Modélisation du raisonnement multiagent : une approche basée sur les labels

Imed Jarras — **Brahim Chaib-draa**

Département d'informatique, Faculté des Sciences
Université Laval, Sainte-Foy, PQ, Canada, G1K 7P4
Email : {jarras, chaib}@ift.ulaval.ca

RÉSUMÉ. *Devant l'intérêt sans cesse grandissant aux systèmes multiagents durant cette dernière décade, le développement d'outils formels pour l'analyse, la description et l'implantation de ces systèmes est, aujourd'hui, plus que nécessaire. La plupart des méthodes formelles développées jusqu'à date sont basées sur la sémantique des mondes possibles. Cette dernière bien qu'élégante est handicapée par deux grands problèmes : 1) le problème de l'omniscience et 2) le problème de mécanisation. Dans notre approche, un agent est défini comme étant un système LDS muni d'un ensemble de mécanismes comme l'action, l'abduction et la mise à jour. Dans le présent article, nous présentons une modélisation d'agents par des systèmes logiques basés sur les LDS (systèmes déductifs étiquetés) de Gabbay. Le modèle obtenu est appliqué par la suite au problème bien connu des n sages (raisonnement sur autrui).*

MOTS-CLÉS : *Systèmes multiagents, logique épistémique, LDS, omniscience logique.*

1. Introduction

Pour certaines recherches, encore en stage théorique, un agent est souvent représenté par ses états mentaux, c'est à dire ses engagements, ses croyances, ses buts, ses intentions, etc. [SC94, Woo92].Ces états mentaux permettent de comprendre, d'analyser et/ou de prédire le comportement de l'agent. Les méthodes formelles sont sans doute nécessaires à la modélisation des systèmes complexes comme les systèmes multiagents. Ces méthodes permettent d'avoir une expression claire des propriétés du système multiagent à modéliser, ainsi qu'une vision cohérente et transparente de ce que le système fait et ne fait pas.

Jusqu'à présent, il y a peu d'outils formels pour l'analyse, la spécification, la conception ou l'implantation des SMA. Devant l'inadéquation de la logique classique, la plupart des travaux formels proposés sont basés sur des logiques non classiques comme la logique modale. Cette dernière est considérée comme l'une des logiques les plus attractives grâce à sa sémantique des mondes possibles. Cependant, elle souffre d'un problème non négligeable, celui de l'*omniscience logique*.

Un autre inconvénient majeur de ces travaux est qu'ils sont, pour la plupart, purement théoriques. Or, la validation sur des exemples pratiques et réels s'avère aujourd'hui nécessaire. Les modèles formels déjà utilisés sont plutôt orientés vers la "description" et ne constituent pas une méthode formelle "mécanisable". Ceci est principalement dû à la "non-mécanisation" des logiques modales non-classiques principalement lorsqu'on veut élaborer un modèle formel où coexistent plusieurs fragments de logiques modales non classiques (dynamique, temporelle, épistémique, etc.) comme c'est le cas pour les SMA.

En plus des connaissances et des croyances, la plupart des travaux prennent en compte dans leur modèle le temps et/ou l'action. La combinaison de plusieurs logiques modales (temporelle et dynamique) s'avère alors nécessaire. Dans ce cas la complexité du modèle augmente et dans presque tous les cas, un tel modèle devient indécidable. Il reste néanmoins que la mécanisation d'un tel modèle comprenant plusieurs fragments non classiques serait très utile pour la conception et l'implémentation d'un système multiagent. Jusqu'ici, hormis quelques tentatives très fragmentaires (par exemple ALX de Huang [Hua94]) il n'y a pas, à notre connaissance, un outil logique tournant sur une machine et permettant de concevoir des agents considérés comme des systèmes logiques capables, entre autres, de faire des déductions de manière autonome. Ce présent travail est une première étape vers cet objectif dans la mesure où nous proposons de concevoir des agents autour d'une logique labélisée traduisible en logique de premier ordre, de manière à utiliser toute la panoplie des prouveurs de théorèmes de cette dernière. Il s'agit des LDS (*Labelled deductive systems*) de Gabbay [Gab94].

Cet article est organisé de la manière suivante. La notion de systèmes logiques vus comme des systèmes déductifs étiquetés est détaillée à la section suivante. La modélisation des agents par des systèmes logiques basés sur les LDS est traitée dans la section 3. À la section 4, nous présentons les règles d'inférences classiques, de connaissance et structurales. Finalement, dans la section 5, le modèle est étendu pour prendre en compte des systèmes multiagents et est appliqué à l'exemple bien connu des n sages.

2. Les systèmes logiques vus comme des systèmes déductifs labélisés

Les LDS sont un formalisme basé sur les étiquettes. Celles-ci donnent généralement un supplément d'information aux formules. Les logiques modales et temporelles sont des logiques pour lesquelles il serait très utile d'utiliser des étiquettes. La logique modale, par exemple, manipule des formules relatives à un ensemble de mondes possibles reliés par une relation d'accessibilité. Il paraît donc naturel de nommer et de référer explicitement à ces mondes et à leur relation en utilisant des étiquettes. Les étiquettes nous permettent également de distinguer le monde dans lequel nous sommes actuellement.

Pour ces raisons, et pour bien d'autres [Gab94], il nous paraît adéquat d'utiliser une approche basée sur les LDS, où nous utilisons les étiquettes pour nommer les mondes et un langage d'étiquetage pour la description de la relation entre mondes.

