



# Doit-on dire la vérité pour se comprendre? Principes d'un modèle collaboratif du dialogue basé sur la notion d'acceptation

Sylvie Saget, Marc Guyomard

## ► To cite this version:

Sylvie Saget, Marc Guyomard. Doit-on dire la vérité pour se comprendre? Principes d'un modèle collaboratif du dialogue basé sur la notion d'acceptation. 2007. hal-00192048

HAL Id: hal-00192048

<https://hal.archives-ouvertes.fr/hal-00192048>

Preprint submitted on 26 Nov 2007

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

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

# Doit-on dire la vérité pour se comprendre ?

## Principes d'un modèle collaboratif du dialogue basé sur la notion d'acceptation

Sylvie Saget

Marc Guyomard

Sylvie.Saget@enssat.fr

Marc.Guyomard@enssat.fr

Institut de Recherche en Informatique et Systèmes Aléatoires (LLI-ENSSAT)  
Université de Rennes 1  
22305 LANNION – FRANCE

### Résumé :

Garrod et Pickering ([16]) prétendent que l'étude fine de la nature collaborative du dialogue doit mener à lever ou à modifier un ensemble d'hypothèses fondamentales qui sont sources de complexité pour les modèles de dialogue. Dans des travaux précédents, il a été proposé que l'une de ces hypothèses est la sincérité des partenaires de dialogue ([35, 37]). Le traitement des énoncés, plus précisément des références, y est ainsi vu comme un processus orienté par un but. L'objet de cet article est de présenter les principes d'un modèle collaboratif du dialogue basé sur une attitude mentale dépendante du contexte : l'acceptation.

**Mots-clés :** Modèle de dialogue, pragmatique, collaboration, attitude mentale, croyance, acceptation

### Abstract:

Garrod and Pickering ([16]) claim that considering spoken dialogue as a collaborative activity must lead to avoid or to modify the fundamental hypothesis which are responsible for complexity limitations of existing spoken dialogue systems. In previous works, we argue that one of these hypothesis is the sincerity of dialogue partners ([35, 37]). Utterance treatment, more precisely reference treatment, is then viewed as a goal-oriented process. The aim of this article is to present the principles of a collaborative model of dialogue based on a context-dependant mental attitude : acceptance.

**Keywords:** Dialogue model, pragmatics, collaboration, mental attitude, belief, acceptance

## Introduction

Le dialogue tient un rôle crucial dans la réalisation des activités collaboratives car il permet d'établir et de maintenir les coordinations nécessaires entre les membres d'une équipe. En outre, dans la lignée de nombreux travaux fondateurs

([9, 32]), le dialogue *en lui-même* est considéré comme une activité collaborative. Le dialogue se caractérise par le but partagé entre les partenaires du dialogue de se comprendre mutuellement. Voir la modélisation du dialogue comme la modélisation d'une activité collaborative permet, entre autres, de fournir un modèle explicatif à de nombreux phénomènes dialogiques tels que les sous-dialogues de négociation portant sur l'interaction (négociation sémantique, négociation des références, etc.). La modélisation fine des phénomènes en question permet d'accroître la flexibilité des systèmes de dialogue.

Comment formaliser le dialogue comme étant explicitement une activité collaborative (i.e. comment formaliser un *modèle collaboratif du dialogue*)? Il s'agit là d'un problème particulièrement difficile et ceci pour deux raisons. Tout d'abord, de manière générale, la définition et la modélisation des phénomènes intervenant dans les activités collaboratives ne sont pas encore clairement définis ([23]). Ensuite, l'étude de la nature collaborative du dialogue est relativement récente ([12, 9]), de nombreux aspects sont toujours à définir et des problèmes difficiles restent à résoudre. En particulier :

- La spécification du but partagé de se comprendre, de manière à pouvoir être intégré dans un modèle de dialogue ([15, 24, 41]),
- La spécification du critère de suffisance ("grounding criterion") : ce critère est

fortement contextuel. Il varie selon les activités, les individus et il varie également au cours du dialogue ([7]).

Que signifie "atteindre une inter-compréhension"? L'un des buts des partenaires du dialogue est d'atteindre une inter-compréhension vis-à-vis de l'intention communicative du locuteur. Cette dernière devant être reconnue par l'auditeur suite à la compréhension et à l'interprétation de l'énoncé prononcé. L'intention communicative correspond à "ce que le locuteur veut dire" ("speaker's meaning", [18]), elle doit être distinguée clairement de la sémantique de l'énoncé (ou encore du contenu propositionnel de l'acte de dialogue correspondant) ([15]).

Quelle est la relation entre l'intention communicative du locuteur et la sémantique de l'énoncé? De la maxime de Qualité de Grice ([18]), en passant par les notions de présuppositions et de terrain commun de Stalnaker ([38, 39]) jusqu'au modèle collaboratif du dialogue de Clark ([9, 12]), on exploite l'hypothèse que l'intention communicative et la sémantique sont liées par une relation de valeur de vérité (Hypothèse de Sincérité). Dans le modèle collaboratif du dialogue proposé par Clark, cela se traduit par l'usage du terrain commun établi ou supposé comme étant un pré-requis au succès de l'inter-compréhension des partenaires du dialogue. Par exemple, si le locuteur veut référer à un objet  $o$  en prononçant une expression référentielle dont la sémantique est  $i.x.descr(x)$ , alors un pré-requis au succès de cette référence est :

$MB_{LA}(descr(o))$ , où :

- $MB_{i,j}(\phi)$  correspond à " $\phi$  est une croyance mutuelle entre les agents  $i$  et  $j$ , du point de vue de  $i$ ",
- On peut définir formellement la croyance mutuelle par :  
 $MB_{i,j}(\phi) \equiv B_i(\phi \wedge MB_{j,i}(\phi))$ ,
- $L$  dénote le locuteur et  $A$  l'auditeur.

