une sémantique mathématique d'un langage formel est de définir chaque élément du langage en terme des pré-conditions qui doivent tenir pour qu'il soit réalisé et des post-conditions qui doivent s'appliquer à son utilisation (comme pour les actions dans STRIPS). C'est ce qu'on nomme une *sémantique axiomatique*. Pour ce qui est de la communication agent, on peut faire la différence entre les sémantiques axiomatiques publiques qui n'utilisent que des éléments publics, accessibles à tous et les sémantiques axiomatiques privées qui font référence à des éléments internes aux agents. Pour illustrer cette idée de sémantique axiomatique, prenons l'exemple de la sémantique des langages de programmation. La méthode la plus fréquemment utilisée pour spécifier la sémantique d'un programme consiste à définir les pré- et post-conditions de chaque élément de code[43]. En guise d'exemple, considérons l'instruction « $x := 3.8$ ». La sémantique de cette instruction inclut la pré-condition : « x est le nom d'un emplacement mémoire pouvant recevoir la valeur d'un nombre de type Réel » et la post-condition : « A l'emplacement mémoire dénoté par x on trouve la valeur 3.8 »[44]. La pré-condition sert à imposer que l'instruction doit pouvoir être exécutée et la post-condition rend compte de l'effet minimal de l'instruction sur l'environnement du programme. La relation entre la pré et la post-condition correspond à la description de la compréhension intuitive que l'on a du fonctionnement de l'instruction, c'est pourquoi on parle de sémantique.

Les pré et post-conditions des instructions dans les langages de programmation séquentiel sont exprimées en termes de valeurs des variables, car la programmation informatique de ce niveau se limite à la manipulation de variables. Mais pour ce qui est des ACLs, on travaille à un niveau supérieur. En fait, les pré-conditions et post-conditions des ACLs devront être

[43] Voir, par exemple, la méthode des triplets de Hoare.

[44] Et ce, quelque soit la valeur dénotée par x auparavant.

exprimées en termes d'attitudes mentales des agents en présence. En effet, les ACLs et leurs sémantiques doivent co-exister et s'aligner avec les théories mentales et comportementales des agents [Dignum, 2000].

Les *sémantiques opérationnelles* sont un autre type de sémantique mathématique qui considèrent les énoncés comme des transitions dans une machine abstraite. Dans ce cadre, un énoncé est défini par le changement d'état qu'il opère dans la machine abstraite. Des sémantiques opérationnelles ont été définies pour certains ACLs. On trouve aussi des *sémantiques dénotationnelles* dans lesquelles chaque élément du langage est mis en relation avec une entité mathématique abstraite appelée dénotation. La sémantique des mondes possible est un exemple de sémantique dénotationnelle. Pour qu'une sémantique dénotationelle soit utile, il faut qu'il soit possible de déterminer le sens sémantique d'un énoncé en fonction du sens de ses éléments, une propriété appelée compositionnalité. Cette propriété est absente dès qu'un énoncé a une dénotation qui ne correspond pas à la composition de celles de ces composants. Finalement, les sémantiques de la théorie des jeux peuvent aussi s'avérer utiles dans certains cas. Dans celles-ci, chaque énoncé bien formé est associé à un jeu à deux joueurs, un protagoniste et un antagoniste. Un énoncé est alors considéré vrai, si et seulement si, il existe une stratégie gagnante (série de coups qui sont gagnants quelle que soit la réponse de l'antagoniste) pour le protagoniste.

Références

[Airenti et al., 1993] Airenti, G., Bara, B., and Colombetti, M. (1993). Conversation and behaviour games in the pragmatics of dialogue. *Cognitive science*, 17(7) :34–49.

[Allen and Perrault,] Allen, J. F. and Perrault, C. R. Analysing intention in dialogues. *Artificial Intelligence*, 15(3) :23–46.

