



HAL
open science

Une formalisation de la détection des actes de langage indirects

Rémi Faure, Andreas Herzig, Dominique Longin

► **To cite this version:**

Rémi Faure, Andreas Herzig, Dominique Longin. Une formalisation de la détection des actes de langage indirects. [Rapport de recherche] IRIT/2000-26-R, IRIT : Institut de Recherche en Informatique de Toulouse, France. 2000. hal-03523414

HAL Id: hal-03523414

<https://hal.science/hal-03523414v1>

Submitted on 12 Jan 2022

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.

Une formalisation de la détection des actes de langage indirects

Rémi Faure, Andreas Herzig et Dominique Longin

Institut de Recherche en Informatique de Toulouse
CNRS – Université Paul Sabatier
118 Route de Narbonne, F-31062 Toulouse Cedex 4
mailto: {faure,herzig,longin}@irit.fr
<https://www.irit.fr/departement/intelligence-artificielle/equipe-lilac/>

RAPPORT DE RECHERCHE IRIT/2000-26-R

Juillet 2000

Mots Clés: Actes de langage, logique épistémique dynamique, représentation (croyance, intention, action).

Résumé

En nous basant sur de récents travaux issus de la linguistique et portant sur la théorie des actes de langage, nous tentons de formaliser le calcul d'*effets ramifiés* de certains actes de langage (nous nous restreignons à l'étude des directifs) dans le cadre d'une logique épistémique dynamique. Bien que le point de vue soit différent, ce que nous appelons ici « effets ramifiés » est à rapprocher des effets des *actes indirects* de la théorie des actes de langage.

Notre approche peut se décomposer en deux étapes distinctes : tout d'abord déterminer si l'acte littéral accompli est susceptible de générer des effets ramifiés (et si oui, lesquels) ; puis tenter de déduire ces effets proprement dit, ce qui ne sera possible que dans un contexte d'énonciation particulier (contexte qui fournit donc un critère d'inférence).

Cette approche est formalisée au sein d'une théorie non logique.

Table des matières

Résumé	3
1 Introduction	7
2 Cadre linguistique	8
2.1 Les actes de langage	8
2.2 Cas d'indirections	12
2.3 Critères d'indirections	13
3 Cadre logique	15
3.1 Langage	15
3.2 Effets ramifiés et critères d'indirection	16
3.3 Lois sur les effets ramifiés	17
4 Conclusion	18
Références	22

1 Introduction

Les dialogues coopératifs homme-machine orientés tâche, qu'ils soient en langage naturel ou artificiel, sont un des défis les plus importants des sciences informatiques. Les participants à de tels dialogues ont un but commun principal, qui est d'accomplir la tâche en question. Chaque participant a quelques informations permettant l'accomplissement de quelques sous-buts, mais aucun d'eux ne peut accomplir le but principal tout seul.

Des travaux ont été entrepris sur la notion d'agent, et parmi ceux-ci, certains se sont basés sur une *théorie de l'Intentionnalité* (Searle, 1983; Searle, 1985; Bratman, 1987). Dans une telle théorie, un *agent rationnel* est représenté par son *état mental*, qui est l'ensemble des informations rendant compte de ses différentes *attitudes mentales* (*i.e.* croyance, but, intention, ...) ¹.

Une telle représentation s'inscrit au sein de théories formelles à deux volets : *l'équilibre rationnel* et *l'interaction rationnelle* (Cohen et Levesque, 1990a; Cohen et Levesque, 1990b; Sadek, 1992), de façon à structurer les *états mentaux* et rendre les agents autonomes. Le premier volet de ces théories décrit les propriétés de chaque attitude mentale, ainsi que les différentes interactions qu'il peut y avoir entre chacune de ces attitudes ; la seconde caractérise, en termes d'attitudes mentales et d'actions, les interactions entre les agents et leur environnement.

Dans des travaux antérieurs (Longin, 1999; Herzig et Longin, 2000a, 2000b, 2000c), nous avons développé notre propre théorie formelle, et avons analysé l'évolution des croyances d'un agent au cours d'un dialogue. Les énoncés sont associés à des actes de langage, et nous nous sommes restreints jusqu'à présent aux effets directs (*i.e.* non ramifiés) des actes de langage.

Nous nous proposons dans ce qui suit d'étendre des travaux antérieurs sur la notion d'*acte de langage indirect* (Herzig et al., 2000; Faure, 2000). Notre but est premièrement d'adapter cette notion aux théories formelles de l'Intentionnalité (que nous appelons parfois à ce titre « théories intentionnelles »), et deuxièmement de la prendre en compte dans le processus d'évolution des croyances de nos agents rationnels ².

Cette prise en compte signifie, d'une part, d'être capable d'inférer les effets ramifiés de l'acte de langage indirect venant d'être accompli et d'autre part d'ajouter aux croyances de l'agent (ou *des* agents) concerné(s), le fait que ces effets ont été

1. Cette théorie est à la base de ce qui est communément appelé les *architectures BDI* (pour *Belief-Desire-Intention*).

2. Nous allons rapprocher ici la notion d'*actes de langage indirects* de celle d'*effets indirects des actions*, ces derniers ayant été étudiés dans le domaine du raisonnement sur les actions. L'inférence de ces effets indirects correspond au problème de la ramification. Pour cette raison, et aussi pour ne pas introduire de confusion par rapport à la notion d'*effet indirect* de SADEK (Sadek, 1991, 2000), nous emploierons le terme d'*effets ramifiés* (plutôt que celui d'*effets indirects*).

produits. Plus précisément, inférer les effets ramifiés de l'acte accompli c'est :

1. déterminer l'ensemble intrinsèque (potentiellement vide) de ces effets ;
2. déterminer si ceux-ci, relativement au contexte d'énonciation, sont déductibles de l'intention du locuteur sous-jacente à cet acte.

