

Actions perceptives et non-perceptives dans une logique de croyances et intentions

Olivier Gasquet, Andreas Herzig, Dominique Longin

▶ **To cite this version:**

Olivier Gasquet, Andreas Herzig, Dominique Longin. Actions perceptives et non-perceptives dans une logique de croyances et intentions. Premières Journées Francophones des Modèles formels de l'interaction (MFI'01), May 2001, Toulouse, France. hal-03523410

HAL Id: hal-03523410
https://hal.science/hal-03523410

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.

Actions perceptives et non-perceptives dans une logique de croyances et intentions*

Olivier Gasquet, Andreas Herzig, and Dominique Longin

Institut de Recherche en Informatique de Toulouse (IRIT)
Équipe Logic, Interaction, Language and Computation (LILaC)
CNRS – Université Paul Sabatier, 118 Route de Narbonne, F-31062 Toulouse cedex 4
{gasquet,herzig,longin}@irit.fr, http://www.irit.fr/ACTIVITES/LILaC/

Résumé Nous présentons dans cet article une logique contenant des opérateurs (croyance, intention, et action) permettant de représenter les états mentaux d'un ensemble d'agents en interaction, dans le but est définir la logique la plus simple capable de rendre compte de ces interactions. Nous analysons les liens entre les différentes attitudes mentales que nous manipulons et l'action. Nous montrons comment formaliser la perception, la rétroaction (« feedback »), et donnons un exemple.

1 Introduction

Notre but est de formaliser à l'aide de la logique l'interaction entre agents. Chaque agent est décrit par ses *états mentaux*, et tout ce qui est extérieur à ces derniers constitue l'*environnement*. Celui-ci est constitué du monde physique dans lequel les agents évoluent, et des autres agents (identifiés avec leurs états mentaux respectifs). Les croyances d'un agent (considérées comme des connaissances subjectives) modélisent son environnement, et ce modèle est bien sûr potentiellement incomplet, voire erroné.

L'interaction entre un agent et son environnement s'effectue au travers de l'action (de cet agent vers son environnement) et de la perception (de l'environnement vers l'agent). D'un certain point de vue, la perception est de nature double, selon qu'elle soit intentionnelle[1] ou non.

Nous souhaitons pouvoir raisonner sur des échanges du type de l'exemple suivant, où un agent u appelle les renseignements téléphoniques (dont le représentant est noté s) pour connaître le numéro d'un agent k. Il sait par ailleurs que l'agent k' connaît également ce numéro. Nous supposons que k est sur la liste rouge, mais que u n'en sait rien :

s_1 : *Bonjour. Que désirez-vous?*
u_1 : *Quel est le numéro de téléphone de* k *?*
s_2 : *Euh . . . k n'est pas dans l'annuaire.*
u_2 : *Ah, il est donc sur la liste rouge ; quel est le numéro de* k' *?*
s_3 : *Le numéro de* k' *est . . .*

* Nous remercions les rapporteurs dont les commentaires ont été fort utiles.

1. « Intention » est pris ici au sens de [2], *i.e.* en tant que but persistant.

Nous souhaitons déduire les croyances et intentions des agents après chaque action (linguistique, dans l'exemple, mais pas nécessairement. Plus généralement, nous nous intéressons aux actions communicatives dans leur ensemble). Dans ce but, nous étudions ici la logique la plus simple permettant de représenter et de raisonner sur ce type d'interaction, en nous restreignant à une analyse propositionnelle.
Dans ce but, nous présentons tout d'abord une formalisation des états mentaux (Sect. 2), puis les liens unissant l'action et la perception (Sect. 3). Après avoir donné quelques pistes relatives à la sémantique adoptée (Sect. 4), nous illustrerons comment formaliser certains principes liés aux architectures d'agents rationnels (Sect. 5). Finalement, nous formalisons notre exemple (Sect. 6).

2 Structure des états mentaux

Fondé sur les théories philosophiques de [2] et de [24], le formalisme adopté [17, 13] s'inscrit dans la lignée de ceux de [4, 5] et [19, 20] et tend, comme ces derniers, à généraliser 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. Le cadre formel que nous utilisons est du type BDI (« Belief, Desire, Intention ») : une logique doxastique dynamique avec des opérateurs de croyance, croyance commune, intention, et action.

2.1 La croyances

L'analyse des notions de connaissance et croyance en termes d'états possibles a été proposé par Hintikka [14]. Chaque état correspond ici à l'état mental d'un agent à un instant donné pendant lequel aucune action n'est accomplie. « Croyance » est à prendre ici au sens du « belief » de l'École d'Oxford, *i.e.* un modèle du monde subjectif à l'agent dont on considère la croyance.

Nous disposons d'un ensemble d'opérateurs de croyance $\{Bel_i, Bel_j, \ldots\}$ correspondant à l'ensemble d'agents $\mathcal{AGT} = \{i, j, \ldots\}$ la formule $Bel_i A$ est lue « l'agent i croit que A ». $BelIf_i A$ est une abréviation de $Bel_i A \vee Bel_i \neg A$ et signifie « L'agent i sait si A est vrai ou non »[2] Comme c'est fréquemment le cas [8], nous adoptons KD45 comme logique de la croyance, et dans ce cadre, un agent ne peut avoir de croyances contradictoires , est conscient de qu'il croit et de ce qu'il ne croit pas.