[Allen et al., 1995] Allen, J. F., Schubert, L. K., Ferguson, G., Heeman, P., Hwang, C. H., Kato, T., Light, M., Martin, N., Miller, B., Poesio, M., and Traum, D. R. (1995). The TRAINS project : A case study in building a conversationnal planning agent. *Journal of Experimental and Theorical AI*, 7(4) :7–48.

[Allwood, 1976] Allwood, J. (1976). *Linguistic comunication as action and cooperation.* PhD thesis, University of Goteborg, Sweden.

[Allwood, 1994] Allwood, J. (1994). Obligations and options in dialogue. *Think Quaterly*, 3(1) :12–34.

[Amgoud et al., 2000] Amgoud, L., Maudet, N., and Parson, S. (2000). Modelling dialogue using argumentation. In *Proceedings of the 4th internationnal conference on multi-agent systems (ICMAS'00).*

[Austin, 1962a] Austin, J. L. (1962a). *How to Do Things With Words.* Oxford University Press : Oxford, England.

[Austin, 1962b] Austin, J. L. (1962b). *Quand dire, c'est faire.* Seuil, Paris.

[Bach and Harnish, 1979] Bach, K. and Harnish, R. M. (1979). *Linguistic Communication and Speech Acts.* MIT Press, Cambridge, Mass.

[Baker, 1991] Baker, M. (1991). *Knowledge Acquisition in Physics and Learning Environments*, chapter An Analysis of Cooperation and Conflict in students' collaborative explanations for Phenomena in Mechanics. Springer-Verlag : Heidelberg, Germany.

[Balkanski and Hurault-Plantet, 2000] Balkanski, C. and Hurault-Plantet, M. (2000). Cooperative requests and replies in a collaborative dialogue model. In Press, A., editor, *Int. J. Human-Computer Studies*, volume 53, pages 915–968.

[Barreau, 1995] Barreau, H. (1995). *L'épistémologie.* Presses Universitaires de France, que sais-je ? collection encyclopédique edition.

[Bates, 1994] Bates, J. (1994). The role of emotion in believable agents. *Communications of the ACM*, 37(7) :122–125.

[Bell, 1995] Bell, J. (1995). Changing attitudes. In Wooldridge, M. and Jennings, N. R., editors, *Intelligent Agents : Theories, Architectures, and Languages (LNAI Volume 890)*, pages 40–55. Springer-Verlag : Heidelberg, Germany.

[Brassac, 1994] Brassac, C. (1994). Speech act and conversation sequencing. *Pragmatics and cognition*, 2(1) :191–205.

[Bratman, 1987] Bratman, M. E. (1987). *Intentions, Plans, and Practical Reason.* Harvard University Press : Cambridge, MA.

[Bratman, 1990] Bratman, M. E. (1990). What is intention ? In Cohen, P. R., Morgan, J. L., and Pollack, M. E., editors, *Intentions in Communication*, pages 15–32. The MIT Press : Cambridge, MA.

[Brehm and Cohen, 1962] Brehm, J. and Cohen, A. (1962). *Explorations in Cognitive Dissonance.* John Wiley and Sons, inc.

[Broersen et al., 2001] Broersen, J., Dastani, M., Hulstijn, J., Huang, Z., and Van der Torre, L. (2001). The BOID architecture : Conflicts between beliefs, obligations, intention and desires. In *International Conference on Autonomous Agent 2001.*

[Bruce, 1975] Bruce, B. (1975). Generation as social action. *Theoretical issues in Natural Language Processing*, 1 :64–67.

[Bunt, 1996] Bunt, H. (1996). *The Structure of Multimodal Dialogue*, chapter Dynamic Interpretation and Dialogue Theory. John Benjamin, Amsterdam.

[Bylander, 1991] Bylander, E. (1991). Complexity results for planning. In *Proceedings of the 12th International Joint Conference on Artificial Intelligence*, pages 274–279, Sydney, Australia.