Dans un premier temps nous présentons toutes les notions philosophico-linguistiques nécessaires à la compréhension de notre démarche (Sect. 2), puis nous formalisons celles-ci en accord nos travaux antérieurs (Sect. 3).

2 Cadre linguistique

2.1 Les actes de langage

Origines philosophiques. Dans la lignée d'AUSTIN et de SEARLE, nous pensons que toute production d'un énoncé réalise l'accomplissement d'un (ou plusieurs) acte(s) de langage. Un acte de langage est traditionnellement composé d'une force illocutoire s'appliquant sur un contenu propositionnel. Par exemple, l'acte réalisé lors de l'énonciation de « Jean travaille-t-il beaucoup? » a pour contenu propositionnel *Jean travaille beaucoup* et pour force illocutoire une question fermée (*i.e.* dont la réponse attendue est *oui* ou *non*).

SEARLE (Searle, 1969) différencie le *sens de l'énoncé* (qui correspond à ce que signifie littéralement cet énoncé) du *sens du locuteur* (qui correspond au sens que le locuteur souhaite donner à son énoncé).

Par exemple, l'énoncé « Passe-moi le sel » constitue une façon directe de demander le sel, car le *sens de l'énoncé* correspond au *sens du locuteur*. En revanche, dans « Peux-tu me passer le sel? », le sens de l'énoncé et celui du locuteur ne coïncident pas forcément : alors que le premier constitue littéralement une question fermée, le second peut correspondre à une requête (identique à l'énoncé direct ci-dessus) ; on dit alors que *la requête a été accomplie de façon indirecte*, ou encore que *l'énoncé accompli constitue un acte indirect*.

D'après Searle, il n'existe pas d'acte indirect sans acte direct (ou littéral). Donc, même en cas d'indirection, l'acte direct est toujours accompli (potentiellement de façon défectueuse, ou avec insuccès). Ceci s'explique simplement par définition même de l'acte indirect : « acte indirect » signifie « acte accompli (indirectement) par le biais d'un autre acte de langage ». Si ce dernier n'est pas accompli, le premier ne peut donc pas l'être. En ce sens, par exemple, SEARLE réfute l'ironie en tant qu'acte indirect (bien que cela soit un acte non littéral), puisque le sens de l'acte d'ironie remplace le sens de l'acte littéral (Searle, 1979).

L'acte accompli au sens du locuteur est appelé *acte principal*. Il désigne l'acte que le locuteur souhaitait accomplir, que ce soit de façon directe ou indirecte.

Par exemple, dans « Peux-tu me passer le sel? », l'acte principal correspond *habituellement* à l'acte indirect (*i.e.* l'acte selon lequel le locuteur demande à son

interlocuteur que celui-ci lui passe le sel). Nous soulignons qu'on peut (toujours) imaginer un contexte où l'acte principal serait l'acte littéral. Dans cet exemple, ce serait le cas si le locuteur accomplit cet énoncé dans le but d'interroger son interlocuteur sur sa capacité physique à accomplir une certaine action : la réponse attendue serait alors *oui* ou *non*.

SEARLE et VANDERVEKEN ont défini les conditions de succès et de satisfaction des actes de langage (Searle et Vanderveken, 1985; Vanderveken, 1991a, 1991b). En reprenant les notations de ces deux derniers ouvrages, nous présentons celles qui nous seront utiles pour la suite (entre parenthèses nous donnons la valeur de chacune de ces conditions relativement à l'acte réalisé par l'énoncé « Passe-moi le sel ») :

- la *condition sur le contenu propositionnel* décrit des restrictions sur le contenu propositionnel de l'acte (l'auditeur réalise l'action *passer le sel* dans un futur proche) ;
- la *condition préparatoire* représente les propositions que le locuteur doit supposer vraies pour faire son énoncé (le locuteur croit que l'auditeur peut lui passer le sel) ;
- la *condition de sincérité* représente les états mentaux exprimés par le locuteur par rapport au contenu propositionnel de l'acte (le locuteur veut que l'auditeur lui passe le sel).

Actes de langage et théories intentionnelles. Se basant en cela sur les travaux de BRATMAN (Bratman, 1987), COHEN et LEVESQUE (Cohen et Levesque, 1990a, 1990b) furent les premiers à jeter les bases formelles de théories philosophiques de l'Intentionnalité, et à généraliser ainsi la théorie des actes de langage en une théorie de la communication, où les propriétés de ces derniers seraient dérivables de principes (plus généraux) de rationalité.

Globalement, deux points de vue sont en concurrence quant à la notion d'acte indirect. Le premier est celui propre à la théorie des actes de langage et décrit dans le paragraphe précédent : l'acte littéral correspond au *sens de l'énoncé*. Si le *sens du locuteur* ne coïncide pas avec ce dernier, l'énoncé au sens du locuteur réalise un second acte, distinct de l'acte littéral, et correspondant à l'acte indirect³. Une façon de produire un acte indirect étant d'*exploiter* (au sens de (Vanderveken, 1997)) la maxime de qualité, l'acte littéral sera dans ce cas un insuccès et pourra être à ce titre incompatible avec l'acte indirect produit. L'*incompatibilité* en question se situe au niveau des ensembles de préconditions associés respectivement à l'acte littéral et à l'acte indirect. Par exemple, dans « Peux-tu me passer le sel ? », la condition préparatoire de l'acte littéral est que le locuteur ne sait pas si l'auditeur peut lui passer le sel, et la condition préparatoire de l'acte indirect est que le locuteur pense

3. En tout rigueur, il correspond à un acte *non littéral* (*vs* indirect), l'ironie étant une forme de communication non littérale, mais ne réalisant pas d'acte indirect (cf. plus haut).

que l'auditeur peut lui passer le sel. D'une certaine façon, nous pouvons donc dire que dans certains cas, le raisonnement conduisant l'auditeur à inférer l'acte indirect de l'acte direct est non monotone. Néanmoins, on reproche souvent à ce point de vue de complexifier inutilement les choses par l'introduction de la notion d'acte indirect dont on peut se passer.