2.2 La croyance commune

Nous avons également besoin de la notion de croyance commune, formalisée par le biais d'opérateurs Bel_{Agt}, pour $Agt \subseteq \mathcal{AGT}$. Nous allons considérer une version simplifiée, en négligeant le fait que la croyance commune est reliée aux croyances individuelles par un axiome de type inductif [6]. Nous allons donc simplement considérer que les opérateurs Bel_{Agt} obéissent aux axiomes de KD45, ainsi qu'à l'axiome (Common$_{Bel}$) : $Bel_{Agt} A \rightarrow Bel_i A$ pour $i \in Agt$

2. Le terme « savoir » est utilisé par « commodité linguistique ».

2.3 L'intention

L'intention est une attitude mentale fondamentale, car c'est elle qui relie la croyance à l'action. Elle est à l'origine de toute action volontaire de l'agent. Notre notion est fidèle à celle de [2].

Contrairement à nos travaux antérieurs [12], notre notion d'intention n'est pas nécessairement fermée sous la vérité logique, la conséquence logique, la conjonction, et l'implication matérielle[3]. Ceci est en accord avec les analyses de l'intention que l'on peut trouver dans [2, 4, 19]. Contrairement à ces dernières approches, l'intention est ici une notion primitive, comme c'est d'ailleurs le cas dans [18] et [15]. Nous généralisons ainsi la sémantique de [15], où seule la fermeture sous la conséquence logique avait été supprimée.

La seule règle d'inférence que nous admettons est (RE_{Int_i}) : $\frac{A \leftrightarrow B}{Int_i A \leftrightarrow Int_i}$

Nous avons choisi cette solution pour trois raisons. Premièrement, construire l'intention à partir de notions primitives telles que les buts ou les désirs conduit à des notions variées et très sophistiquées d'intention, avec de subtiles différences entre elles. Nous n'avons retenu ici que les propriétés communément acceptées. Deuxièmement, les définitions de l'intention, dans la littérature, sont complexes, et il est difficile de leur trouver des méthodes automatiques de démonstration de théorèmes complètes. Notre analyse, au contraire, permet d'utiliser des techniques plus ou moins standard de complétude et des méthodes de preuve. Troisièmement, et c'est peut-être là le plus important, nous pensons que notre notion simplifiée d'intention est suffisante dans la majorité des applications actuelles.

2.4 Relations entre attitudes mentales d'un agent

Nous pensons que plus que le rapport entre les intentions et les buts ou les désirs, ce qui importe est le rapport entre les intentions et les croyances. Par exemple, un agent devrait abandonner son intention de réaliser A quand il croit que A est vrai. Ceci est exprimé par le schéma d'axiome (Réalisme) : $Int_i A \rightarrow Bel_i \neg A$, qui avec (D_{Bel_i}) entraîne: (1) $Bel_i A \rightarrow \neg Int_i A$, ainsi que (2) $Int_i A \rightarrow \neg Int_i \neg A$.

Nous adoptons également (3) $Bel_i Int_i A \leftrightarrow Int_i A$ et (4) $Bel_i \neg Int_i A \leftrightarrow \neg Int_i A$

Finalement, les principes ci-dessus énoncés assurent que l'état mental d'un agent i est défini par un ensemble de formules préfixées par Bel_i, Int_i ou leur négation. Mais si cela suffit du point de vue de la représentation des connaissances (au sens large), nous devons maintenant doter nos agents de la capacité à interagir. C'est l'objet de la section suivante.

3 Action et perception

Pour parler des actions nous utilisons une version de la logique dynamique [9]. Puisque nous nous situons dans un cadre multi-agents, nous considérons que

3. Ainsi, nous rejetons les règles $\frac{A}{Int_i A}$ et $\frac{A \rightarrow B}{Int_i A \rightarrow Int_i B}$, et les axiomes $Int_i A \wedge Int_i B \rightarrow Int_i(A \wedge B)$ et $Int_i A \wedge Int_i(A \rightarrow B) \rightarrow Int_i B$ (valides dans toute logique modale normale).

les actions sont accomplies par des agents. Chaque action sera donc préfixée par son auteur. Nous supposons qu'une action atomique est de la forme $i : a$, où a est un schéma d'action et i un agent (l'auteur de a).

Pour construire des actions complexes, nous disposons des opérateurs de composition :

- λ (l'action vide) ;
- ; (l'exécution séquentielle) ;
- $\cup$ (le choix non déterministe).

Nous n'avons donc ni l'opérateur d'itération « * », ni l'opérateur de test « ? » de la logique dynamique. Nous notons α,α', ... les actions complexes.