Par ailleurs, selon l'auteur des LDS [Gab94], ceux-ci offrent un passage moins compliqué (que les logiques non classiques) vers la logique classique. Par conséquent, cela nous donne la possibilité de tirer profit de tout ce qu'offre la logique classique : prouveurs, outils de vérification, etc.

Dans un premier temps, nous nous intéressons à formaliser, en utilisant les LDS et la déduction naturelle, les systèmes modaux pour les connaissances dans un environnement multiagent. Nous proposons une méthodologie basée sur l'utilisation d'un système logique en vue de concevoir des agents intelligents.

La LDS offre une ossature formelle assez générale comme nous allons le voir. Dans ce contexte précis des LDS, une logique est une paire $(\vdash, \mathbf{S}_\vdash)$, où $\vdash$ est une relation de conséquence structurée sur un langage **L** et $\mathbf{S}_\vdash$ est un système LDS. Pour présenter un système LDS, nous avons besoin de définir son ensemble de formules et son ensemble d'étiquettes. Par exemple, nous pouvons prendre le langage modal comme ensemble de formules (avec variables, constantes et quantificateurs) et prendre certains ensembles de fonctions des mêmes variables et constantes comme ensemble d'étiquettes. Les logiques traditionnelles manipulent des formules alors que les LDS manipulent des paires ⟨*étiquette : formule*⟩ appelées *unités déclaratives*. L'étiquette peut être vue comme un supplément d'information qui n'est pas codé dans la formule. La manipulation de cette extra information est de nature différente de celle contenue dans le prédicat. En général, l'étiquette t d'une unité déclarative $\langle t : \phi \rangle$ peut avoir plusieurs interprétations. t peut représenter une valeur de fiabilité (systèmes à base de connaissances), l'origine de ϕ (bases de données complexes), la date d'entrée ou de validité de ϕ (base de données historique ou temporelle), temps où ϕ a lieu (logique temporelle), monde possible où ϕ est vraie (logique modale), une preuve de ϕ etc. La notion traditionnelle de conséquence entre formules de la forme $A_1, A_2, \cdots, A_n \vdash B$ est par conséquent remplacé par la notion de conséquence entre formules étiquetées $t_1 : A_1, t_2 : A_2, \cdots, t_n : A_n \vdash s : B$

3. Modélisation d'agents par des systèmes logiques basés sur les LDS : un système de déduction naturelle

Dans cette section, nous combinons LDS et déduction naturelle pour construire un système logique pour les agents. Nous adoptons les systèmes modaux déductives à étiquettes MLDS, *Modal Labelled Deductive Systems*, introduites par Alessandra Russo [Rus95, Rus96], qui sont une extension des systèmes déductifs étiquetés, LDS *Labelled Deductive Systems*, de Gabbay [Gab94]. Le choix des MLDS est justifié par le fait que ces derniers offrent un cadre combinant à la fois la représentation relationnelle du premier ordre de la structure des mondes possibles avec le langage modal standard.

Avant de donner les règles de déduction, nous présentons les définitions des langages utilisés.

3.1 Définition. (Langage labélisé $\mathcal{L}_L$) *Un langage labélisé $\mathcal{L}_L$ est un langage de* *premier ordre composé de :*

— $\{w_0, w_1, \cdots, w_n, \cdots\}$ *un ensemble dénombrable de constantes;*

— $\{x, y, z, \cdots\}$ *un ensemble dénombrable de variables;*

— R *prédicat binaire;*

— $\{\neg, \wedge, \vee, \rightarrow, \equiv\}$ *un ensemble de connecteurs logiques;*

— $\forall$ *et* $\exists$ *les quantificateurs universel et existentiel.*

Les constantes et les variables de $\mathcal{L}_L$ dénotent les mondes possibles, tandis que le prédicat binaire R formalise la relation d'accessibilité entre mondes. Le langage suivant permet la construction de formules épistémiques relative à une unité déclarative.

3.2 Définition. (Langage épistémique $\mathcal{L}_E$) *un langage épistémique $\mathcal{L}_E$ est un langage modale propositionnel composé de :*

— $\{p, q, r, \cdots\}$ *un ensemble de propositions atomiques;*

— $\{\neg, \wedge, \vee, \rightarrow, \equiv\}$ *un ensemble de connecteurs logiques;*

— K *l'opérateur de connaissance*[1]*.*

3.3 Définition. (Le langage épistémique déductive étiqueté (ELDL)) *Étant donné un langage étiqueté $\mathcal{L}_L$ et un langage épistémique $\mathcal{L}_E$, un langage épistémique déductive labélisé est la paire ordonnée $\langle \mathcal{L}_L, \mathcal{L}_E \rangle$*

3.4 Définition. (Langage étiqueté semi-étendu *Func*$(\mathcal{L}_L, \mathcal{L}_E)$**)**
Soit $\{\alpha_1, \alpha_2, \cdots, \alpha_n, \cdots\}$ l'ensemble ordonné de toutes les formules bien fondées de $\mathcal{L}_E$. Le langage de premier ordre Func$(\mathcal{L}_L, \mathcal{L}_E)$ est défini comme étant le langage $\mathcal{L}_L$ étendu par :

— $\{f_{\alpha_1}, f_{\alpha_2}, \cdots, f_{\alpha_n}, \cdots\}$ *un ensemble de fonctions unaires correspondante à chaque α_i de $\mathcal{L}_E$.*