Le second point de vue est celui de la conventionnalité : la distinction *acte direct* / *acte indirect* est considérée comme artificielle ; simplement, l'usage a fait que, petit à petit, le sens d'un énoncé n'est pas celui littéralement exprimé. Ainsi, « Peux-tu me passer le sel ? » ne serait pas un acte indirect mais une façon (conventionnelle) de demander à son interlocuteur que celui-ci lui passe le sel. Nous pensons néanmoins que ce point de vue est erroné dès lors que tout énoncé peut TOUJOURS être employé de façon purement littérale (ce qu'il est impossible de considérer si l'on adopte cette vision des choses).

Nous adoptons pour notre part un troisième point de vue qui, d'une certaine façon, est à mi-chemin entre les précédents. Globalement, il n'y a plus de distinction entre acte littéral et acte indirect (alors que c'était le cas dans la théorie des actes de langage), mais simplement un acte (celui qui vient d'être accompli). En revanche, plusieurs ensembles d'effets peuvent être associés à un acte donné, rendant compte de l'accomplissement de cet acte dans divers contextes d'énonciation. Typiquement (sans que ce soit une nécessité), deux ensembles d'effets peuvent être dégagés : celui rendant compte du cas où le sens du locuteur et celui de l'énoncé coïncident, et celui rendant compte du cas où ils ne coïncident pas.

Selon ce dernier point de vue, nous formalisons un acte comme ayant des *effets intrinsèques* (relatifs au sens de l'énoncé), et des *effets ramifiés* (relatifs au sens du locuteur, quand celui-ci diffère du sens de l'énoncé). Les effets ramifiés d'un acte sont inférables à partir de l'accomplissement de cet acte et d'une certaine condition relative au contexte d'accomplissement (*i.e.* à l'état mental de l'auditeur). L'inférence des effets ramifiés peut conduire un agent à remettre en cause les effets intrinsèques (cf. l'exemple donné ci-dessus dans le cadre du point de vue de la théorie des actes de langage). Ainsi, la notion d'acte indirect disparaît, mais pas le processus inférentiel conduisant aux effets ramifiés.

Par exemple, si l'agent *s* dit à destination de *h* : « Peux-tu me passer le sel ? », un effet intrinsèque est que *s* rend manifeste son intention de savoir si *h* est capable de passer le sel ou non. Dans certaines circonstances, un effet ramifié est que *s* rend manifeste son intention que *h* lui passe le sel.

Pour arriver à nos fins, nous devons donc disposer d'une part d'un modèle formel d'actes illocutoires, et d'autre part de certaines lois de rationalité permettant de déduire les effets ramifiés d'un acte de son accomplissement et du contexte de cet accomplissement.

Dans notre approche, préconditions et effets d'un acte sont proches (bien que différents) dans leur esprit, de ceux définis dans (Sadek, 1991, 2000) (en particulier,

certaines préconditions et certains effets sont simplifiés) :

- la *précondition de sincérité*⁴ décrit l’attitude mentale requise pour accomplir l’acte ;
- la *précondition de pertinence au contexte* décrit les attitudes mentales requises pour que l’accomplissement de l’acte soit pertinent ;
- l’*effet indirect* d’un acte est la persistance (à l’issus de l’accomplissement de celui-ci) de sa précondition de sincérité⁵ ;
- l’*effet intentionnel* correspond au point de vue gricéen de la communication (Grice, 1989), à savoir l’intention que le locuteur tente de rendre manifeste en accomplissant l’acte ;
- l’*effet rationnel* est l’effet attendu de l’acte par son auteur.

Par exemple, supposons que l’agent *u* informe l’agent *s* que sa destination est Paris. Alors :

- *u* croit que sa destination est effectivement Paris (précondition de sincérité) ;
- *u* ne croit pas que *s* sait si la destination de *u* est Paris (précondition de pertinence) ;
- juste après l’accomplissement de l’acte, la précondition de sincérité est toujours vraie (effet indirect) ;
- *u* a l’intention que *s* croit que *u* a l’intention de produire sur *s* un certain effet (effet intentionnel) ;
- *s* croit que la destination est Paris (effet rationnel⁶).

Les préconditions et les effets sont associés à un acte (ou une actions) *via* les *lois gouvernant les actes* de notre théorie non logique (Longin, 1999; Herzig et Longin, 2000a), exception faite toutefois de l’effet rationnel qui n’est pas consommé à proprement parlé.

Nous utilisons également la notion de *présupposé* introduite dans (Longin, 1999) et qui représente la présupposition par l’auditeur de la vérification des préconditions d’un acte de langage par l’auteur de cet acte. Par exemple, le présupposé relatif à la précondition de pertinence au contexte est que *s* croit que juste avant l’acte, *u* ne croyait pas que *s* savait si la destination de *u* était Paris.

4. De *capacité* dans (Sadek, 2000).

5. Dans (Sadek, 2000), l’auteur recommande de faire également préserver la condition de pertinence au contexte. Nous pensons que cela constitue un choix *a priori* de l’agent qui consiste à penser que son acte n’a pas atteint son but escompté (qu’il n’a pas été satisfait, au sens de la théorie des actes de langage). Nous avons opté pour une solution où l’agent passe par un état de doute intermédiaire où il ne sait plus rien, et où il attend un feedback de son entourage avant de se forger une opinion (Herzig et al., 2000).

