

# Modélisation de la capacité d'un agent intentionnel en fonction de ses activités

Karl Devooght<sup>1,2</sup>

<sup>1</sup> France Télécom R&D

2, avenue Pierre Marzin, 22300 Lannion

<sup>2</sup> LLI-IRISA, 6, rue de Kerampont, 22300 Lannion

karl.devooght@orange-ftgroup.com

**Résumé** : Depuis plusieurs années, les systèmes d'agents intelligents, en particulier les agents dit BDI (*t* *Belief-Desire-Intention*), font l'objet d'un intérêt croissant. Dans cet article, nous proposons une logique modale de la capacité s'appuyant sur les modèles logiques BDI. En particulier, la capacité d'un agent est comprise comme la capacité à réaliser une activité pour atteindre un but. Une *activité* est un sous-modèle BDI résolvant une classe de problèmes particuliers. Dans cette optique, nous proposons un modèle d'agent dans lequel les activités de ce dernier sont organisées en contextes. La capacité permet alors d'établir le lien entre un agent et ses activités. Pour cela, nous proposons une structure sémantique ainsi qu'un système axiomatique adaptés à la logique de la capacité proposée. **Mots-clés** : capacité, activité, agent intelligent .

---

## 1 Introduction

Depuis plusieurs années, les systèmes d'agents intelligents font l'objet d'un intérêt grandissant. Les agents intelligents permettent de modéliser des processus complexes (par exemple, le contrôle du trafic aérien ou les services de renseignements en dialogue naturel). Les approches agent combinent habituellement logique symbolique et approche cognitive. En particulier, le paradigme d'agent dit BDI (Belief-Desire-Intention) a été largement étudié. Il présente des bases théoriques robustes comme, par exemple, les travaux de (Cohen & Levesque, 1990), de (Rao & Georgeff, 1995), de (Sadek, 1994)). Par ailleurs, plusieurs systèmes implémentés s'appuient sur ce paradigme.

Afin de considérer la dynamique d'un agent, les modèles logiques BDI prennent généralement en compte la notion d'action. Cela permet à l'agent aussi bien de réagir à court terme que de poursuivre des buts à plus long terme. Pour cela, les modèles d'agents doivent prendre en compte un large ensemble de connaissances et de descriptions d'activités. De plus, l'agent doit s'adapter à l'évolution des situations dans

lesquelles il est impliqué. Par ailleurs, pour une situation particulière, seul un sous-ensemble des connaissances et des activités est généralement nécessaire à l'agent.

Comme la plupart des modèles d'agents BDI ne partitionnent pas explicitement les connaissances et les descriptions d'activités de l'agent, deux problèmes importants surviennent. Premièrement, plus la quantité de connaissances et d'activités considérée est grande, plus il est difficile de maintenir un modèle d'agent consistant. Deuxièmement, le manque de modularité au sein des modèles BDI oblige l'agent à considérer à chaque instant toutes ses connaissances. Ainsi, pour garantir la cohérence d'un agent et l'accès aisé aux connaissances pertinentes, on est contraint de limiter les capacités de l'agent.

En pratique, (Busetta *et al.*, 1999) ont proposé de modulariser chaque description d'activité de l'agent. L'idée est de les considérer comme des "modules mentaux" à part entière permettant de résoudre une classe de problème particulier. Pour établir le lien entre l'agent et ses activités modularisées, ils introduisent une notion de capacité. Autrement dit, la capacité comme attitude mentale particulière de l'agent prend en paramètre une activité, et réfère au "module mental" de celle-ci.

Dans ce papier, nous formalisons la notion de capacité suivant cette idée. Elle est comprise comme la capacité d'un agent à réaliser une activité pour atteindre un certain but. Nous proposons en particulier un cadre sémantique dédié aux modèles d'agents BDI, composé de contextes. Sous cet angle, notre opérateur de capacité signifie que l'agent accède au contexte de l'activité eét que la description de l'activité permet d'atteindre un but. Nous montrons par ailleurs l'intérêt de la capacité proposée notamment en termes de généralité.

En section 2, nous introduisons un bref état de l'art des modèles de capacité existants. Ensuite, nous présentons informellement notre modèle de la capacité. En section 3, un cadre sémantique basé sur la sémantique de Kripke est proposé consistant à partitionner les modèles d'agent en contexte. Dans cette optique, une logique de la capacité étendant une logique BDI est considérée. En particulier, les propriétés caractérisant la logique sont précisées. Nous discutons en section 4 des conséquences conceptuelles induites par le modèle d'agent proposé.