[Camps, 1998] Camps, V. (1998). *Vers une théorie de l'auto-organisation dans les systèmes multi-agents basée sur la coopération : application à la recherche d'information dans un système d'information répartie.* PhD thesis, IRIT, Toulouse, France.

[Carberry, 1990] Carberry, S. (1990). *Plan Recognition in Natural Language Dialogue.* MIT Press.

[Chaib-draa and Dignum, 2002] Chaib-draa, B. and Dignum, F. (2002). Trends in agent communication language. *Computationnal Intelligence*, 18(2) :89–101.

[Chaib-draa et al., 2002] Chaib-draa, B., Maudet, N., and Labrie, M. (2002). Request for action reconsidered as dialogue game based on commitments. In *Workshop on Agent Communication Language (AAMASť02).*

[Chaib-draa and Vanderveken, 1998] Chaib-draa, B. and Vanderveken, D. (1998). Agent communication language : Toards a semantics based on success, satisfaction, and recursion. In *ATAL98*, Paris.

[Chaib-draa and Vongkasem, 2000] Chaib-draa, B. and Vongkasem, L. (2000). ACL as a joint project between participants : A preliminary report. In Dignum, F. and Greaves, M., editors, *Issues in Agent Communication*, number 1916 in Lecture Notes in Artificial Intelligence, pages 235–248. Springer-Verlag : Heidelberg, Germany.

[Clark, 1996] Clark, H. H. (1996). *Using Language.* Cambridge University Press.

[Clark and Schaeffer, 1987] Clark, H. H. and Schaeffer, E. F. (1987). Collaborating on contributions to conversations. *Language and Cognitive Processes*, 2 :19–41.

[Clark and Wilkes-Gibbs, 1986] Clark, H. H. and Wilkes-Gibbs, D. (1986). Refering as a collaborative process. *Cognition*, 1(22) :1–35.

[Cohen, 1996] Cohen, P. R. (1996). *Survey of the State of Art in Natural Language Technology*, chapter Discourse and Dialogue, pages p.234–241. Universitat des Saarlandes.

[Cohen and Levesque, 1990a] Cohen, P. R. and Levesque, H. J. (1990a). Intention is choice with commitment. *Artificial Intelligence*, 42 :213–261.

[Cohen and Levesque, 1990b] Cohen, P. R. and Levesque, H. J. (1990b). Performatives in a rationally based speech act theory. In *28th Annual Meeting of the Association for Computational Linguistics*, pages 79–88, Pittsburg. Springer-Verlag : Heidelberg, Germany.

[Cohen and Levesque, 1990c] Cohen, P. R. and Levesque, H. J. (1990c). Rational interaction as the basis for communication. In Cohen, P. R., Morgan, J., and Pollack, M. E., editors, *Intentions in Communication*, pages 221–256. The MIT Press : Cambridge, MA.

[Cohen and Levesque, 1991] Cohen, P. R. and Levesque, H. J. (1991). Teamwork. Technote 504, SRI International, Menlo Park, CA.

[Cohen and Perrault, 1979] Cohen, P. R. and Perrault, C. R. (1979). Elements of a plan based theory of speech acts. *Cognitive Science*, 3 :177–212.

[Colombetti, 1998] Colombetti, M. (1998). Different ways to have something in common. In Christiansen, H., Andreasen, T., and Larsen, H. L., editors, *Procedings of the third internationnal conference on flexible query answering systems*, pages 95–109. Springer, Berlin.

[Colombetti, 2000] Colombetti, M. (2000). Commitment-based semantic for agent communication languages. In *1st Workshop on the History and Philosophy of Logic, Mathematics and Computation*.

[Craig, 1983] Craig, R. (1983). *Conversationnal coherence : form, Structure and strategy.* CA :Sage, Beverly Hills.

[Craig, 1993] Craig, R. (1993). Why are there so many communication theories ? *Journal of Communication*, 43 :26–33.