6. Cet effet n’est pas à proprement parler « consommé » (*i.e.* ajouté tel quel aux croyances), mais on le considère comme tel s’il est dérivable de l’état mental de l’agent après l’accomplissement de l’acte.

Pour formaliser convenablement les effets ramifiés, il nous suffira alors de disposer de lois indiquant que certains effets supplémentaires ont lieu sous réserve qu'un certain critère soit vérifié. Toute la fin de cette section est consacrée à l'élaboration de telles lois sur des bases linguistiques.

2.2 Cas d'indirections

Fidèle à l'esprit de SEARLE et VANDERVEKEN sur la théorie des actes illocutoires, VIRBEL dégage un certain nombre de façons d'accomplir un acte de langage de type *directif* (Virbel, 1999). Reprenant en grande partie son étude, nous partons du principe selon lequel tout acte (de type directif) peut être accompli soit⁷

1. directement (cas littéral) par :
 - (a) Un marqueur lexical (ex. : « Je te (ordonne, commande, ...) de me rapporter le livre que je t'ai prêté. »);
 - (b) Un marqueur syntaxique (ex. : l'emploi de l'impératif dans « Rapporte moi le livre que je t'ai prêté! »);
2. indirectement (cas indirect) par :
 - (a) assertion sur le contenu propositionnel de (1a) ou (1b) (ex. : « Tu me rapportes le livre que je t'ai prêté. »);
 - (b) question sur le contenu propositionnel de (1a) ou (1b) (ex. : « Me rapportes-tu le livre que je t'ai prêté? »);
 - (c) assertion sur la condition préparatoire de (1a) ou (1b) (ex. : « Tu peux me rapporter le livre que je t'ai prêté. »);
 - (d) question sur la condition préparatoire de (1a) ou (1b) (ex. : « Peux-tu me rapporter le livre que je t'ai prêté? »);
 - (e) assertion sur la condition de sincérité de (1a) ou (1b) (ex. : « Je veux que tu me rapportes le livre que je t'ai prêté »);
 - (f)* question sur la condition de sincérité de (1a) ou (1b) (ex. : « Veuille-je que tu me rapportes le livre que je t'ai prêté? »);
 - (g) assertion sur la condition sur le contenu propositionnel de (1a) ou (1b) (ex. : « Tu (me rapporteras, vas me rapporter) le livre que je t'ai prêté. »);
 - (h) question sur la condition sur le contenu propositionnel de (1a) ou (1b) (ex. : « (Vas-tu me rapporter, me rapporteras-tu) le livre que je t'ai prêté? »);
 - (i)* assertion sur la présupposition de la condition préparatoire de (1a) ou (1b) (ex. : « (Tu veux, tu voudrais, tu accepterais de, ...) me rapporter le livre que je t'ai prêté. »);

7. Les cas marqués d'une * correspondent à des cas contre-intuitifs que nous ne retenons pas par la suite.

- (j) question sur la présupposition de la condition préparatoire de (1a) ou (1b) (ex. : « (Veux-tu, voudrais-tu, accepterais-tu de, ...) me rapporter le livre que je t'ai prêté? »).

Dans la classification initiale de VIRBEL, certains cas étaient regroupés (comme (2d) et (2e)). Les cas (2i) et (2j) étaient caractérisés par « la présupposition de l'acceptation de l'auditeur » (Virbel, 1999). Nous préférons y voir une assertion (resp. une question) sur l'intention de l'auditeur de réaliser l'action incluant une présupposition de la condition préparatoire de (1a) ou (1b). Les cas *étoilés* constituent des cas contre-intuitifs, ou non recevables dans le cadre de la théorie de SEARLE. Les cas (2b) et (2i) ont été ajoutés à la classification originale de VIRBEL (cf. à ce sujet (Faure, 2000)).

VIRBEL met en valeur plusieurs autres façons d'accomplir un acte de façon indirecte, que nous ne reprenons pas ici car elles ne peuvent pas être traduites en termes de force illocutoire s'appliquant sur une condition de succès ou le contenu propositionnel de l'acte littéral (ces cas sont donc beaucoup plus difficilement identifiables formellement).

Définition 1 *Tout acte dont la forme syntaxique constitue une façon indirecte d'accomplir un acte de type directif (au sens de la classification ci-dessus) est appelé : forme d'indirection.*

2.3 Critères d'indirections

L'idée est que la classification précédente va nous fournir le support pour l'écriture de lois décrivant les effets ramifiés d'un acte. Reste à définir un critère dont le rôle sera de restreindre l'application de ces lois. Nous partons de l'idée que tout énoncé est pertinent (au sens de SPERBER et WILSON (Sperber et Wilson, 1989)). À ce titre, si la persistance d'un présupposé sur la condition de pertinence au contexte est inconsistante avec la préservation (d'une partie) de la mémoire de l'agent, alors la cause en est peut-être une indirection.