Le terrain commun entre les partenaires du dialogue, modélisé par un état épistémique au niveau des modèles formels du dialogue, est alors considéré comme le contexte à prendre en compte lors de la génération et l'interprétation des énoncés.

Cette hypothèse de sincérité et cette vision de l'usage du terrain commun dans le traitement des énoncés est source de nombreuses limitations et incohérences théoriques et pratiques. Ce constat a mené à les remettre en question. Nous proposons, comme alternative, de considérer le dialogue comme étant dirigé par le *but* de se comprendre et non par la *vérité* ([35, 37, 36]). Un énoncé (par extension, sa sémantique) est alors considéré comme un *outil* permettant d'atteindre ce but. Afin de prendre en compte cette proposition d'un point de vue formel, l'acceptation (attitude mentale dépendante du contexte) d'un énoncé comme étant utile pour permettre l'inter-compréhension est appropriée.

L'objet de cet article est de présenter les principes d'un modèle collaboratif de dialogue basé sur la notion d'acceptation, en se penchant tout particulièrement sur le cas des références. Dans la première partie, les fondements de cette approche sont explicités. Dans la seconde partie, les principes du modèle collaboratif de dialogue sont présentés.

## 1 Quand la sincérité n'est pas de mise

### 1.1 Limites et incohérences

Le modèle collaboratif du dialogue de Clark [12, 9] sert de référence à la grande majorité des modèles formels de dialogue ([42, 20]) utilisés au niveau des systèmes de dialogue. Dans ce type de modèle, le terrain commun, représenté par un état épistémique, est utilisé dans le traitement des énoncés (et donc des références) afin

de garantir le succès des actes de dialogue. Cette hypothèse conduit à un ensemble de problèmes et d'incohérences au niveau du modèle collaboratif du dialogue. Nous passons en revue les principaux.

**Complexité.**

1. Les partenaires du dialogue ont des états mentaux dissemblables et souvent des états épistémiques différents. De plus, ils n'ont qu'un accès indirect aux états mentaux des autres participants. Il existe donc des incertitudes et des différences de point de vue vis-à-vis de leur terrain commun. Les partenaires du dialogue doivent alors s'adapter au point de vue de l'autre lors d'un dialogue ([38, 39]). Taylor et al. [40] ont montré que ceci est difficile à mettre en application. Ceci en raison de la complexité de maintenir les états épistémiques requis et en raison de la complexité des processus de raisonnement associés. En conséquence, la plupart des systèmes de dialogue ne supportent pas les croyances imbriqués <sup>1</sup> au delà du second niveau d'imbrication.
2. L'analyse sémantique d'un énoncé peut être également une source de complexité. C'est le cas en particulier des énoncés vagues ([33]).

**Incohérence vis-à-vis du comportement humain.**

De prime abord une telle complexité semble incompatible avec l'efficacité des êtres humains au cours d'un dialogue. Ensuite, la considération de Clark vis-à-vis de l'usage du terrain commun ("Initial Design Model") dans la génération des énoncés a été remise en question par de nombreuses études ([21] : "Monitoring & Adjustment model", [2]). Il y est montré que les êtres humains peuvent faire preuve d'une attitude beaucoup plus "égoïste" <sup>2</sup> et

ceci s'accroît en particulier lorsque les contraintes temporelles croissent.

Ainsi, les humains utilisent différentes stratégies (considération de leur croyance propres, du terrain commun, du point de vue de l'auditeur de leur terrain commun, etc.) plus ou moins coûteuses lors de la génération des énoncés.

**Difficultés pour caractériser la rationalité des partenaires du dialogue.**

La rationalité des partenaires du dialogue est caractérisée notamment par la cohérence de leurs croyances et par leur sincérité ([27]). Or, il est très fréquent que le locuteur (ou l'auditeur) génère (ou acquiesce) un énoncé présupposant une croyance contradictoire avec ses propres croyances ([11]). Dans le modèle de Clark, ceci s'explique par le processus d'adaptation, par l'un des partenaires du dialogue, à la différence de point de vue sur le terrain commun qu'il a avec son interlocuteur.

Malgré tout, comment caractériser ce type de comportement comme étant rationnel alors qu'il enfreint l'hypothèse de sincérité ? Il existe deux grandes catégories de solutions :

1. Caractériser ce comportement par un raisonnement restreint à un contexte local ([4]). Ce type de méthode n'est pas utilisé en pratique en raison de la complexité requise.
2. Ajouter une attitude mentale supplémentaire, comme l'a proposé Stalnaker [39], dans une caractérisation du terrain commun nécessaire<sup>3</sup> :  
*"Une catégorie d'attitudes propositionnelles (...) qui inclut la croyance, mais également d'autres attitudes (la présomption, l'hypothèse, l'acceptation au profit d'une argumentation ou d'une requête) qui diffèrent de la croyance et les unes des autres. Accepter une proposition c'est accepter de*

<sup>1</sup>"Nested beliefs", permettant de modéliser en particulier le point de vue d'un agent sur l'état épistémique d'un autre agent.

<sup>2</sup>i.e. tenir compte de l'ensemble de ses croyances sans se soucier si elles sont partagées par exemple.