[Dastani et al., 2000] Dastani, M., Hulstijn, J., and der Torre, L. (2000). Negociation protocols and dialogue games. In *Proceedings of the Belgium/Dutch AI Conference (BNAIC'2000)*, Kaatsheuvel.

[Dessalles, 1998] Dessalles, J.-L. (1998). *Formal Semantics and Pragmatics of Dialogue*, volume TWLT-13, chapter The Interplay of Desire and Necessity in Dialogue, pages 89–97. Enschede : University of Twente.

[Dessalles and Ghadakpour, 1999] Dessalles, J.-L. and Ghadakpour, L. (1999). L'activité scientifique en tant que comportement naturel ancré sur le conflit cognitif. In *Actes des Huitièmes Journées de Rochebrune : Conflits des Interprétations et Interprétation des Conflits*, pages 87–98, Paris. ENST 99-S-001.

[Dignum, 2000] Dignum, F. (2000). Issues in agent communication : An introduction. In Dignum, F. and Greaves, M., editors, *Issues in Agent Communication*, Lecture Notes in Artificial Intelligence, pages 1–16. Springer-Verlag : Heidelberg, Germany.

[Ferguson, 1995] Ferguson, G. (1995). *Knowledge Representation and Reasoning for Mixed-Initiative Planning.* Tr-562, University of Rochester, Department of Computer Sciences.

[Festinger, 1954] Festinger, L. (1954). *A Social Communication and Cognition : A very Preliminary and Highly Tentative Draft*, chapter Appendix A. American Psychological Association, Washington DC.

[Festinger, 1957] Festinger, L. (1957). *A Theory of Cognitive Dissonance.* CA : Stanford University Press.

[Fikes and Nilsson, 1971] Fikes, R. and Nilsson, N. (1971). Strips : A new approach to the application of theorem proving to problem solving. *Artificial Intelligence*, 2 :189–208.

[Finin and Fritzson, 1994a] Finin, T. and Fritzson, R. (1994a). KQML - a language and protocol for knowledge and information exchange. In *Proceedings of the Thirteenth International Workshop on Distributed Artificial Intelligence*, pages 126–136, Lake Quinalt, WA.

[Finin and Fritzson, 1994b] Finin, T. and Fritzson, R. (1994b). Specification of the KQML agent communication language plus example agent policies and architectures. Technical report, DARPA Knowledge Sharing Initiative External Interface Working Group.

[Finin and Labrou, 1998] Finin, T. and Labrou, Y. (1998). Semantics for an ACL. In *Agents, Architecture and Languages.* Springer-Verlag : Heidelberg, Germany, Berlin.

[Finin et al., 1999] Finin, T., Labrou, Y., and Peng, Y. (1999). Agent communication languages : The current landscape. *IEEE Intelligent Systems*, pages 45–52.

[FIPA, 2002] FIPA (2002). Fipa [fondation for intelligent physical agents]. http ://www.FIPA.org.

[Flores and Kremer, 2001] Flores, R. and Kremer, R. (2001). Bringing coherence to agent conversation. In Wooldridge, M., Ciancarini, P., and Weiss, G., editors, *Agent-Oriented Software Engineering II*, Lecture Notes in Computer Science, pages 50–67. Springer-Verlag.

[Genesereth and Ketchpel, 1994] Genesereth, M. R. and Ketchpel, S. P. (1994). Software agents. *Communications of the ACM*, 37(7) :48–53.

[Grice, 1969] Grice, H. (1969). Utterer's meaning and intentions. *Philosophical Review*, pages 147–177.

[Grice, 1975] Grice, H. (1975). *Syntax and Semantics : Speech acts*, volume 3, chapter Logic and Conversation. Academic Press.

[Grice, 1957] Grice, H. P. (1957). Meaning. *Philosophical review*, (66) :377–388.

[Grosz and Kraus, 1996] Grosz, B. J. and Kraus, S. (1996). Collaborative plans for complex group action. *Artificial Intelligence*, (86) :269–357.