Pour construire des formules, nous disposons d'opérateurs dynamiques. À chaque action α est associé un opérateur $Done_\alpha$, qui correspond à l'opérateur $\langle\alpha^{-1}\rangle$ de la logique dynamique. Les opérateurs de type $Feasible_\alpha$ et $Before_\alpha$ sont définis respectivement comme le converse et le dual de $Done_\alpha$ (et correspondent aux opérateurs $\langle\alpha\rangle$ et $[\alpha^{-1}]$ de la logique dynamique). Enfin, les opérateurs de type $After_\alpha$ sont définis comme les converses des opérateurs $Feasible_\alpha$.

Un exemple de formule est $Bel_i Feasible_{j:a}\top \wedge After_{j:a}\bot$, exprimant que i croit que j peut accomplir a, alors que j n'en est pas capable. Un autre exemple est $\neg Bel_i p \wedge \neg Bel_i \neg p \wedge After_{i:a} Bel_i p$, exprimant que i ignore si p est vrai ou non, et se met à croire p après avoir exécuté a. Nous notons A,B,C, ... les formules complexes.

Nous notons respectivement ACT et $\mathcal{FORM}$ l'ensemble des actions et formules ainsi définis. Nous disons qu'une formule de $\mathcal{FORM}$ est *objective* si elle ne contient aucune occurrence d'opérateur modal. Nous avons donc les règles d'inférence et les axiomes standards suivants :

$$\frac{A}{After_\alpha A} \qquad (\mathrm{N}_{After_\alpha})$$

$$After_\alpha A \wedge After_\alpha(A \to B) \to After_\alpha B \qquad (\mathrm{K}_{After_\alpha})$$

$$After_\lambda A \;\leftrightarrow\; A \qquad (\mathrm{Déf}_\lambda)$$

$$After_{\alpha;\beta} A \;\leftrightarrow\; After_\alpha After_\beta A \qquad (\mathrm{Déf}_;)$$

$$After_{\alpha\cup\beta} A \;\leftrightarrow\; (After_\alpha A \wedge After_\beta A) \qquad (\mathrm{Déf}_\cup)$$

$$A \to After_\alpha Done_\alpha A \qquad (\mathrm{Conv}_1)$$

$$A \to Before_\alpha Feasible_\alpha A \qquad (\mathrm{Conv}_2)$$

($\mathrm{Déf}_\lambda$) indique que λ est l'action vide ; (Conv_1) et (Conv_2) sont des axiomes de conversion entre un opérateur et le dual de son converse. Ils illustrent notamment que, sémantiquement, nos quatre (ensembles d')opérateurs qui manipulent des actions sont définis à partir d'une seule et même relation d'accessibilité.

Nous disposons de *lois d'action* nous permettant de lier une action avec ses préconditions et ses effets.

3.1 Croyances sur les actions

Au-delà de la simple juxtaposition des axiomes pour les opérateurs de croyance et d'action, il nous faut déterminer des principes nouveaux qui rendent

compte de la relation entre les croyances et les actions. Nous proposons l'axiome suivant (BelAct$_1$) : $Bel_i After_\alpha A \rightarrow Bel_i After_\alpha Bel_i A$
La seule possibilité pour invalider cet axiome est que α efface toute ou une partie de la mémoire de i, cas extrêmes que nous excluons ici.
Nous adoptons également(BelAct$_2$): $Bel_i After_\alpha Bel_i A \rightarrow Bel_i After_\alpha A$
Cet axiome est non intuitif seulement si α rajoute des informations non justifiées à la mémoire du robot Il semble qu'il n'est pas possible d'aller au-delà de ces principes sans une analyse plus fine des actions.

3.2 Principe d'observation totale

Afin de simplifier le formalisme présenté, nous supposons par la suite que l'occurrence d'une action est publique, donc perçue par tous les agents. Les deux axiomes suivants formalisent cette idée :
(Public$_1$) $Bel_i Done_\alpha \top \rightarrow Done_\alpha \top$ et (Public$_2$) $Done_\alpha \top \rightarrow Bel_{\mathcal{AGT}} Done_\alpha \top$
(Public$_1$) stipule que la perception de l'occurrence d'une action est correcte, tandis que (Public$_2$) exprime qu'elle est complète. On peut montrer que (Public$_2$) est déductivement équivalent à $After_\alpha Bel_{\mathcal{AGT}} Done_\alpha \top$

3.3 Actions perceptives et décomposabilité

À partir de maintenant, nous considérons que, du point de vue d'un agent i, toute action α peut être décomposée en :

- une action non perceptive $\alpha^{i\not\hookleftarrow}$ à laquelle i associe *a priori* un ensemble d'effets (ceux associés à toute action de ce type *via* ses lois d'action) ;
- une perception $\alpha^{i\hookleftarrow}$ des effets réels de cette instance d'action.

Du point vue de i, $\alpha^{i\not\hookleftarrow}$ modifie son environnement, et $\alpha^{i\hookleftarrow}$ uniquement ses croyances. Par exemple, l'action de lancer une pièce correspondrait à $(i : \mathsf{lancer})^{j\not\hookleftarrow}; (i : \mathsf{lancer})^{j\hookleftarrow}$, où $(i : \mathsf{lancer})^{j\not\hookleftarrow}$ est l'action où j observe i lancer la pièce sans regarder le résultat (« les yeux fermés »), et $(i : \mathsf{lancer})^{j\hookleftarrow}$ est l'action où j vérifie si la pièce est tombée sur pile ou sur face.