Définition 2 *Toute d'attitude mentale qui, une fois préservée, est susceptible d'être inconsistante avec la préservation d'un présupposé relatif à la précondition de pertinence au contexte est appelé : critère d'indirection.*

Partant du fait qu'un acte vient d'être accompli, plusieurs cas sont envisageables, selon que cet acte constitue ou non une forme d'indirection, et que le critère d'indirection est ou n'est pas vérifié. Ces différents cas peuvent être résumés comme suit :

Si si l'acte accompli est une forme d'indirection

Alors Si le critère d'indirection est vrai

Alors l'acte accompli possède des effets ramifiés

Sinon l'acte accompli ne possède pas, *a priori*, d'effet ramifié, mais certains principes de coopération vont amener l'agent à poser une question sur la production éventuelle d'effets ramifiés

Sinon (aucun effet ramifié n'est alors détectable dans le cadre d'une théorie à la VIRBEL):

Si le critère d'indirection n'est pas vérifié

Alors pas d'autre effet ramifié

Sinon certains autres effets ramifiés peuvent être à envisager

Exemple. Soit α l'acte accompli par l'agent i à destination de l'agent j et réalisé par l'énoncé « Peux-tu me rapporter le livre⁸ de Paul Gochet? ». La précondition de pertinence au contexte de α est que i ne sait pas si j peut ou non rapporter le livre de Paul Gochet. Un critère d'indirection est par exemple que j pense que i est d'avis que j peut lui rapporter le livre. Soit α' l'acte accompli par l'agent i à destination de l'agent j et réalisé par l'énoncé « Rapporte-moi le livre de Paul Gochet ». La précondition de pertinence au contexte de α' est que i ne pense pas que j avait l'intention de lui rapporter le livre. Dans ces conditions :

(Cas 1) α vient d'être accompli (forme d'indirection). Si nous supposons qu'il est manifeste pour j que i pense que j peut lui rapporter le livre (critère d'indirection vérifié), alors j va être amené à penser qu'une autre intention également sous-jacente à l'énoncé de i est que i souhaite que j lui rapporte son livre. La réponse de j attendue sera typiquement « Oui, je te le rapporte », ou « Tiens, je te l'ai apporté »...

(Cas 2) α vient d'être accompli (forme d'indirection). Supposons qu'il n'est pas le cas que j croit que i pense que j peut lui rapporter le livre (critère non vérifié). C'est typiquement le cas si j pense que i a de bonnes raisons de douter de cette capacité (physique). À ce moment-là, la réponse attendue est *oui* ou *non* (quitte à ce que i dise ensuite quelque chose comme « Et bien rapporte-le »). Mais l'acte accompli par i constituant une forme d'indirection, des principes coopératifs devraient pousser j à interroger i sur une éventuelle intention sous-jacente à sa question. Dans le cas présent, il pourrait par exemple faire une réponse plus complète: « Oui. Veux-tu que je te le rapportes? ».

(Cas 3) α' vient d'être accompli (qui n'est pas une forme d'indirection). Si, de plus, le critère d'indirection n'est pas vérifié, alors l'acte accompli ne possède aucun effet ramifié (dans le cadre d'une théorie à la VIRBEL, ou autre). j se contente d'obéir ou de désobéir à l'ordre reçu.

8. (Gochet et al., 2000).

(Cas 4) α' vient d'être accompli (qui n'est pas une forme d'indirection). Si j était d'avis que i pensait que j avait l'intention de lui rapporter son livre (précondition de pertinence au contexte de α' non vérifiée du point de vue de j), alors il y a à cela plusieurs explications possibles : soit j avait jusqu'alors une croyance erronée ou le monde a évolué, et j doit changer ses croyances. À l'issue de ce changement de croyance, il se retrouve donc dans le cas 3 ci-dessus. Soit cela constitue une forme d'insistance, et certains effets ramifiés doivent pouvoir être produits (du genre : j pense qu'il est urgent pour i que j rende le livre à i ...).

Dans ce qui suit, nous formalisons cette démarche dans le cadre de notre logique épistémique dynamique (Longin, 1999; Herzig et Longin, 2000a).

3 Cadre logique

3.1 Langage

Comme beaucoup, nous travaillons dans un cadre multimodal, avec des opérateurs primitifs de croyance, de croyance mutuelle, d'intention, d'action et de temps. Notre langage est celui de la logique multimodale du premier ordre sans égalité ni symbole de fonction.

Soit AGT l'ensemble des agents. Pour $i \in AGT$, $Bel_i A$ est lu « l'agent i croit que A », $Intend_i A$ est lu « l'agent i a l'intention que A », $Bel_{i,j} A$ est lu « les agents i et j croient mutuellement que A ». Par exemple, $Bel_i Dest(\text{Paris})$ exprime que l'agent i croit que la destination est Paris. $Bel_{i,j} A$ est une abbréviation pour $Bel_i A \vee Bel_j A$ qui est lu « l'agent i sait si A est vrai ou non ».

$\mathcal{FI} = \{\text{Assert, Inform, Request, QueryYN, QueryWh}\}$ est un ensemble de forces illocutoires, et $\mathcal{FP} = \{\text{Give, Take, Insert, Dial}\}$ est un ensemble de noms d'actions physiques (que nous appelons *forces physiques*, par analogie avec les forces illocutoires).

Les actions physiques sont des 4-uplets de la forme $\langle \text{FORCE}_{i,j} A \rangle$ où $\text{FORCE} \in \mathcal{FP}$ est le nom de l'action (la *force*) physique, $i \in AGT$, j est le destinataire (au sens large) de l'action (en particulier, tout élément de AGT), et A un symbole de constante. Par exemple, l'action exécutée par l'agent u d'insérer sa carte bleue dans le lecteur de cartes est $\langle \text{Insert}_{u, \text{lecteur}} \text{carte_bleue} \rangle$. On note \mathcal{PACT} l'ensemble des actions physiques.