## 2 Capacité et raisonnement

### 2.1 État de l'art

Modéliser la notion de capacité implique d'explicitier les paramètres qui la composent. Il est évident que, en tant qu'acteur, l'agent lui-même ou, dans certains cas, le groupe d'agents représente un de ces paramètres. Il est cependant plus difficile de déterminer la nature de ce sur quoi porte la capacité. Par exemple, lorsque nous affirmons que *Ronaldo est capable de marquer un but*, est-il entendu que Ronaldo est capable de réaliser une action particulière (ou un plan particulier) consistant à tirer au but ? Ou est-il entendu que Ronaldo est capable de garantir qu'un but soit marqué indépendamment des moyens mis en jeu (en passant la balle, par exemple, à un coéquipier mieux placé que lui) ? Le choix d'une des deux possibilités dépend fortement du modèle d'action considéré. Autrement dit, définir un modèle de capacité exige au préalable d'opter pour un modèle d'action particulier.

Dans la littérature, il existe deux approches principales pour modéliser l'action. La première consiste à ne pas représenter explicitement les actions dans le langage logique. Dans ce cadre, les travaux sur la capacité expriment alors le fait qu'un agent parvient à un état du monde où un but est atteint (ou une propriété est vraie) (cfr. les travaux de (Brown, 1988), de (Horty & Belnap, Jr., 1995)). La seconde, au contraire, prend en compte explicitement les actions de l'agent. C'est, par exemple, le cas de la logique dynamique. L'accent est alors mis sur le fait qu'un agent est capable de réaliser une certaine action atomique ou complexe (van der Hoek *et al.*, 1994; van Linder *et al.*, 1998). Par ailleurs, en connaissant les effets d'une action, une forme dérivée de capacité peut être modélisée. Il s'agit de considérer l'effet d'une action comme une propriété vraie dans l'état du monde atteint après l'exécution de l'action.

En outre, le sens donné à la capacité doit être également désambiguïté. On distingue fréquemment la capacité *générique* de la capacité *occasionnelle*. La capacité générique réfère à l'ensemble des connaissances<sup>1</sup> permettant, par exemple, de marquer un but. Ce type de capacité est indépendant des états intellectuels et physiques ainsi que de l'état du monde. La capacité occasionnelle (appelée aussi *opportunité*<sup>2</sup>) réfère au contraire à la situation courante et aux conditions satisfaites dans celle-ci autorisant à appliquer une capacité générique particulière. Si Ronaldo possède le ballon et se trouve seul face au gardien, on comprend qu'il a l'opportunité de marquer un but.

Techniquement, la considération logique de ces deux types de capacité est complexe. Il est en effet difficile de les définir sans que chacun fasse référence à l'autre (cf. (Singh, 1991)). C'est pourquoi de nombreux travaux proposent une modélisation de la capacité combinant capacité générique et capacité occasionnelle. A notre connaissance, seuls les travaux de van Linder *et al.* (van Linder *et al.*, 1998) ainsi que ceux de Cholvy *et al.* (Cholvy *et al.*, 2005) introduisent une séparation claire.

## 2.2 Capacité et activité

Dans cet article, nous nous intéressons aux modèles de capacité générique dans lesquels l'action est explicitement représentée dans le langage. De tels modèles associent généralement la capacité d'un agent à sa connaissance de plans. Un *plan* se caractérise d'une part, par la connaissance d'un ensemble de préconditions et d'effets caractérisant *a priori* et *a posteriori* l'exécution du plan, et d'autre part, par la description de l'exécution du plan (i.e. appelée aussi le corps du plan). L'exécution d'un plan<sup>3</sup> se définit souvent comme une séquence particulière d'actions. Une capacité générique d'agent est alors comprise comme une capacité à réaliser un plan. La capacité à réaliser un plan s'interprète comme la connaissance par l'agent de la séquence d'actions caractérisant l'exécution du plan. Dans notre exemple, marquer un but pourrait être associé à un plan consistant à dribbler un adversaire puis à tirer au but.