Nous pouvons alors relier plus étroitement les croyances et les actions non perceptives (où α est une action non-perceptive) :

$Bel_i After_\alpha Bel_i A \wedge \neg Bel_i After_\alpha \bot \rightarrow After_\alpha Bel_i A$ (NonPerc$_1$)

$Feasible_\alpha Bel_i A \rightarrow After_\alpha Bel_i A$ (NonPerc$_2$)

$After_\alpha Bel_i A \wedge \neg After_\alpha \bot \rightarrow Bel_i After_\alpha Bel_i A$ (NonPerc$_3$)

Ainsi, pour $\alpha = (i : \mathsf{lancer})^{j\not\hookleftarrow}$ et $A = \neg BelIf_i$, (NonPerc$_1$) et (NonPerc$_3$) expriment que l'incertitude de i sur le résultat de l'action provenant du non déterminisme de celle-ci, est préservée au travers de l'exécution de sa partie non-perceptive $(i : \mathsf{lancer})^{j\not\hookleftarrow}$ [4]. L'axiome (NonPerc$_2$) exprime ce qu'on pourrait appeler le « déterminisme épistémique » des actions non perceptives.

4. Notons que cet axiome ne serait pas acceptable sans nos hypothèses d'observation totale (Public$_1$) et (Public$_2$).

Notons que (NonPerc_2) et (NonPerc_3) peuvent être combinés pour donner $\neg Bel_i After_\alpha A \rightarrow After_\alpha \neg Bel_i A$ [10].

Quant aux actions perceptives, les plus importantes sont les actions communicatives et les actions de test (alias « knowledge-gathering »). Nous avons examiné ces dernières dans [11] et nous allons ici nous concentrer sur les actions communicatives.

3.4 Inertie

Afin de pouvoir utiliser notre logique dans des applications pratiques, il nous faut résoudre le problème du décor (le *frame problem*, encore appelé le *problème de l'inertie*), et trouver une manière de représenter ce qui ne change pas lors d'une action.

Étant donné que nous avons fortement lié l'effet d'une action non-perceptive sur l'état mental d'un agent, aux croyances de cet agent sur l'effet de cette action, il est alors possible de résoudre ce problème. Scherl et Levesque ont utilisé la solution de Reiter et ont appliqué la régression [22], nous pensons adopter la solution présentée dans [13] basée sur la notion de topique.

3.5 Les actions de communication

Nous supposons qu'il y a dans le langage des actions particulières qui sont accomplies, et que l'on appelle *actes de langage*, suivant en cela [1, 23, 26, 27].

Les actes assertifs sont de la forme $i : \mathsf{informs}A$, où A est une formule[5]. Nous supposons que si un acte peut être accompli, alors son auteur croit le contenu propositionnel de cet acte (condition de sincérité des assertifs). Soit, formellement : ($\text{Exec}_{i:\mathsf{informs}A}$) $Feasible_{i:\mathsf{informs}A} \top \rightarrow Bel_i A$ Cet axiome est équivalent à $Before_{i:\mathsf{informs}A} Bel_i A$. À l'aide de notre principe de préservation, nous pouvons déduire $After_{i:\mathsf{informs}A} Bel_i A$ à partir de ($\text{Exec}_{i:\mathsf{informs}A}$)[6]. Ce résultat est particulièrement intéressant, car il rend compte d'une propriété remarquable des actes de langage, à savoir que la condition de sincérité d'un acte n'est pas altérée par l'exécution de cet acte. Dans un cadre comme le nôtre, où les actions sont décomposées en deux phases (exécution et observation du résultat), c'est plus généralement vrai pour toute précondition.

Nous supposons aussi que les actions communicatives sont déterministes :

$$Feasible_{i:\mathsf{informs}A} B \rightarrow After_{i:\mathsf{informs}A} B \qquad (\text{Det}_{i\,:\,\mathsf{informs}A})$$

5. Afin de simplifier le formalisme, nous ne considérons que deux agents. Cela nous permet de considérer que le locuteur ne s'adresse pas à quelqu'un en particulier. De plus, étant donné que les actions sont publiques, nous n'avons pas besoin d'un second agent en argument de nos actions.
6. Selon nos principes de préservation, les actes du type $i : \mathsf{informs}A$ n'influent pas les croyances de i au sujet de A, c.-à-d. nous avons $Bel_i A \rightarrow After_{i:\mathsf{informs}A} Bel_i A$. Nous avons donc $Feasible_{i:\mathsf{informs}A} \top \rightarrow After_{i:\mathsf{informs}A} Bel_i A$. Étant donné que nous avons également $\neg Feasible_{i:\mathsf{informs}A} \top \rightarrow After_{i:\mathsf{informs}A} Bel_i A$ (puisque *After* a les axiomes de la logique modale K), nous obtenons $After_{i:\mathsf{informs}A} Bel_i A$.

Ceci n'est pas une hypothèse forte si l'on conçoit l'indéterminisme comme le fait qu'une même action peut avoir des effets différents dans un contexte d'exécution différent. Nous rendons compte de ce phénomène dans [3] pour les actes de langage indirects.