[Grosz and Sidner, 1986] Grosz, B. J. and Sidner, C. L. (1986). Attention, intentions and the structure of discourse. *Computationnal Linguistics*, 12 :175–204.

[Grosz and Sidner, 1990] Grosz, B. J. and Sidner, C. L. (1990). Plans for discourse. In Cohen, P. R., Morgan, J., and Pollack, M. E., editors, *Intentions in Communication*, pages 417–444. The MIT Press : Cambridge, MA.

[Hamblin, 1970] Hamblin, C. (1970). *Fallacies*. Methuen, London, UK.

[Harmon-Jones and Mills, 1999] Harmon-Jones, E. and Mills, J., editors (1999). *Cognitive Dissonance : Progress on a Pivotal Theory in Social Psychology*. American Psychological Association.

[Houdé et al., 1998] Houdé, O., Kayser, D., Koenig, O., Proust, J., and Rastier, F. (1998). *Vocabulaire des Sciences Cognitives*. PUF.

[Hovland et al., 1953] Hovland, C., Janis, I., and Kelley, H. (1953). *Communication and Persuasion*. Yale University Press, New Haven.

[Hulstiln, 2000] Hulstiln, J. (2000). Dialogue games are recipe for joint action. In *Proceedings of the 4th workshop on the semantics and pragmatics of dialogue (GOTALOG'00)*.

[Keefe, 1991] Keefe, J. (1991). *Cognitive Dissonance*, chapter Persuasion : Theory and research, pages 61–78. Sage, Newbury Park, California.

[Kone et al., 2000] Kone, M., Shimazu, A., and Nakajima, T. (2000). The state of the art in agent communication languages. *Knowledge and Information Systems*, (2) :259–284.

[Kowtko et al., 1991] Kowtko, J., Isard, S., and Doherty, G. (1991). Conversationnal games within dialogue. In *Proceedings of the Esprit Workshop on Discourse Coherence*.

[Labrou, 1996] Labrou, Y. (1996). *Semantics for an Agent Communication Language*. PhD thesis, Computer Science and Electrical Engineering Department, University of Maryland, Baltimore.

[Lambert and Carberry, 1991] Lambert, L. and Carberry, S. (1991). A tripartite plan-based model of dialogue. In *Procedings of the 29th annual meeting of ACL*, pages 47–54.

[Levin and Moore,] Levin, J. A. and Moore, J. A. Dialogue-games : Metacommunication structures for natural language understanding. *Cognitive Science*, 1(4) :384–420.

[Levinson, 1979] Levinson, S. C. (1979). Activity type and language. *Linguistics*, 17 :365–399.

[Levinson, 1983] Levinson, S. C. (1983). *Pragmatics*. Cambridge University Press : Cambridge, England.

[Lewin, 2000] Lewin, I. (2000). A formal model of conversation games theory. In *proceedings of the 4th workshop on the semantics and pragmatics of dialogue (GOTALOG'00)*.

[Litman and Allen, 1990] Litman, D. and Allen, J. (1990). *Intentions in communication*, chapter Discourse Processing and Common Sens Plans, pages 365–388. MIT Press, Cambridge, USA.

[Littlejohn, 1992] Littlejohn, S. (1992). *Theories of Human Communication.* Waldsworth Publishing Company.

[Lochbaum, 1994] Lochbaum, K. E. (1994). *Using Collaborative Plans to Model the Intentional Structure of Discourse.* PhD thesis, Harvard University, Cambridge, MA.

[Mac Burney, 2002] Mac Burney, P. J. (2002). *Rationnal Interaction.* PhD thesis, University of Liverpool.

[MacKenzie, 1979] MacKenzie, J. (1979). Question-begging in non-cumulative systems. *Journal of philosophical logic*, 8 :117–133.

[Mann, 1988] Mann, W. (1988). Dialogues games : conventions of human interaction. *Argumentation*, 4(2) :511–532.