La question qui se pose est la suivante : suffit-il d'un plan pour exprimer une capacité générique ? En pratique, Busetta *et al.* Busetta *et al.* (1999) ont montré que ce n'est pas le cas. Selon eux, déterminer une capacité générique implique au moins la prise en

---

<sup>1</sup>Ces connaissances sont parfois appelées des connaissances *procédurales*

<sup>2</sup>Dans cet article, les deux termes sont utilisés indifféremment.

<sup>3</sup>Les plans considérés sont complètement spécifiés et déterministes.

compte de connaissances <sup>4</sup>, de plans (i.e. éventuellement plus d'un plan) et de règles de comportements (e.g., guidant le choix du plan). Dans cette optique, être capable de marquer un but nécessiterait certaines connaissances (e.g., où suis-je sur le terrain, où se trouvent mes coéquipiers), plusieurs plans (e.g., dribbler un adversaire puis tirer au but ou tirer directement au but), des règles de comportement (e.g., si je suis en position de tirer, je tire au but). L'ensemble de ces éléments caractérise ce que Busetta et al. nomme un composant fonctionnel. Un composant fonctionnel regroupe tous les éléments permettant de résoudre une classe particulière de problème. Dans ce papier, nous nommons cela une *activité* de l'agent.

Dans le contexte des modèles d'agent BDI, à quoi correspond une activité d'agent ? Il existe en fait une analogie profonde entre la modélisation d'un agent et la modélisation d'une activité d'agent. En effet, modéliser un agent BDI particulier, c'est proposer dans une certaine mesure un modèle de son activité globale (e.g., l'activité de contrôler le trafic aérien). Par conséquent, on peut penser que modéliser une activité équivaut à définir un sous-modèle BDI permettant de résoudre une certaine classe de problème. Sous cette hypothèse, il s'agit néanmoins de souligner quelles sont les différences entre le modèle d'agent et le modèle de ses activités. Premièrement, l'agent possède une *vue* sur ses activités. Dans notre approche, cela se traduit par le biais de la capacité. Un agent capable de réaliser une activité est avant tout un agent ayant certaines informations sur le modèle de cette activité. Deuxièmement, l'agent est à un niveau global alors que l'activité est toujours à un niveau local (i.e., caractérisée par la classe de problème qu'elle est censée résoudre). En dernier lieu, nous distinguons la notion de *réalité* mentale de la notion de *projection* mentale. La réalité mentale de l'agent détermine les états mentaux "réels" de l'agent, c'est-à-dire ceux liés à son passé, à sa situation présente et à ses possibles situations futures. Autrement dit, la réalité mentale est la partie *située* du modèle d'agent, c'est-à-dire la partie composée d'éléments effectifs constituant la perception du monde que possède l'agent. La plupart des travaux envisage un modèle d'agent de cette manière. Une projection mentale réfère à des états mentaux imaginés par l'agent (qui ne réfère pas directement à la réalité mentale de l'agent). Par exemple, Ronaldo sait qu'il est en train de jouer un match de football (ce qui représente sa réalité mentale) et à la mi-temps, dans un moment d'égarement, il peut s'imaginer en train de cuisiner sa recette préférée (qui est de l'ordre de la projection mentale). Nous pensons que les modèles d'activité à travers la notion de capacité doivent être perçus par l'agent comme des projections mentales.

La notion de capacité que nous introduisons permet d'établir un lien entre la réalité mentale de l'agent et ses activités considérées comme des projections mentales. Cette vision nous amène à une interprétation plus intuitive de la capacité générique. De nombreux travaux interprètent la capacité générique comme une analyse des états futurs de l'agent envisageant ce qu'il peut faire. Cependant, cette position n'est pas adaptée dans de nombreux cas. Par exemple, imaginons un homme condamné à la prison à vie. Si nous lui demandons naïvement : "Es-tu capable de rouler à vélo ?". Il n'y aurait rien de choquant à ce qu'il nous réponde par l'affirmative. Il est fortement vraisemblable qu'il n'analyse pas cela sur la base de son avenir puisqu'à priori il ne roulera certainement jamais plus à vélo. Par contre, envisager la même situation sous l'angle de la projection

---

<sup>4</sup>Les connaissances sont ici des connaissances *factuelles*.

mentale offre une vision plus adéquate.

En résumé, notre considération de la capacité générique porte non plus sur l'action mais sur l'activité. Une activité se caractérise par un sous-modèle BDI résolvant une classe de problèmes particuliers. De plus, être capable de réaliser une activité ne s'interprète pas comme une analyse des états mentaux futurs de l'agent mais plutôt comme l'analyse du sous-modèle associé à l'activité et projeté mentalement par l'agent. Nous introduisons ainsi un modèle formel dans lequel les activités de l'agent sont organisées en modules.