— $\{box_{\alpha_1}, box_{\alpha_2}, \cdots, box_{\alpha_n}, \cdots\}$ *un ensemble de fonctions unaires correspondante à chaque formule α_i de $\mathcal{L}_E$.*

Pour tout monde possible w et pour toute formule α, $f_\alpha(w)$ dénote un monde *particulier* associé à α, alors que $box_\alpha(w)$ dénote un monde *arbitraire* associé à α. $f_\alpha(w)$ est utilisé pour formaliser les notions de la sémantique de Kripke de la forme *il existe un monde possible....* Par contre $box_\alpha(w)$ formalise la forme *pour tout monde possible....* Dans le cadre d'un système multiagent, il convient d'introduire f^i_α et box^i_α pour spécifier *il existe un monde possible accessible à l'agent* $i \cdots$ et *pour tout monde possible accessible à l'agent* $i \cdots$.

Ces fonctions ne sont utilisées que lors des dérivations. $box_\alpha(w)$ et $f_\alpha(w)$ peuvent dans certains modèles référer au même monde possible. Notons aussi que pour tout monde w et toutes propositions p et p' on a $\{box_p(w)\} = \{box_{p'}(w)\}$. Sans perte de généralité, pour toute proposition p, on peut écrire $box_p(w) = box(w)$. Par contre, on a généralement $\{f_p(w)\} \neq \{f_{p'}(w)\}$. Remarquons en plus que $\{f_p(w)\} \subseteq \{box(w)\}$.

Nous définissons maintenant la notion de R-littéral représentant la relation entre

1. On utilise le terme *connaissance* d'une façon générique : il peut s'agir de toute forme de croyance ou de savoir.

mondes.

3.5 Définition. (R-Littéral) *Étant donné un langage $\langle \mathcal{L}_L, \mathcal{L}_E \rangle$, un R-littéral est un littéral de la forme $R(w_1, w_2)$ ou bien $\neg R(w_1, w_2)$, où $w_1, w_2 \in$ Func$(\mathcal{L}_L, \mathcal{L}_E)$.*

$R(w_1, w_2)$ dénote que le monde w_2 est accessible au monde w_1, alors que son conjugué $\neg R(w_1, w_2)$ signifie que w_2 n'est pas accessible au monde w_1. Généralement, seuls les littéraux positifs, i.e. de la forme $R(w_1, w_2)$ sont décrits. Les littéraux négatifs sont utilisés pour dériver des contradictions ou introduire l'opérateur de négation.

3.1. *Un système de déduction naturelle basé sur les LDS pour la logique épistémique ELDS*

Par opposition aux systèmes modaux classique pour lesquels à tout moment, un et un seul monde possible est initial, les systèmes ELDS permettent à plusieurs mondes possibles d'être des mondes initiaux. Ainsi, la théorie des systèmes déductifs étiquetés est une théorie plus générale que la théorie modale classique. Informellement, une théorie écrite en ELDL consiste en un ensemble d'unité déclarative avec des littéraux de la forme $R(w_1, w_2)$ ou $\neg R(w_1, w_2)$ avec $w_1, w_2 \in$ *Func*$(\mathcal{L}_L, \mathcal{L}_E)$.

Nous appelons une telle théorie une *configuration* et la structure associée aux étiquettes un *diagramme*. Initialement, une configuration ainsi que le diagramme associé contiennent seulement des constantes appartenant à *Func*$(\mathcal{L}_L, \mathcal{L}_E)$ comme étiquettes. Les configurations contenant plusieurs termes de *Func*$(\mathcal{L}_L, \mathcal{L}_E)$ sont générées par l'application des règles d'inférence. Nous décrivons, formellement, les notions de diagramme et de configuration dans ce qui suit.

3.6 Définition. (Diagramme) *Étant donné le langage $\langle \mathcal{L}_L, \mathcal{L}_E \rangle$, un diagramme $\mathcal{D}$ est un ensemble de R-littéraux dont les arguments w_i, w_j, sont des termes de Func$(\mathcal{L}_L, \mathcal{L}_E)$.*

3.7 Définition. (Configuration) *Étant donné un ELDL, une configuration est un tuple $\langle \mathcal{D}, \mathcal{F} \rangle$ où :*

— $\mathcal{D}$ est un diagramme

— $\mathcal{F}$ est une fonction de l'ensemble des termes de Func$(\mathcal{L}_L, \mathcal{L}_E)$ à l'ensemble $\mathcal{P}$(fbf$(\mathcal{L}_E)$) des ensembles des formules bien fondées de $\mathcal{L}_E$.

Un exemple de configuration est représenté comme suit :

$$
\begin{array}{rcll}
\mathcal{D} & = & \{R(w_0, w_1), R(w_0, w_2), \neg R(w_1, w_2)\} & \\
\mathcal{F}(w) & = & \{K(p \rightarrow q), Kr\} & \text{si } w = w_0 \\
 & & \{\neg K \neg p, r \rightarrow s\} & \text{si } w = w_1 \\
 & & \{q, \neg K \neg p\} & \text{si } w = w_2 \\
 & & \{\} & \text{autrement}
\end{array}
$$

Nous avons montré jusqu'ici que l'ELDS est une généralisation des formalismes de la logique modale. Par conséquent, les règles d'inférence ne sauront plus définies entre ensembles de formules (théorie) et une formule, mais plutôt entre configurations. Les notions de *preuve* et de *dérivabilité* et les opérations d'*union*, de *différence* et d'*inclusion* sont généralisées d'une façon naturelle aux configurations.