[Maudet, 2001] Maudet, N. (2001). *Modéliser les Conventions des Interactions Langagières : la Contribution des Jeux de Dialogue.* PhD thesis, ENSEIHT-UPS.

[Maudet and Chaib-draa, 2001] Maudet, N. and Chaib-draa, B. (2001). Commitment-based and dialogue-game based protocols - new trends in agent communication language. *Knowledge Engineering.*

[McBurney and Parson, 2000] McBurney, P. and Parson, S. (2000). Agent ludens : games for agent dialogues. In *Proceedings of the Workshop on Game Theoretic and Decision Theoretic agents.*

[McGuire, 1960] McGuire, W. (1960). Cognitive consistency and attitude change. *Journal of Abnormal Social Psychology*, 60 :345–353.

[Miller and Dollar, 1941] Miller, N. and Dollar, J. (1941). *Social Learning and Imitation.* Yale University Press, Hew Haven, Connecticut.

[Moore, 1990] Moore, R. C. (1990). A formal theory of knowledge and action. In Allen, J. F., Hendler, J., and Tate, A., editors, *Readings in Planning*, pages 480–519. Morgan Kaufmann Publishers : San Mateo, CA.

[Moulin, 1997] Moulin, B. (1997). *Agent and Multi-agent Systems*, chapter The social dimension of interacions in multi-agent systems. LNAI 1441. Springer, Berlin.

[Newcomb, 1953] Newcomb, T. (1953). An approach to the study of communicative acts. *Psychological Review*, 60(6) :393–404.

[Nuyts, 1994] Nuyts, J. (1994). The intentional and socio-cultural in language use. *Pragmatics and Cognition*, 2(2) :237–268.

[Osgood, 1963] Osgood, C. (1963). On understanding and creating sentences. *American Psychologist*, 18 :735–751.

[Pasquier and Chaib-draa, 2002a] Pasquier, P. and Chaib-draa, B. (2002a). The cognitive coherence approach for agent communication pragmatic. In *Proceedings of AAMAS'03.* Soumis.

[Pasquier and Chaib-draa, 2002b] Pasquier, P. and Chaib-draa, B. (2002b). Cohérence et conversations entre agents : vers un modèle basé sur la consonnance cognitive. In Müller, J. and Mathieu, P., editors, *Systèmes multi-agents et systèmes complexes, actes des JFIADSMA'02*, pages 188–203, Paris. Hermes Science Publication.

[Pasquier and Dehais, 2000] Pasquier, P. and Dehais, F. (2000). Approche Générique du Conflit. In Scapin, D. and Vergisson, E., editors, *ErgoIHM 2000*, Biarritz, France. ESTIA.

[Poesio and Mikheev, 1998] Poesio, M. and Mikheev, A. (1998). The predictive power of game structure in dialogue act recognition : Experimental results using maximum entropy estimation. In *Proceedings of the 5th International Conference on Spoken Language Processsing (ICSLP'98)*, Sydney, Australia.

[Pollack, 1990] Pollack, M. E. (1990). Plans as complex mental attitudes. In Cohen, P. R., Morgan, J., and Pollack, M. E., editors, *Intentions in Communication*, pages 77–104. The MIT Press : Cambridge, MA.

[Ramshaw, 1989] Ramshaw, L. (1989). A meta-plan model for problem solving discourse. In *Proceedings of the 4th Conference of the European Chapter of ACL*, pages 35–42.

[Rao and Georgeff, 1995] Rao, A. S. and Georgeff, M. (1995). BDI Agents : from theory to practice. In *Proceedings of the First International Conference on Multi-Agent Systems (ICMAS-95)*, pages 312–319, San Francisco, CA.