### 3 Une logique de la capacité

Dans cette section, nous définissons une logique modale propositionnelle étendant une logique BDI "classique". En particulier, nous considérons les opérateurs modaux de croyance  $B$ , de but (ou désir)  $D$  ainsi que l'opérateur d'intention  $I$ . Cette approche de la logique BDI s'inspirent des travaux de Rao et Georgeff (Rao & Georgeff, 1995). Un opérateur  $C(\alpha, \phi)$  est introduit. Il signifie que l'agent est capable de réaliser l'activité  $\alpha$  pour atteindre le but  $\phi$ . Par ailleurs, nous nous situons dans un cadre mono-agent<sup>5</sup>.

*Syntaxe.*

Un langage  $\mathcal{L}$  est introduit. Il contient un ensemble de propositions atomiques  $\mathcal{P}$ , un ensemble d'actions  $\mathcal{ACT}$  ainsi qu'un ensemble d'activités primitives  $\mathcal{A}$ . Une formule bien formée  $\phi$  de  $\mathcal{L}$  respecte la grammaire suivante :

$$\begin{aligned}\phi &::= p \mid \neg\phi \mid \phi \rightarrow \phi \mid B\phi \mid D\phi \mid I\phi \mid C(\alpha, \phi) \\ a &::= e \mid a; a\end{aligned}$$

Notons que  $p \in \mathcal{P}$ ,  $e \in \mathcal{ACT}$  et  $\alpha \in \mathcal{A}$ . Les connecteurs logiques  $\wedge$  et  $\vee$  sont pris en compte et sont définis de manière habituelle.

*Sémantique.*

Les modèles sémantiques traditionnels pour les logiques modales s'appuient sur des structures de modèles basées sur des cadres sémantiques de Kripke. Un cadre sémantique de Kripke se compose d'un ensemble d'états (ou de mondes possibles)  $W$  et d'un ensemble de relations binaires  $R$  sur  $W$  (i.e.  $W \times W$ ). Nous proposons une forme dérivée de ces cadres. Nous considérons des cadres sémantiques composés de :

- un ensemble  $C$  de contextes. Un contexte  $c$  est un tuple  $(W_c, R_c)$  où :
  - $W_c$  est un ensemble d'états (ou mondes possibles)
  - $R_c$  est un ensemble de relations binaires sur  $W_c$
- un ensemble  $R_{ic}$  de relations binaires "inter-contextes" de la forme  $W_c \times W_{c'}$  (où  $c \neq c'$ ).

Remarquons que ces cadres se présentent comme une forme restreinte des cadres classiques  $(W, R)$  où :

---

<sup>5</sup>C'est pourquoi, les opérateurs modaux ne sont pas relatifs à un agent particulier

- $W = \bigcup_{c \in C} W_c$
- $R = R_{ic} \bigcup (\bigcup_{c \in C} R_c)$

La "réalité mentale" de l'agent est représentée par un contexte particulier  $c_{RM}$ . Chaque activité  $\alpha \in \mathcal{A}$  est associée à un contexte noté  $c_\alpha$ . Ces contextes représentent la réalisation des activités. En particulier, l'ensemble des états  $W_{c_\alpha}$  de chaque contexte  $c_\alpha$  se compose d'un ensemble d'états initiaux  $S_{c_\alpha}$  ainsi qu'un ensemble d'états finaux  $F_{c_\alpha}$ . Autrement dit, nous supposons que la réalisation d'une activité débute à partir d'un état  $s \in S_{c_\alpha}$  et se termine dans un état  $f \in F_{c_\alpha}$ .

L'ensemble des relations d'accessibilité  $R$  pour chaque contexte  $c \in C$  se compose de :

- une relation d'accessibilité de croyance  $\mathcal{B}_c$ ,
- une relation d'accessibilité de but  $\mathcal{D}_c$ ,
- une relation d'accessibilité d'intention  $\mathcal{I}_c$ ,
- une relation d'accessibilité événementielle  $R_c(a)$  pour chaque action  $a \in \mathcal{ACT}$ .