Une autre notion essentielle doit être définie. C'est la notion d'algèbre d'étiquettes, écrite dans le langage *Func*$(\mathcal{L}_L, \mathcal{L}_E)$, et donnée par le sous-ensemble d'axiomes suivants :

$$\forall w(R(w,w)) \qquad \textbf{(T)}$$

$$\forall w, w', w''((R(w,w') \wedge R(w',w'')) \rightarrow R(w,w'')) \qquad \textbf{(4)}$$

$$\forall w, w' \ (R(w,w') \rightarrow R(w',w)) \qquad \textbf{(B)}$$

$$\forall w, \exists w' \ (R(w,w')) \qquad \textbf{(D)}$$

$$\forall w, w', w''((R(w,w') \wedge R(w,w'')) \rightarrow R(w',w'')) \qquad \textbf{(5)}$$

$$\forall w \ \neg R(w,w) \qquad \textbf{(Irr)}$$

Contrairement aux autres règles, l'axiome (**Irr**) n'a pas d'équivalent en logique modal classique comme c'est le cas pour les autres axiomes. Une algèbre d'étiquette contenant cet axiome, identifie les structures dont la relation d'accessibilité est non réflexive.

Deux algèbres différentes spécifient des systèmes différents. Pour différencier ces algèbres nous adoptons la notation suivante.

Noms	Algèbres d'étiquette
$\mathcal{A}_T$	$\{w : \forall w R(w,w)\}$
$\mathcal{A}_{Irr}$	$\{w : \forall w \ \neg R(w,w)\}$
$\mathcal{A}_{S4}$	$\{w : \forall w \ R(w,w), \forall w, w', w''((R(w,w') \wedge R(w',w'')) \rightarrow R(w,w''))\}$
$\mathcal{A}_{K4,Irr}$	$\{w : \forall w \ \neg R(w,w), \forall w, w', w''((R(w,w') \wedge R(w',w'')) \rightarrow R(w,w''))\}$

Il nous est possible maintenant de donner une définition complète de ce qu'est un système épistémique déductif étiqueté (ELDS).

3.8 Définition. (ELDS) *Étant donné un langage ELDL =* $\langle \mathcal{L}_L, \mathcal{L}_E \rangle$*, un système épistémique déductif étiqueté est un tuple de la forme* $\langle \langle \mathcal{L}_L, \mathcal{L}_E \rangle, \mathcal{A}, \mathcal{R} \rangle$ *où :*

$\mathcal{A}$ *est une algèbre écrite en Func*$(\mathcal{L}_L, \mathcal{L}_E)$*;*

$\mathcal{R}$ *un ensemble de règles d'inférences qui "génère" une configuration à partir d'une autre.*

Les systèmes suivants $\langle \langle \mathcal{L}_L, \mathcal{L}_E \rangle, \mathcal{A}_{S_4}, \mathcal{R} \rangle$ et $\langle \langle \mathcal{L}_L, \mathcal{L}_E \rangle, \mathcal{A}_{S_5}, \mathcal{R} \rangle$ représentent respectivement une extension des systèmes classiques S_4 et S_5.

La sous-section suivante présente les différentes règles d'inférence permettant des déductions entre configurations.

$$\mathcal{I}_{\wedge I} \quad \frac{\mathcal{C}\langle w : p, w : q\rangle}{\mathcal{C}'\langle w : p \wedge q\rangle} \qquad\qquad \frac{\mathcal{C}\langle w : p \wedge q\rangle}{\mathcal{C}'\langle w : p\rangle} \quad \mathcal{I}_{\wedge E}$$

$$\mathcal{I}_{\wedge E} \quad \frac{\mathcal{C}\langle w : p \wedge q\rangle}{\mathcal{C}'\langle w : q\rangle} \qquad\qquad \frac{\mathcal{C}\langle w : p\rangle}{\mathcal{C}'\langle w : p \vee q\rangle} \quad \mathcal{I}_{\vee I}$$

$$\mathcal{I}_{\vee I} \quad \frac{\mathcal{C}\langle w : q\rangle}{\mathcal{C}'\langle w : p \vee q\rangle} \qquad\qquad \frac{\mathcal{C}\langle w : p \vee q\rangle \quad \begin{array}{c}\mathcal{C}\langle [w : p]\rangle \\ \vdots \\ \tilde{\mathcal{C}}\langle w : r\rangle\end{array} \quad \begin{array}{c}\mathcal{C}\langle [w : q]\rangle \\ \vdots \\ \tilde{\mathcal{C}}\langle w : r\rangle\end{array}}{\mathcal{C}'\langle w : r\rangle} \quad \mathcal{I}_{\vee E}$$

$$\mathcal{I}_{\rightarrow I} \quad \frac{\begin{array}{c}\mathcal{C}\langle [w : p]\rangle \\ \vdots \\ \tilde{\mathcal{C}}\langle w : q\rangle\end{array}}{\mathcal{C}'\langle w : p \rightarrow q\rangle} \qquad\qquad \frac{\mathcal{C}\langle w : p, w : p \rightarrow q\rangle}{\mathcal{C}'\langle w : q\rangle} \quad \mathcal{I}_{\rightarrow E}$$

$$\mathcal{I}_{\neg I} \quad \frac{\begin{array}{c}\mathcal{C}\langle [w : p]\rangle \\ \vdots \\ \tilde{\mathcal{C}}\langle w' : \bot\rangle\end{array}}{\mathcal{C}'\langle w : \neg p\rangle} \qquad\qquad \frac{\mathcal{C}\langle w : \neg\neg p\rangle}{\mathcal{C}'\langle w : p\rangle} \quad \mathcal{I}_{\neg E}$$