Finalement, le fait de dire quelque chose ne change pas le monde physique :

$$C \rightarrow \mathit{After}_{i:\mathsf{informs}A} C \text{ si } C \text{ est une formule objective} \qquad (\mathrm{Pres}_{i\,:\,\mathsf{informs}A})$$

Nous avons supposé que les agents sont tous coopératifs. Ceci justifie de considérer que d'autres types d'actes de communication peuvent être déduits des actes assertifs (comme c'est montré dans [3]). Nous ne conservons donc qu'un seul type d'acte : celui d'information. Ainsi, en respect avec la théorie des actes de langage, la question fermée i : $\mathsf{queryYN}A$ (i demande si A est vrai) peut être exprimée par l'acte i : $\mathsf{informs}\mathit{Int}_i\mathit{BelIf}_i A$, et le directif i : $\mathsf{request}j$: a (i demande que j fasse a) par l'acte i : $\mathsf{informs}\mathit{Int}_i\mathit{Done}_{j:a}\top$ [7].

4 Sémantique

Comme nous l'avons vu, à l'exception des opérateurs Int_i, nous ne faisons intervenir que des opérateurs modaux normaux, et pour tous les axiomes les caractérisant il existe un résultat général de complétude dû à Sahlqvist [21] garantissant la complétude de notre système par rapport à la classe de modèles correspondante.

Les opérateurs Int_i sont du type des opérateurs des logiques modales classiques, logiques qui admettent une sémantique « de voisinage ». Ces logiques sont décrites dans [7] où, par ailleurs, est présentée une méthode permettant de traduire ces logiques dans des logiques multimodales *normales*. Ainsi, *via* la traduction évoquée, nous obtenons une logique ne faisant intervenir que des opérateurs normaux. Qui plus est, les axiomes les caractérisant sont tous des axiomes de Sahlqvist [21] ce qui garantit les bonnes propriétés de l'ensemble de notre système.

5 Applications

Comme nous le montrons dans ce qui suit, nous pouvons, à l'aide des notions définies jusqu'à présent, formaliser certains principes liés aux architectures BDI, et en particulier la coopération (Sect. 5.1), la perception (Sect. 5.2) et la rétroaction (Sect. 5.3).

7. Plus précisément, nous considérons qu'un acte de type assertif constitue en réalité un directif si son contenu propositionnel est une intention du locuteur de réaliser une certaine propriété. Il s'ensuit l'algorithme suivant : si i : $\mathsf{informs}A$ est l'acte accompli, alors : si A est de la forme $\mathit{Int}_i A$, alors l'acte est un directif, sinon c'est un assertif.

5.1 Formalisation de la coopération

Le comportement coopératif d'un agent i vis-à-vis d'un autre agent j peut être analysé en terme d'influence. Plus précisément, il s'agit de l'influence des croyances de i à propos d'attitudes mentales de j, sur les propres attitudes mentales de i. Nous distinguons :

- l'adoption de croyances ;
- l'adoption d'intentions ;
- la génération de nouvelles intentions.

Un agent i adopte la croyance qu'un autre agent j lui a communiquée d'après un critère de compétence qu'il accorde à j à propos de cette croyance. Soit formellement

$$Bel_i(Bel_j A \rightarrow A) \text{ si pour } i,\ j \text{ est compétent sur } A \qquad (\text{Adoption}_{\text{Croyance}})$$

En l'état, ceci est un axiome non logique. Dans [13] nous avons proposé de l'intégrer dans la logique en formalisant la compétence à l'aide de la notion de topique associé à une formule et un agent.

Si un agent i apprend qu'un agent j a l'intention que A alors, selon le statut doxastique de A [8] :

- i va adopter l'intention que j croie A si i croit lui-même que A, ou
- il va adopter l'intention de croire A d'abord s'il ignore si A (sous réserve que i n'ait pas déjà une intention contraire), ou
- il va adopter l'intention que A s'il croit que $\neg A$ (sous réserve que i n'ait pas déjà une intention contraire).

Nous notons que les deux derniers principes sont en accord avec $(\text{Exec}_{i:\mathsf{informs}A})$. Formellement :

$$Bel_i Int_j A \wedge Bel_i A \rightarrow Int_i Bel_j A \qquad (\text{Adoption}_{\text{Intention1}})$$
$$Bel_i Int_j A \wedge \neg BelIf_i A \wedge \neg Int_i \neg A \rightarrow Int_i Bel_i A \qquad (\text{Adoption}_{\text{Intention2}})$$
$$Bel_i Int_j A \wedge Bel_i \neg A \wedge \neg Int_i \neg A \rightarrow Int_i A \qquad (\text{Adoption}_{\text{Intention3}})$$

Nous soulignons que dans le second cas, dès que l'intention de i est satisfaite, il va automatiquement adopter l'intention que j croie A à son tour.