<sup>3</sup>"A category of propositional attitudes (...) that includes belief, but also some attitudes (presumption, assumption, acceptance for the purposes of argument or an inquiry) that contrast with belief and with each other. To accept a proposition is to treat it as true for some reason."

*la traiter comme si elle était vraie."*

Le but est ici de séparer clairement l'état épistémique d'un agent des pré-supposition qu'il va "accepter", afin que la contrainte de cohérence entre les deux soit levée. Nous proposons d'aller encore plus loin, en utilisant une autre définition de l'acceptation qui permet d'éviter "de faire semblant d'être sincère".

**Incohérence entre les propriétés des états épistémiques et celles des éléments identifiés du Terrain Commun Conversationnel.** Dans le cas particulier du traitement des références, même pour des tâches quotidiennes, nécessitant des objets qui possèdent des noms communs pour y faire référence, il existe un large panel de manières possibles pour y faire référence. Afin de garantir une intercompréhension, les humains<sup>4</sup> "*associent les objets avec des expressions (et les perspectives qu'elles transcrivent), autrement dit en faisant des pactes conceptuels, ou des accords temporaires et flexibles de considérer un objet d'une façon particulière*" [5]. Ces pactes conceptuels font partie du Terrain Commun Conversationnel (TCC, le Terrain Commun spécifique au dialogue). Selon le modèle collaboratif du dialogue de Clark, le TCC doit être modélisé par des croyances partagées ou par un état épistémique plus faible. Mais les propriétés identifiées des pactes conceptuels (flexibilité importante, durée limitée à celle du dialogue, spécifique à un partenaire) ne correspondent pas avec les propriétés de la croyance (idéal d'agglomération, non limitée dans le temps, indépendante du contexte, etc.) ([35, 37]).

## 1.2 Se comprendre : une activité orientée par des buts

**Les buts liés à l'intercompréhension.** Kronfeld, [25], définit les buts des actes de ré-

<sup>4</sup>"associate objects with expressions (and the perspectives they encode), or else from achieving conceptual pacts, or temporary, flexible agreements to view an object in a particular way"

férence comme suit :

1. Le but littéral ("Literal goal") : en résultat de la reconnaissance par l'auditeur d'un groupe nominal, l'auditeur générera une interprétation qui détermine le même objet.
2. Le but lié à la tâche ("Discourse purpose") : l'auditeur doit appliquer différentes opérations à l'ensemble individualisant de telle sorte qu'il respectera les contraintes d'identification appropriées.

Le but lié à la tâche est, pour les dialogues finalisés, un pont reliant le dialogue en lui-même à la tâche motivant le dialogue (l'activité de base). Les contraintes d'identification correspondent aux propriétés de l'objet que le destinataire doit connaître afin d'être en mesure de réaliser la tâche. C'est par ce biais que la valeur de vérité d'un énoncé peut être imposée ([36]).

En outre, selon Clark et al. [10], le but littéral doit être décomposé en deux buts interdépendants :

- 1.1. Identification : le locuteur tente de faire en sorte que le destinataire identifie un référent particulier à partir d'une description particulière.
- 1.2. Grounding : le locuteur et le destinataire tentent d'établir le fait que le destinataire a identifié le référent suffisamment bien par rapport à leurs buts courants.

Une expression référentielle, et par extension sa sémantique, n'est donc qu'un *outil* dans la réalisation de ces buts. La sincérité<sup>5</sup> peut être utile mais pas nécessaire. Elle est coûteuse et peut être un mauvais choix si les contraintes temporelles sont strictes. Sans compter que le but n'est pas ici d'établir la vérité vis-à-vis des propriétés d'un objet, mais d'être capable de se comprendre suffisamment (sur l'objet dont

<sup>5</sup>Autrement dit, la croyance de l'existence d'un unique objet (dans le cas de descriptions définies) respectant effectivement la description spécifiée par l'expression référentielle.

il est question) afin de réaliser la tâche motivant le dialogue, dans le cadre de dialogues finalisés.

**Le dialogue finalisé : une activité subordonnée.**

En effet, dans le cadre des dialogues finalisés, il est important de faire la distinction entre deux types de sous-activités : les activités qui sont des *sous-parties* d'une autre activité (*activité de base*) et les (sous-)activités *subordonnées*. Les sous-activités qui sont des *sous-parties* d'une autre activité reflètent la compositionnalité des activités de base. Les (sous-)activités *subordonnées* désignent les sous-activités *au service* d'une autre activité, telles que la planification, la résolution de problèmes, l'interaction avec d'autres agents (dialogues finalisés), etc. Ces activités subordonnées comprennent les différentes fonctions remplies par la macrocognition définies par Klein et al. [22].

D'un point de vue logique, la rationalité des agents est traduite par un état mental cohérent et par la notion d'action rationnelle [14, 34]. Par exemple, les croyances et les intentions forment un ensemble consistant et les actions des agents sont également consistantes avec leurs croyances et leurs intentions.

Au premier abord, la cohérence d'une action d'un agent avec ses croyances semble être irréfutable. Mais, étant donné que le succès d'une activité subordonnée est gouverné par une généralisation du critère de suffisance et les contraintes temporelles, on peut raisonnablement prétendre que la rationalité d'un agent engagé dans une activité subordonnée n'implique pas de manière stricte la cohérence entre les actions qui font partie de l'activité subordonnée et l'état épistémique des agents concernés.