Nous supposons que chaque relation  $R_c(a)$  est contrainte sous les hypothèses de passé linéaire <sup>6</sup>, de futur ramifié <sup>7</sup> et de déterminisme <sup>8</sup>. Par ailleurs, pour toute séquence d'action  $a1; a2$ , nous notons par abus de langage  $R_c(a1; a2)$ . Si nous avons  $R_c(a1; a2)(w, w'')$ , cela est équivalent à dire qu'il existe  $w'$  tel que  $R_c(a1)(w, w')$  et  $R_c(a2)(w', w'')$ .

L'ensemble  $R_{ic}$  des relations d'accessibilité inter-contextes correspond dans notre approche à un ensemble de relations d'accessibilité  $R(\alpha)$  pour chaque activité  $\alpha \in \mathcal{A}$ . Ces relations se présentent sous la forme  $W_{c_{RM}} \times S_{c_\alpha}$  reflétant la possibilité d'accéder à certains états initiaux des contextes  $c_\alpha$ , à partir d'états  $W_{c_{RM}}$  du contexte  $c_{RM}$ . Ces relations d'accessibilité caractérisent ce que nous avons nommé les "projections mentales" de l'agent.

Chaque contexte possède son propre vocabulaire. Un *vocabulaire* est un sous-ensemble de propositions atomiques pour lesquelles une interprétation sémantique est possible. Une proposition atomique n'appartenant pas au vocabulaire d'un contexte  $c$  est supposée fautive dans chacun des états de  $c$ .

Ainsi, les modèles  $M$  dont nous nous servons pour interpréter le langage  $L$  sont des tuples  $(C, R_{ic}, V, \pi)$  où :

- $C$  est l'ensemble de contextes composé de  $c_{RM}$  et des contextes  $c_\alpha$  pour chaque activité  $\alpha \in \mathcal{A}$ ,
- $R_{ic}$  est l'ensemble de relations d'accessibilité composé de  $R_{ic}(\alpha)$  où  $\alpha \in \mathcal{A}$ ,
- $V$  est l'ensemble des vocabulaires  $V_c \subseteq \mathcal{P}$  pour  $c \in C$ ,
- $\pi$  est l'ensemble des fonctions d'interprétation  $\pi_c : V_c \times W \rightarrow \text{Bool}$  pour  $c \in C$ .

### Satisfiabilité.

Nous pouvons à présent définir la relation de satisfiabilité  $\models$ . Soit  $\phi$  une formule bien formée (fbf) de  $\mathcal{L}$ ,  $M$  un modèle  $(C, R, V, \pi)$ ,  $w_c$  appartenant au contexte  $c$  de  $C$ . On note  $M, w_c \models \phi$  pour signifier que le monde  $w_c$  satisfait (ou rend vraie) la fbf  $\phi$ . Elle est *contextuellement satisfaite* ssi  $M, w_c \models \phi$  pour tout  $w_c \in W_c$ . Elle est valide ssi  $M, w \models \phi$

<sup>6</sup>C'est-à-dire, pour tout état  $w$ , il n'existe qu'une et une seule action  $a$  et un état  $w'$  tel que  $R_c(a)(w, w')$ .

<sup>7</sup>C'est-à-dire, pour chaque état  $w$ , il existe une action  $a$  et un état  $w'$  tel que  $R_c(a)(w, w')$ .

<sup>8</sup>C'est-à-dire, si  $R_c(a)(w, w')$  et  $R_c(a)(w, w'')$ , alors  $w' = w''$

pour tout couple  $(M, w)$ . La relation de satisfiabilité  $\models$  est définie récursivement comme suit :

- $M, w_c \models p$  ssi  $p \in V_c$  et  $\pi_c(p, w_c) = \text{vrai}$  ;
- $M, w_c \models \neg \phi$  ssi  $M, w_c \not\models \phi$
- $M, w_c \models \phi \wedge \psi$  ssi  $M, w_c \models \phi$  et  $M, w_c \models \psi$
- $M, w_c \models B\phi$  ssi pour tout  $w'_c$  tq  $B_c(w_c, w'_c)$ , on a  $M, w'_c \models \phi$
- $M, w_c \models D\phi$  ssi pour tout  $w'_c$  tq  $D_c(w_c, w'_c)$ , on a  $M, w'_c \models \phi$
- $M, w_c \models I\phi$  ssi pour tout  $w'_c$  tq  $I_c(w_c, w'_c)$ , on a  $M, w'_c \models \phi$