[Reed, 1998] Reed, C. (1998). Dialogue frames in agent communication. In *Proceedings of the Third Internationnal Conference on MultiAgent Systems (ICMAS).*

[Reed and Long, 1997] Reed, C. and Long, D. (1997). Collaboration, cooperation and dialogue classification. In *IJCAI 1997*.

[Reiter, 1996] Reiter, R. (1996). Knowledge in action : Logical fondation for describing and implementing dynamical systems. Draft, CS departement Toronto University.

[Sadek, 1991a] Sadek, D. (1991a). *Attitudes mentales et interaction rationnelle : vers une théorie formelle de la communication*. PhD thesis, Université de Rennes 1, France.

[Sadek et al., 1997] Sadek, D., Bretier, P., and Panaget, F. (1997). ARTIMIS : Natural dialogue meets rational agency. In *Proceedings of the Fifteenth International Joint Conference on Artificial Intelligence (IJCAI-97)*, Yokohama, Japan.

[Sadek, 1991b] Sadek, M. (1991b). Dialogue acts are rational plans. In *Proceedings of the ESCA/ETRW Workshop on the Structure of Multimodal Dialogue*, pages 1–29, Maratea, Italy.

[Sakai, 2001] Sakai, H. (2001). *A Multiplicative Power-Function Model of Cognitive Dissonance : Toward an Integrated Theory of Cognition, Emotion and Behavior After Leon Festinger*, chapter Computer Simulation, pages 267–294. American Psychological Association.

[Searle, 1969] Searle, J. R. (1969). *Speech Acts : An Essay in the Philosophy of Language.* Cambridge University Press : Cambridge, England.

[Searle, 1979] Searle, J. R. (1979). *Expression and Meaning*. Cambridge University Press : Cambridge, England.

[Searle, 1983] Searle, J. R. (1983). *Intentionality : An Essay in the Philosophy of Mind*. Cambridge University Press : Cambridge, England.

[Searle, 1992a] Searle, J. R. (1992a). *La Redécouverte de l'Esprit*. NRF Essais. Gallimard.

[Searle, 1992b] Searle, J. R. (1992b). *(On) Searle on Conversation*, chapter Conversation, pages 7–29. Benjamins Pub, Philadelphia.

[Searle and Vanderveken, 1985] Searle, J. R. and Vanderveken, D. (1985). *Foundations of Illocutionary Logic*. Cambridge University Press, NY.

[Shannon and Weaver, 1975] Shannon, C. E. and Weaver, C. (1975). *Théorie mathématique de la communication, traduit de l'anglais par Cosnier, Dahan et Economides*. Retz CEPL.

[Sherif and Hovland, 1961] Sherif, M. and Hovland, C. (1961). *Social Judgement*. Yale University Press, New Haven.

[Shoham, 1993] Shoham, Y. (1993). Agent-oriented programming. *Artificial Intelligence*, 60(1) :51–92.

[Shultz and Lepper, 1999] Shultz, R. and Lepper, R. (1999). *Cognitive Dissonance : progress in a pivotal theory in social psychology*, chapter Computer simulation of the cognitive dissonance reduction, pages 235–265. American Psychological Association.

[Singh, 2000] Singh, M. (2000). A social semantics for agent communication languages. In Dignum, F. and Greaves, M., editors, *Issues in Agent Communication*, pages 31–45. Springer-Verlag : Heidelberg, Germany.

[Singh, 1994] Singh, M. P. (1994). *Multiagent Systems : A Theoretical Framework for Intentions, Know-How, and Communications (LNAI Volume 799)*. Springer-Verlag : Heidelberg, Germany.

[Singh, 1998] Singh, M. P. (1998). Agent communication languages : rethinking the principles. *IEEE Computer*, 12(31) :40–47.

[Smith et al., 1995] Smith, R. W., Hipp, D. R., and Biermann, A. W. (1995). An architecture for voice dialogue systems based on prolog-style theorem proving. *Computationnal Linguistics*, (21) :281–320.