Figure 1. *Les règles d'inférence classiques.*

3.2. *Les règles d'inférence*

Une règle d'inférence $\mathcal{I}$ est un ensemble de paires de configurations, où chaque paire est dénotée $\mathcal{C}/\mathcal{C}$'. Si $\mathcal{C}/\mathcal{C}' \in \mathcal{I}$, alors $\mathcal{C}$ est dite la *configuration antécédente* de $\mathcal{I}$, et $\mathcal{C}$' une configuration *de conséquence* ou *inférée* de $\mathcal{I}$. On dit que $\mathcal{I}$ infère $\mathcal{C}$' de $\mathcal{C}$, ou bien $\mathcal{I}$ génère $\mathcal{C}$' de $\mathcal{C}$. On distingue trois types de règles : les classiques, celles de l'opérateur de connaissance et les structurales.

3.2.1. *Les règles d'inférence classiques*

Les règles classiques se construisent aisément de la déduction naturelle classique [Lal90, RM90]. Elles sont représentées par la figure 1. À titre d'exemple, la règle $\mathcal{I}_{\wedge I}$ dénote l'introduction de l'opérateur de disjonction alors que la règle $\mathcal{I}_{\rightarrow E}$ n'est rien d'autre que le modus ponens dans un contexte de configurations.

3.2.2. *Les règles d'inférence de l'opérateur de connaissance*

Nous décrivons maintenant, les règles d'introduction et d'élimination de l'opérateur de connaissance K.

K**-élimination** $\mathcal{I}_{KE}$

Pour toute configuration $\mathcal{C}$, pour tous mondes w_1 et w_2, et toute formule bien fondée p, $\mathcal{C}/\mathcal{C} + [w_2 : p]$ est membre de la règle d'inférence $\mathcal{I}_{KE}$ si : $\mathcal{C}\langle w_1 : Kp\rangle$ et $\mathcal{C}\langle R(w_1, w_2)\rangle$. Cette règle d'inférence peut être représentée par

$$\frac{\mathcal{C}\langle w_1 : Kp, R(w_1, w_2)\rangle}{\mathcal{C}'\langle w_2 : p\rangle}$$

K**-introduction** $\mathcal{I}_{KI}$

Pour toute configuration $\mathcal{C}$, tout monde w et toute formule bien fondée p, $\mathcal{C}/\mathcal{C} + [w : Kp]$ est membre de la règle d'inférence $\mathcal{I}_{KI}$ si : $\mathcal{C} + [R(w, \mathit{box}(w))] \vdash_{\text{ELDS}} \mathit{box}(w) : p$

À noter que d'une façon générale $\vdash_{\mathbf{F}}$ signifie que la dérivation n'utilise que les axiomes et les règles du modèle **F**. Cette règle d'inférence peut être représentée par

$$\begin{array}{c} \mathcal{C}\langle [R(w, \mathit{box}(w))]\rangle \\ \vdots \\ \tilde{\mathcal{C}}\langle \mathit{box}(w) : p\rangle \\ \hline \mathcal{C}'\langle w : Kp\rangle \end{array}$$

Si p est vraie dans un monde arbitraire accessible à w alors on peut déduire que Kp est vraie dans w.

Remarquons que dans tous les cas qu'on vient de voir, on a une expansion de la théorie, i.e. de l'ensemble des unités déclaratives. Par contre, le diagramme n'a subit aucun changement. Les règles d'inférence structurales peuvent avoir des expansions et des réductions au niveau de la structure de la configuration (diagramme).

3.2.3. *Les règles d'inférence structurales*

Nous introduisons maintenant un troisième ensemble de règles d'inférence impliquant le raisonnement sur le diagramme $\mathcal{D}$ d'une configuration $\mathcal{F}$.

R**-Assertion,** $\mathcal{I}_{R-A}$

Pour toute configuration $\mathcal{C} = \langle \mathcal{D}, \mathcal{F}\rangle$ et tout R-littéral Δ, $\mathcal{C}/\mathcal{C} + [\Delta]$ est membre de la règle d'inférence $\mathcal{I}_{R-A}$ si : $\mathcal{D}, \mathcal{A} \vdash_{\text{FOL}} \Delta$ où $\mathcal{A}$ est une algèbre étiquetée et $\vdash_{\text{FOL}}$ est l'induction de la logique de premier ordre.

Cette règle facilite la déduction de nouveaux R-littéraux relativement à une relation d'accessibilité donnée par une algèbre étiquetée particulière $\mathcal{A}$. En effet, supposons que $\mathcal{A} = \mathcal{A}_{S4}$ et que par dérivation, on arrive à la situation suivante : $\{R(w, w'), R(w', w'')\} \subset \mathcal{D}$, alors d'après $\mathcal{A}_{S4}$ on déduit que $\{R(w, w'')\} \subset \mathcal{D}$.