Avant de donner l'interprétation sémantique de l'opérateur de capacité  $C$ , nous introduisons la notion de stratégie. Une *stratégie* d'activité, noté  $\sigma(a, s, f)$ , pour une activité  $\alpha$  associée à un contexte  $c_\alpha$  est une action primitive ou une séquence d'actions  $a$ , telle que  $R(a)_\alpha(s, f)$  où  $s \in S_{c_\alpha}$  et  $f \in F_{c_\alpha}$ . Autrement dit, une stratégie d'activité est une action particulière partant d'un état initial de l'activité et atteignant un de ses états finaux. L'opérateur de capacité  $C$  est interprété sémantiquement comme suit :

- $M, w_c \models C(\alpha, \phi)$  ssi
  - pour tout  $s_{c_\alpha} \in S_{c_\alpha}$  tq  $R_{ic}(\alpha)(w_c, s_{c_\alpha})$  et
  - pour toute stratégie  $\sigma(a, s_{c_\alpha}, f_{c_\alpha})$ , on a  $M, f_{c_\alpha} \models \phi$ .

Remarquons que cette formule ne peut éventuellement être satisfaite que pour  $w_c$  où  $c = c_{RM}$ . Autrement dit, nous supposons que la capacité est seulement satisfiable dans les états mentaux "réels" de l'agent.

Attardons nous maintenant sur l'interprétation sémantique de cet opérateur. Informellement, cet opérateur signifie que l'agent se projette mentalement dans un état initial particulier de l'activité  $\alpha$ . A partir de cet état, toutes les stratégies (projetées) de l'agent atteignent un état dans lequel  $\phi$  est vraie. Notre opérateur modal de capacité s'apparente à un opérateur de nécessité. La nécessité est exprimée par l'accès aux états initiaux de l'activité. De plus, la nécessité porte sur les états finaux de l'activité et notamment sur le fait qu'une propriété doit être vraie dans chacun de ces états. Le caractère nécessaire de la capacité est critiqué dans certaines approches. L'argument le plus souvent avancé est qu'envisager une réussite totale est une hypothèse irréaliste. Dans notre approche, l'ensemble des états finaux  $F_{c_\alpha}$  pour une activité  $\alpha$  doit être compris comme les états de *succès* de l'activité. Dès lors, exiger qu'une propriété soit vraie dans tous les états de succès est une hypothèse acceptable.

*Axiomes.*

L'opérateur de capacité que nous proposons est un opérateur modal normal. Les propriétés de notre opérateur caractérisent une logique modale KD. En particulier, l'opérateur est clos sous l'implication logique, la conjonction et la disjonction. Il est par ailleurs consistant. La consistance de l'opérateur signifie, par exemple, qu'être capable de réaliser une activité qui me rend riche implique le fait de ne pas être capable de réaliser cette même activité avec, pour but, d'être pauvre. Ainsi, l'opérateur de capacité

admet donc les axiomes suivants :

$$C(\alpha, \phi) \wedge C(\alpha, \phi \Rightarrow \psi) \Rightarrow C(\alpha, \psi) \quad (\text{K})$$

$$C(\alpha, \phi) \Rightarrow \neg C(\alpha, \neg\phi) \quad (\text{D})$$

$$C(\alpha, \phi \wedge \psi) \Leftrightarrow C(\alpha, \phi) \wedge C(\alpha, \psi)$$

$$C(\alpha, \phi \vee \psi) \Leftrightarrow C(\alpha, \phi) \vee C(\alpha, \psi)$$

En particulier, la validité de l'axiome D est due à la nature des relations d'accessibilité  $R_{ic}(\alpha)$  et  $R_c(a)$ . Nous imposons, en effet, que  $R_{ic}(\alpha)$  soit sérielle, c'est-à-dire, pour tout état  $w_{c_{RM}}$ , il existe un état  $s_{c_\alpha}$  tel que  $R(\alpha)(w_{c_{RM}}, s_{c_\alpha})$ . Par ailleurs, les états finaux atteints (à travers les relations d'accessibilité  $R_c(a)$ ) à partir d'un état initial particulier  $s_{c_\alpha}$  sont, par nature, des états consistants.