Cette distinction est importante tant au niveau de la distinction des différents terrains communs mis en jeu lorsque la

tâche est de nature collaborative ([35])<sup>6</sup> qu'au niveau des différents degrés d'effort déployés pour chacune des tâches ([8]).

**2 Un modèle de dialogue basé sur la notion d'acceptation**

**2.1 L'acceptation : une attitude mentale dépendante du contexte**

L'interaction entre les différents domaines concernés par les activités collaboratives peut être une source de richesses et d'avancées significatives. En particulier, en philosophie, la nécessité de distinguer l'attitude mentale *dépendante du contexte* (pragmatique) qu'est l'acceptation, de l'attitude mentale *indépendante du contexte* qu'est la croyance a été (re)mise en avant par de nombreux philosophes tels que Cohen ([13]).

L'acceptation et la croyance diffèrent par leur rôle fonctionnel : la croyance est orientée par la vérité, alors que l'acceptation est orientée par un (des) but(s) [31] :

<b>Croyance</b>	vs	<b>Acceptation</b>
$B_i(\phi)$	vs	$Acc_i(\psi)$
<b>orientée par la vérité</b>	vs	<b>orientée par un but</b>
$\phi$ est vraie	vs	$\psi$ est appropriée pour le succès d'un certain but
↓		↓
involontaire	vs	volontaire
graduelle	vs	tout-ou-rien
indépendante du contexte	vs	dépendante du contexte

Ainsi, la croyance et l'acceptation ont différents critères de correction a posteriori :

<sup>6</sup> Il faut alors distinguer le terrain commun lié à la tâche du terrain commun spécifique au dialogue

$B_i(\phi)$  **est correcte** si elle se révèle correspondre à l'état actuel du monde

$Acc_i(\psi)$  **est correcte** si les choix faits et les actions réalisées en se basant sur ce qui est accepté font que le choix de  $\psi$  se révèlent porter ses fruits

Cette distinction et les différents travaux philosophiques qui en découlent ont été mis à profit au niveau des modèles formels de l'interaction :

- Dans le domaine des SMAs, Gaudou et al. ([17]) utilisent la notion d'acceptation collective afin de modéliser les "faits publics" de manière cohérente.
- Dans le cadre des dialogues argumentatifs, Baker [1] considère que le résultat d'une négociation est une acceptation collective. Paglieri & Castelfranchi ([31]) utilisent la distinction entre croyance et acceptation afin de distinguer les argumentaires dirigés par la vérité des argumentaires dirigés par un but.
- Au niveau du dialogue en lui-même, dans le cadre d'un modèle collaboratif du dialogue, nous avons mis en avant l'intérêt de cette distinction pour séparer clairement la sémantique d'un énoncé de l'intention communicative du locuteur ([35, 37, 36]).

Dans les derniers travaux, la notion d'acceptation est utilisée pour traduire le fait qu'un énoncé est vu comme un *outil* permettant de déterminer l'intention communicative du locuteur. La négociation de l'interprétation est alors considérée comme la *co-construction* d'un outil linguistique permettant aux partenaires du dialogue d'atteindre un niveau d'intercompréhension suffisant vis-à-vis de leur activité courante. Le résultat de cette co-construction est modélisé par une acceptation collective.

## 2.2 Le modèle rationnel de référence

Les modèles rationnels de dialogue, basés principalement sur les travaux de Cohen

et Levesque [14], peuvent être considérés comme une reformulation logique des modèles de dialogue différentiels (modèles à base de plans). Ils intègrent, en plus, une formalisation précise des états mentaux des partenaires du dialogue (leurs croyances, leurs choix (ou désirs) et leurs intentions) et de l'équilibre rationnel qui les relie entre eux et qui relie les attitudes mentales d'un agent avec ses actions. En outre, les préconditions et les effets des actes de dialogue sont exprimés en fonction des états mentaux des partenaires du dialogue.

Le modèle choisi, fondé sur le modèle rationnel proposé par Sadek [34], se base sur un ensemble de principes (de schémas d'axiomes) dont les actes de dialogue découlent. Un système de dialogue est alors considéré comme un agent rationnel qui a une attitude coopérative envers les autres agents (tel que l'utilisateur de ce système). Ce système étant capable de communiquer avec les autres agents.

Les états mentaux (croyances, intentions, etc.) et les actions sont formalisées dans une logique modale du premier ordre. Dans la suite de cet article, les symboles  $\neg, \wedge, \vee, \Rightarrow$  désignent les connecteurs de la logique classique (respectivement la négation, la conjonction, la disjonction et l'implication);  $\forall, \exists$  désignent les quantificateurs universel et existentiel;  $p$  désigne une formule close dénotant une proposition;  $i, j$  dénotent des agents,  $\phi$  est un schéma d'axiome et  $\psi$  est une formule complexe. Nous avons seulement besoin d'introduire ici deux attitudes mentales, la croyance et l'intention :

$B_i(p)$  signifie "*i* croit que *p* est vraie"

$I_i(p)$  signifie "*i* a l'intention d'établir *p*"

Les expressions d'actions sont formées à partir d'actes primitifs, par  $(a_1; a_2)$  qui désigne une action séquentielle (où  $a_1$  et  $a_2$  sont des expressions d'actions) et par  $(a_1|a_2)$  qui désigne le choix non

déterministe.