Les actes illocutoires sont des 4-uplets de la forme $\langle \text{FORCE}_{i,j} A \rangle$ où $\text{FORCE} \in \mathcal{FI}$ est la force illocutoire de l'acte, $i, j \in AGT$, et A est une formule du langage représentant le contenu propositionnel de l'énoncé. Par exemple, $\langle \text{Inform}_{i,j} Dest(\text{Paris}) \rangle$ est réalisé par un énoncé informatif de l'agent i à destination de l'agent j selon lequel la destination est Paris; $\langle \text{QueryYN}_{i,j} Bel_j Prix(150 \text{ €}) \rangle$ signifie « l'agent i demande à l'agent j s'il croit que le prix est 150 € ». On note \mathcal{ACT} l'ensemble des schémas

d'actes illocutoires (isomorphe au produit cartésien $\mathcal{FI} \times \mathcal{AGT} \times \mathcal{AGT}$) et où le contenu propositionnel est une variable schématique de formules du langage.

Soit $\mathcal{ACT} = \mathcal{ACT} \cup \mathcal{PACT}$ l'ensemble des schémas d'actes illocutoires et des actions physiques. À chaque $\alpha \in \mathcal{ACT}$ nous associons un opérateur modal $Done_\alpha$. $Done_\alpha A$ est lu « α vient juste d'être accompli, avant quoi A était vrai »⁹. $Done_\alpha \top$ est lu « α vient juste d'être accompli ».

$Always A$ est lu « la formule A est toujours vrai dans tous les états mentaux » (sémantiquement, c'est la fermeture réflexive et transitive des relations d'accessibilité des actions).

$Bel_i Dest(\text{Paris})$ est un exemple de formule. De même, $Done_{\langle \text{Inform}_{i,j} p \rangle} Bel_i p$ exprime la sincérité de i : i vient d'informer j que p , juste avant quoi i croyait p .

Les formules atomiques sont notées p, q, \dots ou $P(t_1, \dots, t_n)$. \mathcal{ATM} est l'ensemble de toutes les formules atomiques. Les formules sont notées A, B, \dots

3.2 Effets ramifiés et critères d'indirection

Nous disposons dans notre théorie non logique de certaines lois (les *lois gouvernant les actes*) décrivant les préconditions d'un actes de langage ainsi que ses effets (indirect et intentionnel). C'est sous forme de lois similaires (mais dont l'application sera restreinte par le critère d'indirection) que nous allons décrire les effets ramifiés.

On note $\beta \in \mathcal{ACT}$ l'acte de langage ou l'action dont l'auteur est j , le destinataire est i , et dont il est question dans l'acte α venant d'être accompli par i à destination de j .

Par exemple, que α soit l'acte réalisé par l'énoncé « Passe-moi le sel » ou par l'énoncé « Peux-tu me passer le sel? », l'action β est $\langle \text{Give}_{j,i} \text{sel} \rangle$: dans le premier cas, i demande à j d'accomplir β , et dans le second cas, i questionne j sur la capacité de ce dernier à accomplir β (si on se base sur le sens de l'énoncé). Notre propos consiste à montrer ici que parfois, même dans le second cas, i demande (indirectement) à j d'accomplir β .

Puisque nous souhaitons formaliser les effets ramifiés d'actes dont le sens du locuteur correspond à un directif, ces effets ramifiés sont en relation avec les présupposés et les effets engendrés par l'accomplissement d'un acte de ce type.

Exemple 1 Soit $\alpha = \langle \text{QueryYN}_{i,j} \text{Feasible}_\beta \top \rangle$ l'acte réalisé par l'énoncé « Peux-tu me passer le sel? » (et où $\beta = \langle \text{Give}_{j,i} \text{sel} \rangle$).

α étant une forme d'indirection, on suppose que le critère d'indirection est vérifié et que l'acte réalisé au sens du locuteur est que i demande à j que j passe le sel à i .

Si j considère qu'une telle demande vient d'être accomplie par i , alors il présuppose que i croit que j est capable de lui passer le sel, et que i ne croit pas que j ait déjà l'intention de le lui passer. Soit, formellement : $Bel_j Done_\alpha (Bel_i \text{Feasible}_\beta \top \wedge$

9. $Done_\alpha A$ est similaire à $\langle \alpha^{-1} \rangle A$ de la logique dynamique (Harel, 1984).

$\neg Bel_i Intend_j Done_\beta \top$). C'est l'effet ramifié en relation avec les présupposés liés à l'accomplissement d'une requête.

À l'issue de l'accomplissement de cette demande, j doit également considérer que i a l'intention qu'il croie que i a l'intention qu'il lui passe le sel, et que i continue à croire qu'il peut le lui passer. Soit, formellement : $Bel_j(Intend_i Bel_j Intend_i Done_\beta \top \wedge Bel_i Feasible_\beta \top)$. C'est l'effet ramifié lié aux effets de l'accomplissement d'une requête.

Définition 3 *Soit*

$$A^\alpha \stackrel{\text{déf}}{=} Intend_i Bel_j Intend_i Done_\beta \top \wedge Bel_i Feasible_\beta \top \wedge Done_\alpha (Bel_i Feasible_\beta \top \wedge \neg Bel_i Intend_j Done_\beta \top)$$

la conjonction des effets ramifiés de type directif liés aux effets et présupposés d'une requête accomplie par le biais de l'acte α .

Les effets ramifiés étant ainsi formalisés (dans l'exemple précédent, β peut en fait être n'importe quelle action), reste à formaliser les critères d'indirection.

En relation avec la définition 2 p. 13, le critère C recherché est n'importe quelle formule inconsistante avec le précondition de pertinence au contexte de l'acte α venant d'être accompli. D'après les formes d'indirection recensées précédemment (cf. Sect. 2.2 p. 12), l'acte accompli est :

- soit $\alpha = \langle \text{Inform}_{i,j} A \rangle$ (dont la précondition de pertinence au contexte est $\neg Bel_i BelIf_j A$);
- soit $\alpha = \langle \text{QueryYN}_{i,j} A \rangle$ (dont une des préconditions de pertinence au contexte est $\neg BelIf_i A$).