$\perp$-Introduction, $\mathcal{I}_{\perp I}$

Nous avons introduit précédemment les règles d'inférence relatives à la négation (voir $\mathcal{I}_{\neg I}$, tableau 1). Sachant que $\neg p$ est une abréviation de $p \to \perp$, il est nécessaire de spécifier le sens du connecteur constant $\perp$ représentant le *faux* ou l'*absurde*. L'absurde en ELDS est soit un R-littéral et son conjugué, soit une unité déclarative et sa négation. Cette dernière est donnée pas la règle $\mathcal{I}_{\neg I}$ (Tableau 1), alors que le R-littéral et son conjugué est identifiée par la règle $\mathcal{I}_{\perp I}$ suivante :
Pour toute configuration $\mathcal{C}$, tout R-littéral Δ et toute unité déclarative $w : p, \mathcal{C}/\mathcal{C} + [w : p]$ est un membre de la règle d'inférence $\mathcal{I}_{\perp I}$ si : $\Delta \in \mathcal{C}$ et $\overline{\Delta} \in \mathcal{C}$. Ce qu'on peut représenter par

$$\frac{\mathcal{C}\langle \Delta, \overline{\Delta} \rangle}{\mathcal{C}'\langle w : p \rangle}$$

En utilisant les règles d'inférences $\mathcal{I}_{\neg I}$ et $\mathcal{I}_{\perp I}$, seules des unités déclaratives peuvent être dérivées à partir d'informations inconsistantes. Un raisonnement similaire peut être appliqué cependant aux R-littéraux. Dans la règle d'inférence suivante, un R-littéral est déduit à partir d'inconsistances dans une configuration.

R-Introduction, $\mathcal{I}_{R-I}$

Pour toute configuration $\mathcal{C}$, tout R-littéral Δ, $\mathcal{C}/\mathcal{C} + [\Delta]$ est un membre de la règle d'inférence $\mathcal{I}_{R-I}$ si pour tout terme w', on a $\mathcal{C} + [\overline{\Delta}] \vdash_{\text{ELDS}} w' : \perp$. Ce qu'on peut représenter par

$$\frac{\begin{array}{c}\mathcal{C}\langle \overline{\Delta} \rangle \\ \vdots \\ \tilde{\mathcal{C}}\langle w' : \perp \rangle\end{array}}{\mathcal{C}'\langle \Delta \rangle}$$

Dans la section suivante, nous traitons un autre type de raisonnement un peu plus complexe à savoir le raisonnement sur autrui. Ceci nous permet d'utiliser des opérateurs de connaissances enchâssés, (i.e. imbriqués). Le type de problème qu'on traitera n'est pas nouveau mais le type de démonstration contient une originalité. En effet, nous utilisons dans cette démonstration les règles d'inférences structurales. Ce qui est un des avantages de la LDS qui nous permet de manipuler les formules et les structures dans une même configuration.

Étant donné que nous allons utiliser une logique multi-modale (2 agents), une extension est par conséquent nécessaire. Nous associons pour chaque agent i, une relation R_i et des fonctions f_p^i et *box*^{i} pour chaque formule bien fondée p (voir [Jar97, Jar98]).

4. Raisonnement sur autrui

Pour illustrer ce qu'on vient de développer, nous allons donner un exemple, montrant à la fois l'application des règles d'inférence et un autre aspect intéressant de la logique modale en l'occurrence le *raisonnement sur autrui*. L'un des exemples le mieux adapté pour ce type de raisonnement est le *problème des n sages* (wise-men puzzle) [Fit83]. Nous ne considérons que le cas où $n = 2$.

Dans cet exemple, un roi s'adresse à ses deux sages a et b (logiciens, bons raisonneurs, ...) après avoir déposé une tâche blanche sur leur front respectif, en leur disant qu'au moins un des deux a une tâche blanche sur le front. Chaque sage peut voir le front de l'autre mais ne peut voir le sien. Ainsi, chacun sait si l'autre a une tâche blanche ou non mais n'a aucune idée s'il a une tâche ou non. Dans la littérature, on trouve plusieurs versions de ce problème. Dans notre cas, nous supposons le scénario suivant : Le roi pose à l'agent b, la question "As-tu une tâche blanche sur le front ?" et ce dernier répond qu'il n'a aucune idée. Par la suite, le roi pose la même question à l'agent a et celui-ci répond affirmativement. Nous allons formaliser ce problème en utilisant les règles d'inférence sur les configurations définis précédemment et montrer que le sage a saura bel et bien qu'il a une tâche blanche sur le front.

Au début, nous devons définir un langage formel permettant de définir les conditions du problèmes. Ce langage contient :

— une proposition A : l'agent a a une tâche blanche,

— une proposition B : l'agent b a une tâche blanche,

— deux opérateurs de connaissance K_a et K_b. Si p est une formule alors $K_i p$ ($i \in \{a, b\}$) dénote que l'agent i connaît p.

Formalisons maintenant le problème, en prenant en compte le raisonnement de l'agent a sur les connaissances de l'agent b.

AX1 : $K_a K_b (A \vee B)$

Le roi a annoncé qu'au moins un des deux sages a une tâche blanche sur le front, i.e. $A \vee B$. Par conséquent, on a $K_a(A \vee B)$ et $K_b(A \vee B)$. Comme l'annonce était faite publiquement, chaque agent sait que l'autre sait aussi, $K_a K_b (A \vee B)$ et $K_b K_a (A \vee B)$.

AX2 : $K_a(K_b A \vee K_b \neg A)$

En regardant le front de l'agent a, l'agent b sait si l'agent a a une tâche blanche ou non $K_b A \vee K_b \neg A$. Évidemment, l'agent a sait ce fait.