$Faisable(a, p) : "a$  peut être réalisée, et si elle l'est, alors  $p$  sera vraie"

$Faisable(a) = Faisable(a, true)$

$Fait(a) : "a$  vient d'être réalisée"

Le modèle des actes communicatifs est :

$\langle i, TypeActeCommunicatif(j, \phi) \rangle$

PF (Préconditions de Faisabilité) : les conditions qui doivent être satisfaites afin de planifier l'action

EP (Effet Perlocutoire) : la raison pour laquelle l'acte a été sélectionné

Par exemple, le modèle de l'acte communicatif de " $i$  informant  $j$  que  $\phi$ " est :

$\langle i, INFORM(j, \phi) \rangle$

PF :  $B_i(\phi) \wedge \neg B_i(B_j(\phi))$

EP :  $B_j(\phi)$

Dans ce modèle, les processus de génération et d'interprétation des énoncés sont vus comme des activités individuelles isolées. Ainsi, l'interprétation du locuteur est supposée réalisée et correcte dès lors que l'énoncé est prononcé. Le dialogue et la référence ne sont pas considérés comme étant des activités collaboratives.

### 2.3 Modèle collaboratif du dialogue

Il n'existe pas de consensus sur la définition d'une activité collaborative. Nous considérons ici qu'un groupe d'agents est engagé dans une activité collaborative dès lors qu'ils partagent une intention collective.

Comment formaliser le dialogue comme étant une activité collaborative? Par un modèle rationnel intégrant la notion d'intention collective permettant de spécifier explicitement l'intention des partenaires du dialogue de se comprendre mutuellement? Ceci est à proscrire dans une premier temps car, comme stipulé dans l'introduction, les formalisations existantes de ce but sont impossibles à mettre

en oeuvre ([15, 24, 41]).

Faut-il intégrer la notion de plan partagé ?

Le concept de plan partagé est utilisé pour modéliser le plan correspondant à la réalisation concrète d'une activité collaborative. Ce plan doit être mutuellement connu à un certain degré, afin que les membres de l'équipe puissent coordonner leurs actions. Ce paradigme est utilisé dans les systèmes de dialogue pour modéliser une tâche (motivant le dialogue) qui est de nature collaborative ([19, 29]) et *non* pour modéliser la nature collaborative du dialogue.

De notre point de vue, l'utilisation d'un plan partagé pour spécifier la nature collaborative du dialogue n'est pas nécessaire car le dialogue est une activité acquise au cours de l'enfance, elle fait donc appel à des routines ([28]). De plus, étant une activité sociale, elle est régie par tout un ensemble de lois et de normes.

Notre modèle collaboratif du dialogue considère donc que le processus de négociation de l'interprétation, par lequel l'inter-compréhension est atteinte, est une co-construction d'un outil linguistique (d'une expression référentielle) permettant d'atteindre une inter-compréhension suffisante de l'intention communicative du locuteur. Le résultat de cette co-construction étant modélisé par une acceptation collective :

$AccColl_{LA}(refererPar(o, ix.descr(x)))$ , où :

- $AccColl_{ij}(\psi)$  correspond à " $\psi$  est collectivement acceptée par les agents  $i$  et  $j$ , du point de vue de  $i$ ",
- $refererPar(R, D)$  correspond à "le référent  $R$  est référé par la description  $D$ ",
- $o$  est une représentation mentale de  $L$  d'un objet.

Le succès d'une telle co-construction aboutit donc à une inter-compréhension :

$AccColl_{LA}(refererPar(o, ix.descr(x))) \Rightarrow MB_{LA}(I_L(refer_{LA}(o)))$ , où :



–  $I_i(\text{refer}_{ij}(o))$  correspond à "l'intention communicative de  $i$  de référer à  $o$ , le destinataire étant  $j$ ".

• L'acceptation collective est une attitude mentale intentionnelle, i.e. qui est formée suite à des actions réalisées par les agents mis en jeu :

$$((\exists \alpha, \beta \in \{i, j\} \wedge \exists a). \\ \text{Fait}(\text{Prop}_{\alpha\beta}(\psi)); a; \text{Accept}_{\beta\alpha}(\psi)) \\ \Rightarrow \text{AccColl}_{ij}(\psi))$$

où :

- $\text{Prop}_{ij}(\psi)$  correspond à " $i$  propose à  $j$  de considérer  $\psi$ ",
- $\text{Accept}_{ji}(\psi)$  correspond à " $j$  accepte de considérer  $\psi$  (vis-à-vis de  $i$ )",
- $\text{Prop}_{ij}(\psi)$  et  $\text{Accept}_{ji}(\psi)$  sont des actions individuelles,
- $a$  est une expression d'action.

• Suite à une proposition, l'autre agent est obligé de réagir en raison d'une loi sociale :

$$\text{Fait}(\text{Prop}_{i,j}(\psi)) \\ \Rightarrow (I_j(\text{Fait}((\text{Accept}_{j,i}(\psi) \\ |(\text{Prop}_{j,i}(\psi')) \\ |(\text{Prop}_{i,j}(\psi'')) \\ |(\text{Reporter}) \\ |(\text{Stopper})))))) \\ \wedge ((\psi' \neq \psi) \wedge (\psi'' \neq \psi)))$$

où :

- *Reporter* est une action individuelle qui permet de reporter la réaction, de la "mettre en suspend",
- *Stopper* est une action individuelle qui permet de sortir de la négociation.