En conséquence, le critère C sera respectivement tel que :

- soit $C = Bel_i Bel_j A$;
- soit $C = Bel_i A$.

3.3 Lois sur les effets ramifiés

Les axiomes permettant de déduire les effets ramifiés seront donc de la forme :

$$Always Bel_k (Done_\alpha Bel_k C \rightarrow A^\alpha)$$

qui signifie : « il est toujours le cas que n'importe quel agent k observateur de l'acte¹⁰ croit que : si l'acte α vient juste d'être accompli, avant quoi il croyait que le critère C était vérifié, alors les effets ramifiés lié à l'accomplissement de α se sont produits ».

Axiome 1 $Always Bel_k (Done_\alpha Bel_k Bel_i Bel_j Done_\beta \top \rightarrow A^\alpha)$ où $\alpha = \langle \text{Assert}_{i,j} Done_\beta \top \rangle$

¹⁰. Dans l'exemple précédent, nous avons assimilé k à un observateur particulier qu'est les destinataire de cet acte (j dans l'exemple).

Axiome 2 $Always Bel_k(Done_\alpha Bel_k Bel_i Done_\beta \top \rightarrow A^\alpha)$ où
 $\alpha = \langle QueryYN_{i,j} Done_\beta \top \rangle$

Axiome 3 $Always Bel_k(Done_\alpha Bel_k Bel_i Bel_j Feasible_\beta \top \rightarrow A^\alpha)$ où
 $\alpha = \langle Assert_{i,j} Feasible_\beta \top \rangle$

Axiome 4 $Always Bel_k(Done_\alpha Bel_k Bel_i Feasible_\beta \top \rightarrow A^\alpha)$ où
 $\alpha = \langle QueryYN_{i,j} Feasible_\beta \top \rangle$

Axiome 5 $Always Bel_k(Done_\alpha Bel_k Bel_i Bel_j Intend_i Done_\beta \top \rightarrow A^\alpha)$ où
 $\alpha = \langle Assert_{i,j} Intend_i Done_\beta \top \rangle$

Axiome 6 $Always Bel_k(Done_\alpha Bel_k Bel_i Bel_j \exists \varepsilon : Feasible_\varepsilon Done_\beta \top \rightarrow A^\alpha)$ où
 $\alpha = \langle Assert_{i,j} \exists \varepsilon : Feasible_\varepsilon Done_\beta \top \rangle$

Axiome 7 $Always Bel_k(Done_\alpha Bel_k Bel_i \exists \varepsilon : Feasible_\varepsilon Done_\beta \top \rightarrow A^\alpha)$ où
 $\alpha = \langle QueryYN_{i,j} \exists \varepsilon : Feasible_\varepsilon Done_\beta \top \rangle$

Axiome 8 $Always Bel_k(Done_\alpha Bel_k Bel_i Intend_j Done_\beta \top \rightarrow A^\alpha)$ où
 $\alpha = \langle QueryYN_{i,j} Intend_j Done_\beta \top \rangle$

4 Conclusion

Nous avons montré comment inférer certains effets ramifiés à partir de l'acte accompli, de lois gouvernant les actes et du contexte (*i.e.* de l'état mental de l'agent considéré au moment de l'accomplissement de l'acte). Cette démarche s'appuie sur la reconnaissance (syntaxique) dans l'acte accompli d'une *forme d'indirection*, puis sur l'application d'une certaine loi décrivant les effets ramifiés si un certain *critère d'indirection* est vérifié dans l'état mental de l'agent considéré.

Nous pensons qu'un des avantages de cette approche est son originalité quant à la phase de reconnaissance d'une forme d'indirection : elle exploite des travaux récents en linguistique, offrant ainsi un outil générique permettant de travailler sur des classes d'actes (les directifs, les assertifs...) plutôt que sur des actes particuliers. Cette approche a également l'avantage de permettre l'inférence d'effets ramifiés pour un agent particulier (puisque le critère est relatif à un état mental donné) : ainsi, il est tout à fait possible que l'acte accompli ait des effets ramifiés du point de vue d'un observateur de l'acte accompli, mais pas du point de vue d'un autre observateur de cet acte. Un autre avantage est la souplesse d'utilisation de ces lois sur les effets ramifiés : la contextualisation de leur application nous permet d'affecter, au besoin, plusieurs ensembles distincts d'effets ramifiés à un même acte, correspondant à des contextes mentaux différents. Enfin, cette approche offre une certaine solution au problème de la ramification : rien n'interdit à ce qu'un critère d'application d'une

des lois contienne elle-même des effets ramifiés, ce qui affine l'ensemble global des effets ramifiés associés à un acte qui ne sera pas le même dans tous les contextes.

La difficulté principale réside plutôt dans l'élaboration du critère d'indirection. Nous sommes conscients que d'une façon générale, notre processus de génération des effets ramifié n'est ni sain, ni complet : il existe certainement des contre-exemples pour lesquels notre critère s'applique alors qu'aucun effet ramifié n'aurait du être inféré, ou au contraire pour lesquels des effets ramifiés auraient du être appliqués alors que notre critère n'était pas vérifié.