## 4 Conséquences conceptuelles

### 4.1 Analyse locale

Modéliser la capacité sur la base d'une sémantique contextuelle offre deux avantages importants. Premièrement, elle permet de décrire et d'analyser *localement* une activité par le biais de la capacité. L'intérêt logique est de ne pas ou plus chercher à établir un système axiomatique au modèle d'agent qui soit adapté à toutes les activités de l'agent. Par contre, il est possible d'établir les systèmes axiomatiques convenant individuellement à chacune des activités. Dans notre approche, cela correspond à identifier ce qui est contextuellement satisfait. Deuxièmement, la capacité comme attitude mentale abstrait la description d'une activité. Ainsi, un état "réel" d'agent donné ne prend pas en compte chaque description détaillée d'activité dans sa phase de raisonnement. Par ailleurs, l'agent est préservé d'un effort éventuel de raisonnement qu'il fournirait si chacune des descriptions d'activités étaient "physiquement" présentes dans chacun de ces états mentaux.

Pour illustrer le premier point, prenons un exemple. Rao et Georgeff dans (Rao & Georgeff, 1995) proposent un système axiomatique pour la logique BDI. Le système contient huit axiomes. Ils introduisent un neuvième axiome exprimant les conditions sous lesquelles un agent s'engage à accomplir une intention. Trois alternatives sont considérées : l'agent poursuit son intention (1) tant qu'il ne croit pas que le but est atteint ("*blind commitment*"), (2) tant qu'il croit que le but n'est pas atteint et que le but est possible à réaliser ("*single-mind commitment*") ou (3) tant qu'il croit que le but n'est pas atteint et que ce dernier est encore un but de l'agent ("*open mind commitment*"). Pour eux, le choix d'une alternative dépend du profil d'agent à définir.

Dans notre approche, nous pensons que cela dépend non seulement de l'agent mais particulièrement de l'activité de l'agent à considérer. Par exemple, imaginons que Ronaldo ait son brevet de secouriste. Supposons qu'il soit capable de marquer des buts et de sauver une vie. Intuitivement, ces deux activités n'adhèrent pas au même type d'engagement. D'un côté, pour sauver une vie, Ronaldo poursuivra son intention de la sauver jusqu'au moment où il y arrive. L'axiome "*blind commitment*" est dans ce cas adapté. D'un autre côté, s'engager à marquer un but semble être contraint par le fait que

cet objectif soit encore réalisable. L'axiome "*single-mind commitment*" semble mieux convenir. Ces axiomes sont dès lors parfaitement envisageables au sein d'un même modèle d'agent. Seulement, l'ensemble de ces axiomes ne sont pas consistants mais, par contre, sont contextuellement satisfaits dans le cadre des activités données en exemple.

## 4.2 Opportunité

Comme nous l'avons mentionné en section 2, la capacité générique accompagne automatiquement la capacité occasionnelle. Bien que nous n'ayons pas formalisé cette dernière, nous devons évoquer quelques perspectives intéressantes.

Traditionnellement, quand elle possède une action comme paramètre, la capacité occasionnelle est caractérisée par le fait qu'il existe une exécution de l'action. Supposons que  $w$  soit un état du monde et  $a$  une action. En termes logiques, un agent a la capacité occasionnelle de réaliser l'action  $a$  à partir de l'état  $w$  si et seulement s'il existe un état du monde  $w'$  résultant de l'exécution de  $a$ . Cependant, nous pensons que la capacité occasionnelle ne doit pas être liée à quelconque état résultant puisque, par définition, celle-ci réfère "strictement" à l'état du monde *avant* l'exécution d'une action sans pré-supposer *a priori* l'existence d'une telle exécution. Par exemple, une personne peut avoir l'opportunité de jouer au tennis car elle possède une raquette et un adversaire, et se trouve sur un terrain de tennis. Supposons qu'elle ne sache pas comment jouer au tennis i.e. qu'elle n'ait pas la compétence. Dans ce cas, elle n'a donc *a priori* pas "conscience" de l'existence d'une exécution particulière d'action permettant de jouer au tennis.

Dans notre approche, la capacité générique est indépendante de l'état mental "réel" dans lequel se trouve l'agent. En particulier, elle réfère à un état initial de l'activité projeté mentalement par l'agent. Dans cette optique, la capacité occasionnelle peut être interprétée comme la compatibilité entre l'état mental "réel" et l'état mental initial de l'activité. Cela signifie, par exemple, que avoir l'opportunité de jouer au tennis résulte de la comparaison entre un état courant "réel" et un état initial de l'activité "jouer au tennis" imaginé par l'agent. Il s'agit alors de définir quand deux états sont compatibles. Une solution extrême est d'exiger que les états mentaux soient identiques. Une solution plus raisonnable est de définir un ensemble fini de conditions devant être satisfaites dans chacun des deux états. De cette manière, les notions de capacité générique et de capacité occasionnelle sont indépendantes. Avoir l'opportunité n'influence pas le fait d'avoir ou non la capacité générique (et réciproquement).