Dans la lignée des travaux de Boella et al. [3], nous considérons que les obligations sociales en tant que pro-attitudes ne sont pas nécessaires pour formaliser le dialogue. Le locuteur procède à une coordination par anticipation. Ce phénomène est gouverné par une loi sociale, acquise au cours d'interactions sociales précédentes. Cette loi sociale est transcrite

par son usage répété en réaction à la réalisation d'une action particulière (du point de vue du locuteur) et en réaction à l'occurrence d'un événement qui est l'occurrence d'une action particulière (du point de vue de l'auditeur). Étant donné qu'une réaction est une action non intentionnelle, nous devons étendre la catégorie des actions tolérées par notre modèle de base. En effet, ce modèle ne considère que les actions intentionnelles. Nous appelons ici *actions intentionnelles* les actions d'un agent qui sont générées par une chaîne d'intentions. Dans le modèle de Sadek ([34]), elles sont générées par l'activation de l'axiome rationnel suivant :

$$I_i(p) \Rightarrow I_i(\text{Fait}(a_1 | \dots | a_n))$$

L'intention de l'agent, de réaliser un but donné, génère l'intention de réaliser l'un des actes satisfaisant les conditions suivantes :

1.  $(\exists x)B_i(a_k = x) \equiv \text{Bref}_i(a_k)$  :  
L'agent  $i$  connaît l'action  $a_k$ ,
2.  $EP_{a_k} = p$  et
3.  $\neg I_i(\neg \text{Faisable}(\text{Fait}(a_k)))$ .

Il faut donc rajouter les réactions aux actions intentionnelles. Les *réactions* d'un agent sont définies comme les actions générées par l'activation d'un axiome tel que :

$$\phi \Rightarrow I_i(\text{Fait}(a_1 | \dots | a_n))$$

où  $\phi$  résulte de la perception d'un événement ou de l'occurrence d'une action.

Considérons à présent le cas particulier des références. Tout d'abord, il faut préciser qu'un modèle collaboratif du dialogue nécessite de considérer la référence comme un acte de dialogue à part entière, [15], comme c'est le cas dans les travaux de Kronfeld [26].

Quand le locuteur veut référer à un objet, sa première tentative se concrétise par un acte de référence qui :

1. A pour effet la reconnaissance par l'auditeur de l'intention du locuteur de référer à un certain objet :

$$B_A((\exists o)I_L(refer_{ij}(o)))$$

2. A pour effet de générer chez l'auditeur l'adoption de l'intention de comprendre le locuteur, c'est-à-dire de construire ou d'identifier une représentation mentale correcte :

$$I_A(Bref_A(o')) \wedge RepMemeObj(o, o'), \text{ où :}$$

- $o$  et  $o'$  sont des représentations mentales d'objets ;
- $RepMemeObj(o, o')$  signifie que "les représentations mentales  $o$  et  $o'$  représentent le même objet".

3. Correspond à l'engagement dans un processus de négociation entre les partenaires du dialogue. En générant une expression référentielle, le locuteur propose à l'auditeur de considérer sa description comme permettant de construire ou d'identifier une représentation mentale qui soit correcte.

En raison de la loi sociale, le locuteur attend une réaction de l'auditeur à sa proposition. La loi sociale oblige également l'auditeur à réagir à la proposition du locuteur.

Ceci est synthétisé dans l'extension du modèle des actes de références [6] que nous proposons. L'acte de référence d'un agent  $i$  envers un autre agent  $j$ , en utilisant la conceptualisation  $x$  (qui correspond à la sémantique de l'expression référentielle) de référer à un objet  $o$  est formalisé comme suit :

$$\langle i, REFER(j, x, o) \rangle$$

$$PF : I_i(refer_{ij}(o)) \wedge Bref_i(o)$$

$$EP : B_j((\exists o)I_i(refer_{ij}(o)))$$

$$\wedge I_j(Bref_j(o'))$$

$$\wedge RepMemeObj(o, o')$$

$$\wedge Fait(Prop_{i,j}(refererPar(o, x)))$$

$$\wedge B_j(Fait(Prop_{i,j}(refererPar(o, x))))$$

## 2.4 Formalisation de l'acceptation

L'acceptation collective est formée par un processus de négociation qui fait intervenir les actions  $Prop_{ij}(\psi)$  et  $Accept_{ij}(\psi)$ . La précondition de ces deux actions est identique : l'acceptation individuelle de  $\psi$  par  $i$ . Dans cette section, nous proposons les premiers éléments de formalisation de l'opérateur modal d'acceptation individuelle.

**Cadre philosophique.** Il n'existe pas de consensus sur la caractérisation de l'acceptation<sup>7</sup>. Nous adoptons ici la vision de Paglieri, introduite en 2.1., en raison de sa simplicité et de notre adéquation avec son point de vue.

Par définition, les acceptations sont considérées comme étant le seul arrière-plan cognitif à la délibération. Le modèle traditionnel BDI devient donc ([30], p.36) :

$$(\text{Desire} + \text{Acceptation}) \Rightarrow \text{Intention}$$

Les croyances tiennent un rôle indirect et non nécessaire dans le raisonnement pratique. En effet, elles sont une raison fréquente d'accepter.

**Formalisation.** Qu'est-ce qui est accepté ? Tout ce qui peut permettre et/ou faire avancer le raisonnement va être soumis à l'acceptation. D'après Paglieri ([30]), cela comprend :

- les relations qui permettent d'arriver à ses fins qui instancient une action adéquate pour un but,
- les préconditions des actions que l'on a l'intention de réaliser,
- les préconditions nécessaires pour avoir une intention particulière, l'intention n'est pas déjà réalisée et est réalisable,
- les résultats des actions, afin de pouvoir évaluer si le but a été atteint ou si on doit continuer dans la ligne d'action

<sup>7</sup>Pour plus de détails sur ce débat, consulter [30], Chap.1.

courante.