[Sperber and Wilson, 1986] Sperber, D. and Wilson, D. (1986). *Relevance*. Harvard University Press, Cambridge MA.

[SRI Internationnal, 1999] SRI Internationnal (1999). Open Agent Architecture. Internet, http ://www.ai.sri.com/ oaa.

[Sun and Browne, 1999] Sun, R. and Browne, A. (1999). Connectionist variable binding. *Expert Systems*, 16(3) :189–207.

[Sun and Browne, 2001] Sun, R. and Browne, A. (2001). Connectionist inference models. *Neural Networks*, 14(10) :1331–1355.

[Thagard, 2000] Thagard, P. (2000). *Coherence in Thought and Action*. The MIT Press.

[Thagard and Verbeurgt, 1998] Thagard, P. and Verbeurgt, K. (1998). Coherence as constraint satisfaction. *Cognitive Science*, 22 :1–24.

[Traum, 1994] Traum, D. (1994). Discourse obligations in dialogue processing. In *Proceedings of ACL'94*.

[Traum, 1997] Traum, D. R. (1997). A reactive-deliberatie model of dialogue agency. In Müller, J.-P., Wooldridge, M., and Jennings, N., editors, *Intelligent Agent III*, number 1193 in LNAI, pages 157–171. Springer Verlag.

[Traum and Hinkelman, 1992] Traum, D. R. and Hinkelman, E. A. (1992). Conversation acts in task-oriented spoken dialogue. *Computationnal Intelligence*. Special Issue on Non-literal Language.

[Traum and Poesio, 1997] Traum, D. R. and Poesio, M. (1997). Conversational actions and discourse situations. *Computationnal Intelligence*, 13(3) :309–347.

[Vanderveken, 1990] Vanderveken, D. (1990). *Meaning and Speech Acts : Principles of Language Use*. Cambridge University.

[Vanderveken, 1999] Vanderveken, D. (1999). *Analyse et simulation de conversations : de la théorie des actes de langage aux systèmes multi-agents*, chapter La structure logique des dialogues intelligents, pages 61–100. L'interdiscilplinaire informatique.

[Velasquez, 1997] Velasquez, J. (1997). Modeling emotions and other motivations in synthetic agents. In *Proceedings of the AAAI Conference 1997*, pages 10–15. Providence, RI.

[Vignaux, 1991] Vignaux, G. (1991). *Les Sciences Cognitives : une Introduction*. le Livre de Poche : biblio essais, Paris, La Découverte edition.

[Walton and Krabbe, 1995] Walton, D. and Krabbe, E. (1995). *Commitment in Dialogue*. Suny Press.

[Werner, 1992] Werner, E. (1992). The design of multi-agent systems. In Werner, E. and Demazeau, Y., editors, *Decentralized AI 3 — Proceedings of the Third European Workshop on Modelling Autonomous Agents in a Multi-Agent World (MAAMAW-91)*, pages 3–30. Elsevier Science Publishers B.V. : Amsterdam, The Netherlands.

[Wickland and Brehm, 1976] Wickland, R. and Brehm, J. (1976). *Perspectives on Cognitive Dissonance*. NY : Halsted Press.

[Winograd and Flores, 1986] Winograd, T. and Flores, F. (1986). *Understanding Computers and Cognition : A New Foundation for Design*. Ablex, Norwood, NJ.

[Wittgenstein, 1953] Wittgenstein, L. (1953). *Philosophical Investigations*. MacMillan.

[Wooldridge and Jennings, 1994] Wooldridge, M. and Jennings, N. (1994). The cooperative problem solving process : A formal model. Technical report, Department of Computing, Manchester Metropolitan University, Chester St., Manchester M1 5GD, UK.

[Wooldridge et al., 2002] Wooldridge, M., Mac Burney, P., and Parsons, S. (2002). Desiderata for agent argumentation protocols. In *Procedings of the Autonomous Agent and Multi-Agent Systems Conference(AAMAS'02)*, pages 402–409.