## 5 Conclusion

Nous avons proposé dans cet article un modèle logique de la capacité générique. La notion d'activité a été introduite et a été considérée comme la nature de ce sur quoi porte la capacité. Chaque activité est associée à un sous-modèle BDI explicitant "mentalement" sa réalisation. Le modèle général d'agent est ainsi décomplexifié. Par ailleurs, la capacité comme attitude mentale permet d'établir un lien entre l'agent et ses activités.

Formellement, la capacité générique est comprise comme la capacité générique à réaliser une activité pour atteindre un certain but. Un système axiomatique a été proposé

caractérisant notre logique de la capacité comme une logique modale KD.

De plus, nous avons discuté de deux importantes conséquences conceptuelles de notre modèle. Premièrement, une logique d'agent BDI devrait être analysée en fonction des activités qu'un agent est amené à réaliser. Il s'agit alors de déterminer ce qui est vrai pour *une* activité et non pour l'ensemble des activités. Deuxièmement, la capacité générique est indépendante des états mentaux "réels" (intellectuels ou physiques) de l'agent. En particulier, nous avons proposé des pistes pour définir la capacité occasionnelle dans le cadre du modèle logique proposé.

En termes de perspectives, plusieurs pistes sont ouvertes. Premièrement, avec l'opérateur de capacité présenté, la réalisation du but est *nécessairement* vrai alors que, dans certaines activités, certains buts sont simplement *possibles* selon le cas de figure. Par exemple, Ronaldo peut marquer un but avec son pied gauche (même s'il est droitier). Si son objectif est de marquer un but avec son pied gauche, il est intuitif que celui-ci est possiblement (et non nécessairement) atteignable. Deuxièmement, comme pour les modèles d'actions, il serait intéressant que nous puissions paramétrer une activité. Ainsi, nous pourrions distinguer un type d'activité et une instance d'activité (i.e. un type d'activité dont les paramètres sont instanciés). Troisièmement, nous avons restreint la satisfaisabilité d'une capacité aux états "réels" de l'agent. Il est envisagé de généraliser cela à n'importe quel état. Dans ce cas, cela signifie que, dans les contextes associés aux activités, il serait possible de faire référence à une autre capacité (et, par conséquent, à un autre contexte).

## Références

- BROWN M. A. (1988). On the logic of ability. *Journal of Philosophical Logic*, **17**, 1–26.
- BUSETTA P., HOWDEN N., RÖNNQUIST R. & HODGSON A. (1999). Structuring BDI agents in functional clusters. In N. R. JENNINGS & Y. LESPÉRANCE, Eds., *ATAL*, volume 1757 of *Lecture Notes in Computer Science*, p. 277–289 : Springer.
- CHOLVY L., GARION C. & SAUREL C. (2005). Ability in a multi-agent context : A model in the situation calculus. In F. TONI & P. TORRONI, Eds., *CLIMA VI*, volume 3900 of *Lecture Notes in Computer Science*, p. 23–36 : Springer.
- COHEN P. R. & LEVESQUE H. J. (1990). *Intention is choice with commitment*. Toronto : Computer Science Dept., University of Toronto.
- HORTY J. F. & BELNAP, JR. N. D. (1995). The deliberative STIT : A study of action, omission, and obligation. *Journal of Philosophical Logic*, **24**(6), 583–644.
- RAO A. S. & GEORGEFF M. P. (1995). BDI agents : From theory to practice. In *ICMAS*, p. 312–319.
- SADEK M. D. (1994). Communication theory = rationality principles + communicative act models.
- SINGH M. P. (1991). A logic of situated know-how. In *IJCAI-91 Proceedings* : IJCAI.
- VAN DER HOEK W., VAN LINDER B. & MEYER J.-J. C. (1994). A logic of capabilities. *Proceedings of Laboratory for Foundations of Computer Science*, **813**, 366–378.
- VAN LINDER B., VAN DER HOEK W. & MEYER J.-J. C. (1998). Formalising abilities and opportunities of agents. *Fundamenta Informatica*, **34**.