À notre connaissance il n'existe pas de caractérisation logique de l'acceptation telle qu'elle est définie ci-dessus. Voici les premiers éléments d'une proposition de formalisation :

- $Acc_i(\psi, \phi)$  signifie que "L'agent  $i$  accepte  $\psi$  dans l'optique de réaliser  $\phi$ ", où :
- $Acc_i(\psi, \phi)$   
 $\equiv Acc_i(\psi) \Rightarrow B_i(\exists e. Faisable(e, \phi))$

Notre opérateur modal d'acceptation individuelle respecte l'axiomatique suivante (axiomatique KD45) :

- (K)  $(Acc_i(\psi) \wedge Acc_i(\psi \Rightarrow \psi')) \Rightarrow Acc_i(\psi')$
- (D)  $Acc_i(\psi) \Rightarrow \neg Acc_i(\neg \psi)$
- (5)  $\neg Acc_i(\psi) \Rightarrow Acc_i(\neg Acc_i(\psi))$
- (4)  $Acc_i(\psi) \Rightarrow Acc_i(Acc_i(\psi))$

La relation d'accessibilité est ainsi définie comme étant  $R_{Acc}$  transitive, euclidienne et sérielle. Le système axiomatique proposé est donc complet et adéquat. Il est également intéressant de remarquer que notre opérateur a la même axiomatique que celle couramment utilisée pour la croyance. De plus,  $R_{Acc}$  et  $R_B$  ne sont ni réflexives, ni symétriques. Cependant, la raison en diffère :  $R_{Acc}$  n'est ni sérielle, ni réflexive *par définition* car le rôle fonctionnel de l'acceptation n'a rien à voir avec la réalité objective.

## Conclusion

Dans cette article, nous avons proposé une alternative à la considération du traitement des énoncés comme étant des processus orientés par la vérité (et le terrain commun) en mettant au premier plan le caractère finalisé de ces processus. La distinction entre les attitudes mentales que sont la croyance et l'acceptation est nécessaire pour prendre en compte cette alternative dans un modèle rationnel de dialogue.

Ces travaux sont porteurs de nombreuses perspectives. Au niveau des extensions :

- La génération d'un énoncé est vu comme la détermination d'un outil acceptable, au niveau individuel, c'est-à-dire *utile* pour atteindre une intercompréhension suffisante,
- Le processus dual d'interprétation d'un énoncé est vu comme l'évaluation de l'acceptabilité de l'énoncé par le destinataire.

Incorporer l'acceptation individuelle, et étudier ses propriétés logiques, au niveau du modèle collaboratif du dialogue consiste donc à faire le lien avec les processus de traitement des énoncés.

Ces travaux procurent également des perspectives intéressantes vis-à-vis de la prise en compte du contexte, l'acceptation étant une attitude mentale dépendante du contexte :

- Dans le modèle collaboratif du dialogue décrit ci-dessus, le contexte qui sert de base à l'interprétation des énoncés n'est plus systématiquement l'état épistémique (les croyances) des participants. Il peut donc être élargi en incorporant le contexte motivationnel, les contraintes temporelles, etc.
- L'hypothèse de sincérité étant levée, les partenaires du dialogue ont le choix entre différentes stratégies possibles : considération uniquement de leurs propres croyances (attitude "égoïste"), de celles de son interlocuteur (attitude coopérative), de baser une interprétation sur la reconnaissance des mots-clés, etc.

## Remerciements

Ces travaux sont financés par la subvention A3CB22 / 2004 96 70 de la Région Bretagne.

## Références

- [1] M. J. Baker. A model for negotiation in teaching-learning dialogues. *Journal of Artificial Intelligence in Education*, 5(2) :199–254, 1994.