AX3 : $K_a \neg K_b B$

L'agent b a répondu négativement à la question du roi et l'agent a le sait.

À ces axiomes il faut ajouter la règle de sincérité de l'agent b suivante :

$$\frac{\mathcal{C}\langle w : K_b \phi \rangle}{\mathcal{C}'\langle w : \phi \rangle} \; R_{sinc}$$

En effet, on suppose que les deux agents sont sincères et ont même faculté de raisonnement.

Essayons maintenant d'appliquer les règles d'inférence et utiliser les axiomes ainsi que la règle de sincérité ci-dessus afin de construire une configuration nous permettant de conclure que l'agent a peut déduire qu'il a une tâche blanche. Pour faciliter la lecture de cette démonstration, seules les unités déclaratives auxquelles nous nous intéressons sont montrées à chaque étape.

1. Commençons par l'axiome 2.

$$\mathcal{C}\langle w_0 : K_a(K_bA \vee K_b\neg A)\rangle$$

2. Supposons que $R_a(w_0, w_1)$. On obtient alors

$$\mathcal{C}\langle w_0 : K_a(K_bA \vee K_b\neg A), [R_a(w_0, w_1)]\rangle$$

3. Posons $w_1 = \mathit{box}^a(w_0)$. D'après $\mathcal{I}_{KE}$ et 2, on obtient

$$\mathcal{C}_1\langle w_1 : K_bA \vee K_b\neg A\rangle$$

4. En utilisant la logique classique, on peut écrire $K_bA \vee K_b\neg A = (K_bA \wedge \neg K_b\neg A) \vee (\neg K_bA \wedge K_b\neg A) \vee (K_bA \wedge K_b\neg A)$. Il est facile de montrer que :

$$\begin{aligned} K_bA \wedge \neg K_b\neg A &= K_bA \\ \neg K_bA \wedge K_b\neg A &= K_b\neg A \\ K_bA \wedge K_b\neg A &= \bot \end{aligned}$$

Ainsi, on peut "décomposer" w_1 en w_{11} et w_{12} avec $w_{11} = f^a_{K_bA}(w_0)$ et $w_{12} = f^a_{K_b\neg A}(w_0)$. Notons ici qu'on a

$$\begin{aligned} \{f^a_{K_bA}(w_0)\} \cap \{f^a_{K_b\neg A}(w_0)\} &= \emptyset \\ \{f^a_{K_bA}(w_0)\} \cup \{f^a_{K_b\neg A}(w_0)\} &= \{\mathit{box}^a(w_0)\} \end{aligned}$$

5. La configuration est alors

$$\begin{gathered} \mathcal{C}_1\langle w_{11} : K_bA \vee K_b\neg A, [R_a(w_0, w_{11})], \\ w_{12} : K_bA \vee K_b\neg A, [R_a(w_0, w_{12})]\rangle \end{gathered}$$

6. Posons $w_{111} = \mathit{box}^b_A(w_{11})$ et $w_{121} = \mathit{box}^b_{\neg A}(w_{12})$ et appliquons $\mathcal{I}_{KE}$ à 5, on aura alors

$$\mathcal{C}_1\langle w_{111} : A, w_{121} : \neg A\rangle.$$

À cette étape, on obtient la configuration $\mathcal{C}_1$ représentée par la figure 2.

7. Introduisons AX1.

$$\mathcal{C}_2 = \mathcal{C}_1 + [w_0 : K_aK_b(A \vee B)]$$

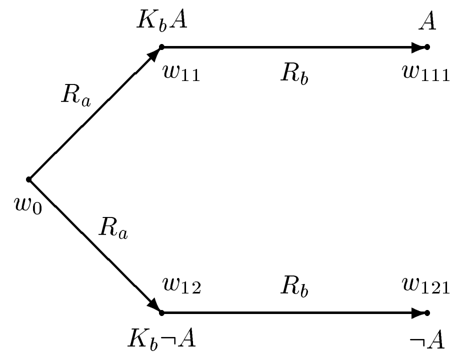

Figure 2. *Configuration* $\mathcal{C}_1$

8. D’après $\mathcal{I}_{KE}$, on obtient

$$\mathcal{C}_2\langle box^a(w_0) : K_b(A \vee B)\rangle.$$

Étant donné que $box^a(w_0)$ est un monde arbitraire accessible à w_0, on obtient alors la configuration suivante : $\mathcal{C}_2\langle w_{11} : K_b(A \vee B), w_{12} : K_b(A \vee B)\rangle$

9. D’après $\mathcal{I}_{KE}$, on obtient $\mathcal{C}_2\langle w_{111} : A \vee B, w_{121} : A \vee B\rangle$.

10. En appliquant $\mathcal{I}_{\wedge I}$ à 6 et à 8, puis $\mathcal{I}_{\wedge E}$ au résultat obtenu, on obtient alors

$$\mathcal{C}_2\langle w_{121} : B\rangle$$

11. D’après $\mathcal{I}_{KI}$, on a $\mathcal{C}_2\langle w_{12} : K_bB\rangle$. La configuration actuelle est représenté par la figure 3.