Au delà de l'adoption de croyances et d'intentions, un autre principe coopératif consiste à générer des intentions destinées à satisfaire des besoins de ses interlocuteurs. Ce principe peut être formalisé par différents axiomes dépendant des caractéristiques que l'on souhaite affecter aux agents. Par exemple :

$$Bel_i(A \wedge Bel_j \neg A \rightarrow Int_i Bel_j A) \qquad (\text{Correction}_{\text{Croyance}})$$
$$Bel_i(A \wedge Done_{j:\mathsf{informs}A} Bel_i \neg A \rightarrow Int_i Bel_j Bel_i A) \qquad (\text{Acquittement}_{\text{Croyance}})$$
$$Bel_i(Done_\alpha(Done_\gamma \top \wedge Bel_i Done_\beta \top) \rightarrow Int_i Bel_j Bel_i Done_\alpha Done_\gamma \top) \qquad (\text{Interprétation}_{\text{Erronée}})$$

8. Afin d'être cohérent avec l'axiome (Réalisme)

Enfin, l'axiome suivant rend compte du fait qu'un agent i ne peut chercher à ce qu'un agent j croit A sans que lui-même croie A (cet axiome est relié au principe de la sincérité) :

$$Int_i Bel_j A \rightarrow Bel_i A \vee Int_i Bel_i A. \qquad (\text{Sinc}_1)$$

Il existe d'autres axiomes dans la littérature que nous ne détaillons pas ici (cf. [4, 19, 18]).

5.2 Communication et perception

L'action i : $\mathsf{informs}A$ n'est pas perceptive pour i (mise à part la perception de sa propre exécution, garantie par nos hypothèses d'observation totale (Public_1) et (Public_2) puisque i n'apprendra rien de nouveau en l'exécutant (*i.e.* rien de plus que ce que ses croyances sur cet acte ne lui auraient déjà permis d'inférer avant de l'exécuter). Formellement, ceci correspond au fait que les formules ci-dessous constituent des instances acceptables des axiomes (NonPerc_1), (NonPerc_2) et (NonPerc_3) :

$$Bel_i After_{i:\mathsf{informs}A} Bel_i B \wedge \neg Bel_i After_{i:\mathsf{informs}A} \bot \rightarrow After_{i:\mathsf{informs}A} Bel_i B$$
$$Feasible_{i:\mathsf{informs}A} Bel_i B \rightarrow After_{i:\mathsf{informs}A} Bel_i B$$
$$After_{i:\mathsf{informs}A} Bel_i B \wedge \neg After_{i:\mathsf{informs}A} \bot \rightarrow Bel_i After_{i:\mathsf{informs}A} B$$

Ceci est relié à la préservation des préconditions d'un acte de communication (mais, comme on peut le montrer, ne l'implique pas). En particulier il s'agit ici de la précondition de pertinence au contexte de l'action i : $\mathsf{informs}A$, à savoir $Bel_i \neg BelIf_j A$. Nous avons choisi de considérer que cette précondition n'est abandonnée que lorsqu'il y a perception. [9]

Un peu plus surprenant est que nos axiomes de non perception sont également acceptables pour Bel_j, c.-à-d.

$$Bel_j After_{i:\mathsf{informs}A} Bel_j B \wedge \neg Bel_j After_{i:\mathsf{informs}A} \bot \rightarrow After_{i:\mathsf{informs}A} Bel_j B$$
$$Feasible_{i:\mathsf{informs}A} Bel_j B \rightarrow After_{i:\mathsf{informs}A} Bel_j B$$
$$After_{i:\mathsf{informs}A} Bel_j B \wedge \neg After_{i:\mathsf{informs}A} \bot \rightarrow Bel_j After_{i:\mathsf{informs}A} B$$

Ceci peut s'expliquer par le fait que les actions de communication en général, et i : $\mathsf{informs}A$ en particulier, sont déterministes ($(\text{Det}_{i:\mathsf{informs}A})$).

5.3 Formalisation de la rétroaction

Enfin, il est intéressant de se demander à quel niveau de la communication intervient la perception. Notre hypothèse est qu'à part la perception correcte et complète des occurrences des actions (exprimée par les axiomes (Public_1) et (Public_2) peut être assimilée à la *rétroaction* (le *feedback*).

9. $After_{i:\mathsf{informs}A} Bel_i A$ peut être déduit à partir de ($\text{Exec}_{i:\mathsf{informs}A}$), cf. la note 6 de bas de page.

Nous considérons que la rétroaction de l'auditeur j suite à un acte α de i est la réponse immédiate de l'auditeur relative au statut des conditions de succès et de non défectuosité de cet acte[10].

Comme nous nous situons dans le cadre d'un dialogue et que chaque agent est coopératif, nous pouvons considérer que la rétroaction est une réaction immédiate, que nous notons $j : \mathsf{retro}(\alpha)$. De plus, notre hypothèse de sincérité implique que tout acte est accompli avec succès, nous tâchons d'analyser la rétroaction par rapport à la non défectuosité.