- [2] E. G. Bard, A. H. Anderson, Y. Chen, H. Nicholson, and C. Havard. Let's you do that : Enquiries into the cognitive burdens of dialogue. In *Proceedings of DIALOR'05*, 2005.
- [3] G. Boella, R. Damiano, and L. Lesmo. Social goals in conversational cooperation. In *Proceedings of the First SIGdial Workshop on Discourse and Dialogue*, pages 84–93. ACL, Somerset, New Jersey, 2000.
- [4] A. Bonomi. Truth and reference in context. *Journal of Semantics*, 23(2) :107–134, 2006.
- [5] S. E. Brennan and H. H. Clark. Conceptual pacts and lexical choice in conversation. *Journal of Experimental Psychology : Learning, Memory and Cognition*, 22 :482–1493, 1996.
- [6] P. Bretier, F. Panaget, and M. D. Sadek. Integrating linguistic capabilities into the formal model of rational agent : Application to cooperative spoken dialogue. In *AAAI-95, Fall Symposium of Rational Agency*, Stanford, MA, 1995.
- [7] J. E. Cahn and S. E. Brennan. A psychological model of grounding and repair in dialog. In *Proceedings of the AAAI Fall Symposium on Psychological Models of Communication in Collaborative Systems*, pages 25–33, 1999.
- [8] M. Cherubini and J. van der Pol. Grounding is not shared understanding : Distinguishing grounding at an utterance and knowledge level. In *Proceedings of CONTEXT'05*, 2005.
- [9] H. H. Clark. *Using language*. Cambridge University Press, Cambridge, UK, 1996.
- [10] H. H. Clark and A. Bangerter. Changing conceptions of reference. In I. Noveck & D. Sperber, editor, *Experimental pragmatics*, pages 25–49. Palgrave Macmillan, Basingstoke, England, 2004.
- [11] H. H. Clark and C. R. Marshall. Definite reference and mutual knowledge. In *Elements of discourse understanding*, pages 10–63. Cambridge University Press, Cambridge, 1981.
- [12] H. H. Clark and D. Wilkes-Gibbs. Referring as a collaborative process. *Cognition*, 22 :1–39, 1986.
- [13] J. L. Cohen. *An Essay on Belief and Acceptance*. Oxford University Press, Oxford, 1992.
- [14] P. R. Cohen and H. J. Levesque. Rational interaction as the basis for communication. In *Intentions in Communication*, pages 221–256. MIT Press, Cambridge, MA, 1990.
- [15] P. R. Cohen and H. J. Levesque. Preliminaries to a collaborative model of dialogue. *Speech Communication*, 15 :265–274, 1994.
- [16] S. Garrod and M. J. Pickering. Why is conversation so easy? *Trends in Cognitive Sciences*, 8 :8–11, 2004.
- [17] B. Gaudou, A. Herzig, and D. Longin. A logical framework for grounding-based dialogue analysis. In *IJCAI'05, Proceedings of LCMAS 2005, Edinburgh*, 2005.
- [18] P. Grice. Logic and conversation. In *Syntax and semantics*. Academic Press, 1975.
- [19] B. J. Grosz and C. Sidner. Plans for discourse. In *Intentions in Communication*. The MIT Press, 1990.
- [20] P. A. Heeman and G. Hirst. Collaborating on referring expressions. *Computational Linguistics*, 21(3) :351–382, 1995.
- [21] W. S. Horton and B. Keysar. When do speakers take into account common ground? *Cognition*, 59 :91–117, 1996.
- [22] G. Klein, K. G. Ross, B. M. Moon, D. E. Klein, R. R. Hoffman, and

- E. Hollnagel. Macro-cognition. *IEEE Intelligent Systems*, 18(3) :81–85, 2003.
- [23] G. Klein, D. D. Woods, J. Bradshaw, R. R. Hoffman, and P. J. Fel-tovich. Ten challenges for making automation a 'team player' in joint human-agent activity. *IEEE Intelligent Systems*, pages 91–95, November/December 2004.
- [24] K. Korta. Mental states in conversation. Technical Report ILCLI-95-LIC-2, ILCLI, Donostia, 1995.
- [25] A. Kronfeld. Goals of referring acts. *Proceedings of TINLAP-3*, pages 143–149, 1987.
- [26] A. Kronfeld. *Reference and Compu-tation : An Essay in Applied Philosophy of Language*. Cambridge University Press, 1990.
- [27] M. Lee. Rationality, cooperation and conversational implicature. In *Proceedings of the AICS'97*, 1997.
- [28] A. N. Leont'ev. *Activity, Consciousness, Personality*. Englewood Cliffs, NJ, Prentice Hall, 1978.
- [29] K. Lochbaum. The use of knowl-edge preconditions in language pro-cessing. In *Proceedings of IJCAI'95*.
- [30] F. Paglieri. *Belief dynamics : From formal models to cognitive architec-tures, and back again*. PhD thesis, University of Siena, 2006.
- [31] F. Paglieri and C. Castelfranchi. Belief and acceptance in argumentation. Towards an epistemological tax-onomy of the uses of argument. In *Proceedings of ISSA'06*, 2006.
- [32] M. J. Pickering and S. Garrod. Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences*, 27(169-225), 2004.
- [33] G. Pitel and J. P. Sansonnet. A dif-ferential representation of predicates for extensional reference resolution. In *Proceeding of the 2003 Interna-tional Symposium on Reference Res-olution and its Application to Ques-tions Answering and Summurization*, Venice, Italy, 2003.
- [34] M. D. Sadek. Communication theory = rationality principles + commu-nicative act models. In *AAAI-94, Workshop on Planning for Interagent Communication*, 1994.
- [35] S. Saget. In favour of collective ac-ceptance : Studies on goal-oriented dialogues. *Collective Intentionality V*, Helsinki, Finland, 2006.
- [36] S. Saget. Using collective accep-tance for modelling the conversational common ground : Consequences on referent representation and on re-ference treatment. In *IJCAI-07, Pro-ceedings of the Workshop on Knowl-edge and Reasoning in Practical Dialog Systems*, pages 55–58, Hyde-rabad, India, 2007.
- [37] S. Saget and M. Guyomard. Goal-oriented dialog as a collaborative sub-ordinated activity involving colla-borative acceptance. In *Proceedings of Brandial'06*, pages 131–138, Uni-versity of Potsdam, Germany, 2006.
- [38] R. Stalnaker. Pragmatic presupposi-tions. In *Semantics and Philosophy*. New York University Press, 1974.
- [39] R. Stalnaker. Common ground. *Linguistics and Philosophy*, 25 :701–721, 2002.
- [40] J. A. Taylor, J. Carletta, and C. Mel-lish. Requirements for belief models in cooperative dialogue. *User Mo-deling and User-Adapted Interaction*, 6(1) :23–68, 1996.
- [41] M. Tirassa. Mental states in commu-nication. In *Proceedings of ECCS'97*, pages 103–114, Manchester, UK, 1997.
- [42] D. Traum. *A computational theory of grounding in natural language conversation*. PhD thesis, University of Rochester, 1994.