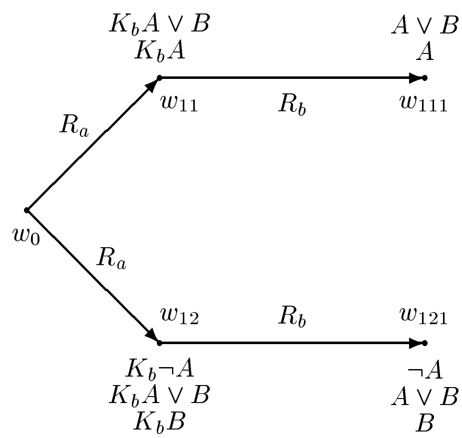

Figure 3. *Configuration* $\mathcal{C}_2$

12. Introduisons maintenant l’axiome AX3.

$$\mathcal{C}_3 \; = \; \mathcal{C}_2 + [w_0 : K_a\neg K_bB]$$

13. D'après $\mathcal{I}_{KE}$, on a $\mathcal{C}_3\langle w_{11} : \neg K_b B, w_{12} : \neg K_b B\rangle$

14. Appliquons $\mathcal{I}_{\wedge I}$ à 11 et 14,

$$\mathcal{C}_3\langle w_{12} : \perp\rangle$$

15. Nous obtenons une contradiction, alors ce qu'on a supposé au départ n'est pas vrai (point 5). Appliquons alors $\mathcal{I}_{R-I}$, on obtient

$$\mathcal{C}_3\langle \neg R_a(w_0, w_{12})\rangle$$

16. En appliquant $\mathcal{I}_{KI}$, on arrive à la configuration $\mathcal{C}_3\langle w_0 : K_a K_b A\rangle$ représentée par la figure 4.

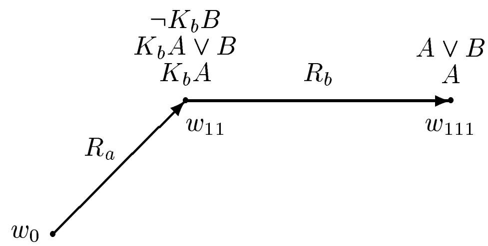

Figure 4. *Configuration* $\mathcal{C}'$

17. Finalement, en appliquant la règle de sincérité R_{sinc}, nous arrivons à la configuration finale $\mathcal{C}'\langle w_0 : K_a A\rangle$.

Cette démonstration est légèrement différente de celle présentée dans [Jar98]. En effet, nous avons ajouté une règle de sincérité qui nous semble indispensable. Sans elle l'agent a ne peut raisonner sur les connaissances et les déductions de l'agent b.

Un autre point intéressant qu'il convient de souligner dans cette démonstration est que les configurations nous ont permis de voir l'évolution des connaissances et des déductions de l'agent a.

5. Conclusion et travaux futurs

Nous avons présenté un système pour la modélisation des connaissances et des croyances d'un agent à l'aide des systèmes labélisés du type LDS. Les variantes de ce système logique sont équivalentes à des systèmes de logique épistémique de type S_4, S_5, etc.

Nous avons ensuite étendu ce système logique pour représenter et modéliser le raisonnement dans un environnement multiagent. Vu que les configurations reflètent bien l'évolution du système, nous avons pu appliquer ce modèle au problème des n sages, d'une façon simple et claire.

Nos perspectives sont maintenant de nous intéresser à la modélisation des agents à capacités limités. Parallèlement, nous désirons introduire le temps, et/ou les actions afin de considérer l'aspect dynamique de l'agent.

Références

[Fit83] M. Fitting. Proof Methods for Modal and Intuitionistic Logics. D. Reidel Publishing Company: Dordrecht, The Netherlands, 1983. (Synthese library volume 169).

[Gab94] D. M. Gabbay. LDS - labelled Deductive Systems -, Foundations. Technical Report MPI-I-94/223, Max-Plank-Institut Für Informatik, 1994.

[Hua94] Z. Huang. Logics for Agents with Bounded Rationality. PhD thesis, ILLC-Univ. v. Amsterdam, NL, 1994.

[Jar97] I. Jarras. Vers la modélisation des interactions entre agents considérés comme des systèmes logiques : Un approche basée sur les labels. Technical report, Université Laval, Départment d'informatique, Québec, 1997.

[Jar98] I. Jarras. Raisonnement d'agents considérés comme des systèmes logiques : approche basée sur les labels. In P. Marquis, M. Zacklad, and J. Lang, editors, RJCIA'98, pages 117–126, 1998.

[Lal90] R. Lalement. Logique réduction résolution. Masson, 1990.

[RM90] S. Reeves and C. Michael. Logic for computer science. Addison-Wesley: Reading, MA, 1990.

[Rus95] A. Russo. Modal labelled deductive systems. Technical Report 95/7, Imperial College of Science, Technology and Medecine, Department of Computing, London, 1995.

[Rus96] A. Russo. Modal Logics as Labelled Deductive Systems. PhD thesis, Imperial College of Science, Technology and Medicine, University of London, 1996.

[SC94] Y. Shoham and S. B. Cousins. Logics of mental attitudes in AI. In Lakemeyer G. and Nebel B., editors, Foundations of Knowledge Representation and Reasoning. Springer-Verlag: Heidelberg, Germany, 1994.

[Woo92] M. Wooldridge. The Logical Modelling of Computational Multi-Agent Systems. PhD thesis, Department of Computation, UMIST, Manchester, UK, October 1992. (Also available as Technical Report MMU–DOC–94–01, Department of Computing, Manchester Metropolitan University, Chester St., Manchester, UK).