Considérons un acte de type informatif $i : \mathsf{informs}A$. Il est non défectueux si j ignore si A, donc si $\neg \mathit{BelIf}_j A$. En fonction des croyances de j, la rétroaction de j sera alors $j : \mathsf{retro}(i : \mathsf{informs}A) = \begin{pmatrix} \mathsf{if}\ \mathit{BelIf}_j A\ \mathsf{then}\ j : \mathsf{informs}\mathit{BelIf}_j A \\ \mathsf{else}\ j : \mathsf{informs}\neg \mathit{BelIf}_j A \end{pmatrix}$

6 En guise conclusion : traitement de l'exemple

Nous traitons dans ce qui suit l'exemple de dialogue présenté en introduction. Comme il n'existe pas une infinité de numéros de téléphone, par soucis de simplification il est acceptable de représenter l'énoncé u_1 « Quel est le numéro de téléphone de k? » par la question fermée « Est-ce que le numéro de téléphone de k est n? », que nous représentons par l'action $u : \mathsf{queryYN}n$. Nous avons identifié ici cet acte à $u : \mathsf{informs}\mathit{Int}_u(\mathit{Done}_{s:\mathsf{informs}n}\top \vee \mathit{Done}_{s:\mathsf{informs}\neg n}\top))$.

Dans ces conditions, nous souhaitons savoir à quoi correspond la réaction que u peut attendre de s (*i.e.* à quoi correspond la rétroaction $s : \mathsf{retro}(u : \mathsf{queryYN}n)$).

Nous pouvons supposer que le fait suivant fait partie de la croyance commune : *l'interrogation des renseignements téléphoniques aboutira* ($\mathit{BelIf}_u n$) *dès lors que l'abonné n'est pas sur la liste rouge* ($\neg r$). Soit :

$$\mathit{Bel}_{\{u,s\}}(\neg r \rightarrow \mathit{After}_{u:\mathsf{queryYN}n;s:\mathsf{retro}(u:\mathsf{queryYN}n)}\, \mathit{BelIf}_u n)$$

Soit $\alpha = u : \mathsf{queryYN}n; s : \mathsf{retro}(u : \mathsf{queryYN}n)$. Les actes de communication ne modifiant pas le monde physique, nous avons $\neg r \rightarrow \mathit{After}_\alpha \neg r$. Des deux formules précédentes nous pouvons alors déduire $\mathit{Bel}_u(\mathit{After}_\alpha r \vee \mathit{After}_\alpha \mathit{BelIf}_u n)$
Comme *After* obéit à l'axiome K, nous obtenons $\mathit{Bel}_u \mathit{After}_\alpha(r \vee \mathit{BelIf}_u n)$.
L'axiome relatif aux actions non perceptives ($\mathrm{NonPerc}_1$) garantit alors que $\mathit{Bel}_u \mathit{After}_\alpha \bot \vee$ ∎

10. La distinction entre ces deux ensembles de conditions est la suivante [25]. Les conditions de succès, dans la théorie des actes de langage, correspondent à ce qui est communément appelé « préconditions » : ce sont toutes les conditions que doit respecter le locuteur pour accomplir « avec succès » son acte de langage. Mais même si les conditions de succès, du point de vue du locuteur, sont remplies, celui-ci peut se tromper (il a fait une supposition erronée, par exemple), ou ne pas exprimer les bons états mentaux (conditions inadaptées à la situation) : on dit alors que l'acte est défectueux. Par exemple, si le Maire déclare marié deux jeunes personnes, mais qu'il s'avère par la suite que l'une d'entre elle était déjà mariée, alors la déclaration du Maire est défectueuse, et les personnes n'ont pas été mariés (bien que la déclaration du Maire ait été faite avec succès.)

$After_\alpha Bel_u(r \vee BelIf_u n)$ (en l'appliquant en deux temps).
Si nous supposons que u croit que α est exécutable $Bel_u \neg After_\alpha \bot$ et nous obtenons finalement $After_\alpha Bel_u(r \vee BelIf_u n)$.
Avec les axiomes de KD45 pour la croyance, nous pouvons donc en déduire $After_\alpha Bel_u(Bel_u r \vee BelIf_u n)$ ce qui est équivalent à :

$$After_\alpha Bel_u(\neg BelIf_u n \rightarrow Bel_u r)$$

Nous avons ainsi établi que si après l'acte u_1 suivi de la rétroaction s_2, u continue à ignorer si n, alors u apprend que l'agent dont il cherche le numéro de téléphone est sur la liste rouge.

Notons que notre exemple (dans la version formelle simplifiée présentée ici) semble poser des problèmes à presque tous les formalismes proposés récemment en Intelligence Artificielle pour raisonner sur les connaissances et les actions[11].