Néanmoins, il s'agit pour nous ici simplement d'ébaucher une méthode de calcul et d'application des effets ramifiés dans un cadre particulier d'applications telles que les dialogues orientés tâche. L'idée sous-jacente à cette remarque est que si les formes d'indirection sont relativement identiques quels que soient les types de dialogues analysés (orientés tâche, argumentatif, ...), il n'en n'est pas de même pour les critères d'indirection, tout simplement parce que ceux-ci sont fortement contextuels. Ce n'est donc que dans un contexte particulier qu'un agent inférera de l'acte accompli certains effets ramifiés, alors qu'il ne les aurait pas inférés dans un contexte différent. De plus, nous pensons pouvoir, à terme, le raffiner en faisant par exemple intervenir d'autres préconditions, et capturer ainsi un plus large éventail de cas.

Au niveau de l'évolution des croyances, ce qui précède montre clairement que les effets indirects de l'acte accompli peuvent être inconsistants avec certains effets ramifiés. Il s'agit alors de gérer l'évolution des croyances des agents en prenant en compte cette nouvelle source d'inconsistance.

Références

- BRATMAN, M. E. (1987). *Intention, Plans, and Practical Reason*. Harvard University Press, Cambridge, MA.
- COHEN, P. R. et LEVESQUE, H. J. (1990a). « Intention is Choice with Commitment ». *Artificial Intelligence Journal*, 42(2–3).
- COHEN, P. R. et LEVESQUE, H. J. (1990b). Rational Interaction as the Basis for Communication. Dans COHEN, P. R., MORGAN, J., et POLLACK, M. E., éditeurs, *Intentions in Communication*. MIT Press.
- FAURE, R. (2000). « « Peux-tu me passer le sel? » : formalisation de la détection des actes indirects directifs dans des dialogues entre agents coopératifs autonomes. ». Mémoire de DEA, IRIT, Université Paul Sabatier, Toulouse, France.
- GOCHET, P., GRIBOMONT, P., et THAYSE, A. (2000). *Logique*, volume 3, méthodes pour l'intelligence artificielle. Hermès Science Publications.
- GRICE, H. P. (1989). *Studies in the way of words*. Harvard University Press, USA, 3rd edition.
- HAREL, D. (1984). Dynamic Logic. Dans GABBAY, D. et GUENTHNER, F., éditeurs, *Handbook of Philosophical Logic*, volume II. D. Reidel Publishing Company.
- HERZIG, A. et LONGIN, D. (2000a). « Belief Dynamics in Cooperative Dialogues ». *Journal of Semantics*, 17(2):91–118.
- HERZIG, A. et LONGIN, D. (2000b). « Beliefs, intentions, speech acts and topics ». Rapport technique 00-08-R, IRIT. Available on <http://www.irit.fr/ACTIVITES/LILaC>.
- HERZIG, A. et LONGIN, D. (2000c). « A Topic-Based Framework for Rational Interaction ». Dans *Proc. Conf. Int. sur le Traitement Automatique du Langage Naturel (TALN-2000)*, pages 431–434. Poster Session.
- HERZIG, A., LONGIN, D., et VIRBEL, J. (2000). « Towards an Analysis of Dialogue Acts and Indirect Speech Acts in a BDI Framework ». Dans POESIO, M. et TRAUM, D., éditeurs, *Proc. Fourth Int. Workshop on the Semantics and Pragmatics of Dialogue (Götaglog-2000)*, Göteborg University, Sweden.
- LONGIN, D. (1999). « *Interaction rationnelle et évolution des croyances dans le dialogue: une logique basée sur la notion de topique* ». Thèse de doctorat, IRIT, Université Paul Sabatier, Toulouse, France. <http://www.irit.fr/ACTIVITES/LILaC>.
- SADEK, M. D. (1991). « *Attitudes mentales et interaction rationnelle: vers une théorie formelle de la communication* ». Thèse de doctorat, Université de Rennes I, Rennes, France.
- SADEK, M. D. (1992). « A Study in the Logic of Intention ». Dans NEBEL, B., RICH, C., et SWARTOUT, W., éditeurs, *Proc. Third Int. Conf. on Principles*

- of Knowledge Representation and Reasoning (KR'92)*, pages 462–473. Morgan Kaufmann Publishers.
- SADEK, M. D. (2000). « Dialogue Acts are Rational Plans ». Dans TAYLOR, M., NÉEL, F., et BOUWHUIS, D., éditeurs, *The structure of multimodal dialogue*, pages 167–188, Philadelphia/Amsterdam. John Benjamins publishing company. From ESCA/ETRW, Workshop on The Structure of Multimodal Dialogue (Venaco II), 1991.
- SEARLE, J. R. (1969). *Speech acts: An essay in the philosophy of language*. Cambridge University Press.
- SEARLE, J. R. (1979). *Expression and Meaning*. Cambridge University Press.
- SEARLE, J. R. (1983). *Intentionality: An essay in the philosophy of mind*. Cambridge University Press.
- SEARLE, J. R. (1985). *Du cerveau au savoir*. Hermann.
- SEARLE, J. R. et VANDERVEKEN, D. (1985). *Foundation of illocutionary logic*. Cambridge University Press.
- SPERBER, D. et WILSON, D. (1989). *La pertinence: communication et cognition*. Proposition. Les éditions de Minuit, Paris.
- VANDERVEKEN, D. (1991a). *Meaning and Speech Acts*, volume 1: *Principles of language use*. Cambridge University Press.
- VANDERVEKEN, D. (1991b). *Meaning and Speech Acts*, volume 2: *Formal semantics of success and satisfaction*. Cambridge University Press.
- VANDERVEKEN, D. (1997). « Formal Pragmatics of Non Literal Meaning ». *Linguistische Berichte*.
- VIRBEL, J. (1999). Contributions de la théorie des actes de langage à une taxinomie des consignes. Dans VIRBEL, J., CELLIER, J.-M., et NESPOULOUS, J.-L., éditeurs, *Cognition, Discours procédural, Action*, volume II. PRESCOT.