Références

1. John L. Austin. *Quand dire, c'est faire.* Éditions du Seuil, France, 1970.
2. Michael E. Bratman. *Intention, Plans, and Practical Reason.* Harvard University Press, Cambridge, MA, 1987.
3. Maud Champagne, Rémi Faure, Andreas Herzig, Dominique Longin, Christophe Luc, Jean-Luc Nespoulous, et Jacques Virbel. « Formalisation logique de la communication non littérale à la lumière d'aperçus pragmatiques et neuropsycholinguistiques ». Dans *Journées Francophones Modèles Formels de l'Interaction (MFI'01)*, 2001. 12 pages.
4. Philip R. Cohen et Hector J. Levesque. « Intention is Choice with Commitment ». *Artificial Intelligence Journal*, 42(2–3), 1990.
5. Philip R. Cohen et Hector J. Levesque. Rational Interaction as the Basis for Communication. Dans Philip R. Cohen, Jerry Morgan, et Martha E. Pollack, éditeurs, *Intentions in Communication.* MIT Press, 1990.
6. Ronald Fagin, Joseph Y. Halpern, Yoram Moses, et Moshe Y. Vardi. *Reasoning about knowledge.* MIT Press, 1995.
7. Olivier Gasquet et Andreas Herzig. From Classical to Normal Modal Logics. Dans Heinrich Wansing, éditeur, *Proof Theory of Modal Logics*, numéro 2 dans Applied Logic Series, pages 293–311. Kluwer Academic Publishers, 1996.
8. Joseph Y. Halpern et Yoram Moses. « A guide to the modal logics of knowledge and belief: preliminary draft ». Dans *Proc. Int. Joint Conf. on Artificial Intelligence (IJCAI'85).* Morgan Kaufmann Publishers, 1985.
9. David Harel. Dynamic Logic. Dans Dov M. Gabbay et Franz Günthner, éditeurs, *Handbook of Philosophical Logic*, volume II, pages 497–604. D. Reidel, Dordrecht, 1984.
10. Andreas Herzig, Jérôme Lang, Dominique Longin, et Thomas Polacsek. « A logic for planning under partial observability ». Dans *Proc. Seventeenth National Conf. on Artificial Intelligence (AAAI-2000)*, 2000.

11. Merci à Jorge Lobo pour avoir attiré notre attention sur cet exemple.

11. Andreas Herzig, Jérôme Lang, et Thomas Polacsek. « A modal logic for epistemic tests ». Dans *Proc. Eur. Conf. on Artificial Intelligence (ECAI'2000)*, Berlin, août 2000.
12. Andreas Herzig et Dominique Longin. « Belief Dynamics in Cooperative Dialogues ». Dans Jan van Kuppevelt, Noor van Leusen, Robert van Rooy, et Henk Zeevat, éditeurs, *Proc. Third Int. Workshop on the Semantics and Pragmatics of Dialogue (Amstelogue'99)*, 1999. 13 pages.
13. Andreas Herzig et Dominique Longin. « Belief Dynamics in Cooperative Dialogues ». *Journal of Semantics*, 17(2), 2000. 20 pages.
14. Jaakko K. K. Hintikka. *Knowledge and belief.* Cornell University Press, Ithaca, N.Y., 1962.
15. Kurt Konolige et Martha E. Pollack. « A Representationalist Theory of Intention ». Dans *Proc. 13th Int. Joint Conf. on Artificial Intelligence (IJCAI'93).* Morgan Kaufmann Publishers, 1993.
16. Hector J. Levesque. « What is planning in the presence of sensing ». Dans *Proc. Nat. Conf. on AI (AAAI'96).* AAAI Press, 1996.
17. Dominique Longin. « *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, 1999. `http://www.irit.fr/ACTIVITES/LILaC`.
18. Anand S. Rao et Michael P. Georgeff. « Modeling Rational Agents within a BDI-Architecture ». Dans J. A. Allen, R. Fikes, et E. Sandewall, éditeurs, *Proc. Second Int. Conf. on Principles of Knowledge Representation and Reasoning (KR'91)*, pages 473–484. Morgan Kaufmann Publishers, 1991.
19. M. D. Sadek. « *Attitudes mentales et interaction rationnelle : vers une théorie formelle de la communication* ». Thèse de doctorat, Université de Rennes I, Rennes, France, 1991.
20. M. D. Sadek. « Dialogue Acts are Rational Plans ». Dans M.M. Taylor, F. Néel, et D.G. Bouwhuis, éditeurs, *The structure of mutimodal dialogue*, pages 167–188, Philadelphia/Amsterdam, 2000. John Benjamins publishing company. From ESCA/ETRW, Workshop on The Structure of Multimodal Dialogue (Venaco II), 1991.
21. H. Sahlqvist. « Completeness and correspondence in the first and second order semantics for modal logics ». Dans S. Kanger, éditeur, *Proc. 3rd Scandinavian Logic Symposium*, volume 82 de *Studies in Logic*, 1975.
22. Richard Scherl et Hector J. Levesque. « The frame problem and knowledge producing actions ». Dans *Proc. Nat. Conf. on AI (AAAI'93)*, pages 689–695. AAAI Press, 1993.
23. John R. Searle. *Les actes de langage.* Hermann, Paris, 1972.
24. John R. Searle. *L'Intentionalité : Essai de philosophie des états mentaux.* Les Éditions de Minuit, 1985.
25. D. Vanderveken. *Les actes de discours : essai de philosophie du langage et de l'esprit sur la signification des énonciations.* Mardaga, Liège – Bruxelles, 1988.
26. D. Vanderveken. *Meaning and Speech Acts*, volume 1 : *Principles of language use.* Cambridge University Press, 1991.
27. D. Vanderveken. *Meaning and Speech Acts*, volume 2 : *Formal semantics of success and satisfaction.* Cambridge University Press, 1991.