Logique de la mise à jour des croyances objectives

Philippe Balbiani[1,2] et Pablo Seban[1,3]

[1] Université de Toulouse
CNRS, Institut de recherche en informatique de Toulouse
118 route de Narbonne, 31062 Toulouse CEDEX 9 (France)

[2] balbiani@irit.fr

[3] seban@irit.fr

Résumé : Le système modal que nous développons consiste à donner une logique pour la mise à jour des croyances objectives. Son langage est construit de la même manière que celui de la logique classique propositionnelle en ajoutant les opérateurs $\square_i$ ("il est cru par l'agent i que"), $[\phi, G]$ ("après que les agents du groupe G aient appris que ϕ") et $[G]$ ("après que les agents du groupe G aient appris quoi que ce soit"). Nous définissons les notions de modèle, de formule valide et de formule satisfaite. L'axiomatisation complète de notre système modal est établie. L'étude de la décidabilité et de la complexité de notre système modal est proposée.

Mots-clés : Logique épistémique, croyances objectives, mise à jour des croyances.

1 Introduction

Deux joueurs, Alex et Béa, sont en face à face au poker *Texas Hold'em*. Chacun connaît ses deux cartes et ne sait pas encore quelles seront les cartes communes posées face visible sur la table. Cette donnée est pourtant fixée à l'avance par le mélange du paquet : en effet nos deux joueurs ne peuvent pas en modifier la distribution. Ils l'apprennent par contre au fur et à mesure de leurs enchères, en mettant ainsi à jour leurs croyances sur ces faits objectifs. Ils peuvent se tromper sur les croyances des autres agents ("Alex croit que Béa croit avoir le meilleur jeu") mais leurs croyances concernant telle ou telle carte posée face visible sur la table sont des croyances vraies. Quel modèle nous permet d'exprimer ce caractère objectif de leur croyance, qui fait que s'ils croient un fait objectif c'est que celui-ci est vrai ?

Rentrons dans le cours du jeu. Initialement, chaque joueur a 2 cartes face cachées. Après une phase d'enchères, le donneur pose trois cartes face visible sur la table (le *flop*). S'ensuivent trois phases d'enchères, entrecoupées par la découverte d'une nouvelle carte (le *turn* et la *rivière*). Même s'il pense avoir un jeu perdant, Alex peut tenter un bluff s'il pense que Béa n'est pas sûre d'avoir le jeu gagnant. Comment représenter alors le fait que Béa est sûre d'avoir le jeu gagnant quoi qu'il arrive ultérieurement ?

Supposons maintenant qu'Alex apprenne le jeu de Béa à son insu. Il a alors un avantage sur son adversaire. Il sait en particulier qui a le jeu gagnant, et il peut également déduire si son adversaire est sûre ou non que son jeu sera gagnant. Quelle opération de mise à jour sur notre modèle permet de différencier la découverte commune d'une carte d'une tricherie (dans laquelle seul un joueur apprend quelque chose) ?

Cet article propose un formalisme répondant à ces exigences. Nous présenterons d'abord la logique des croyances objectives et une opération de mise à jour exprimable dans cette logique, ce qui nous permettra d'exprimer des phrases telles que "Après la distribution du flop, Béa envisage d'avoir un jeu perdant" ou "Alex croit que si l'As de pique sort à la rivière, il aura un jeu gagnant". Dans un second temps, nous présenterons le concept de "mise à jour arbitraire" qui nous permettra d'exprimer par exemple "Béa peut apprendre quelque chose de sorte qu'elle sache que son jeu est gagnant et pourtant quoi qu'apprenne Alex, il envisage que Béa envisage d'avoir un jeu perdant".

2 Croyances objectives

2.1 Syntaxe et sémantique

Nous construisons le langage de la logique des croyances objectives à partir d'un ensemble dénombrable AG d'agents et d'un ensemble dénombrable AT d'atomes. L'ensemble des formules est obtenu à partir de AG et AT en itérant un nombre indéfini de fois les opérations suivantes :

- pour tout $p \in AT$, p est une formule,
- $\perp$ ("falsum") est une formule,
- si ϕ est une formule alors $\neg\phi$ ("non ϕ") est une formule,
- si ϕ est une formule et ψ est une formule alors $(\phi \vee \psi)$ ("ϕ ou ψ") est une formule,
- si ϕ est une formule alors pour tout agent $i \in AG$, $\Box_i\phi$ ("il est cru par l'agent i que ϕ") est une formule.

Nous utilisons les abréviation habituelles pour simplifier les écritures. En particulier, $\Diamond_i\phi$ ("il est admis par l'agent i que ϕ") remplace $\neg\Box_i\neg\phi$. La longueur d'une formule ϕ, notation $\mid \phi \mid$, est le nombre d'occurrences de symboles dans ϕ. La formule ϕ est booléenne lorsque pour tout agent $i \in AG$, ϕ ne contient aucune occurrence de l'opérateur $\Box_i$. Il s'agit maintenant d'interpréter sémantiquement les objets syntaxiques que nous avons introduits. Un modèle est une structure relationnelle de la forme $\mathcal{M} = (W, R, V)$ dans laquelle W est un ensemble non vide de mondes possibles, R est une fonction associant à chaque agent $i \in AG$ une relation binaire R_i sur W et V est une fonction associant à chaque monde possible $x \in W$ un ensemble $V(x)$ d'atomes. Relativement à un modèle $\mathcal{M} = (W, R, V)$, nous définissons la relation de satisfaisabilité entre un monde possible $x \in W$ et une formule ϕ, notation $(\mathcal{M}, x) \models \phi$, de la façon suivante :

- $(\mathcal{M}, x) \models p$ ssi $p \in V(x)$,
- $(\mathcal{M}, x) \not\models \perp$,
- $(\mathcal{M}, x) \models \neg\phi$ ssi non $(\mathcal{M}, x) \models \phi$,
- $(\mathcal{M}, x) \models \phi \vee \psi$ ssi $(\mathcal{M}, x) \models \phi$ ou $(\mathcal{M}, x) \models \psi$,
- $(\mathcal{M}, x) \models \Box_i\phi$ ssi pour tout $y \in W$, xR_iy seulement si $(\mathcal{M}, y) \models \phi$.

Une formule ϕ est valide dans un modèle $\mathcal{M} = (W, R, V)$, notation $\mathcal{M} \models \phi$, lorsque pour tout $x \in W$, $(\mathcal{M}, x) \models \phi$. Une formule ϕ est satisfaite dans un modèle $\mathcal{M} = (W, R, V)$, notation $\mathcal{M}\ sat\ \phi$, lorsqu'il existe $x \in W$ tel que $(\mathcal{M}, x) \models \phi$. Soit $\mathcal{M} = (W, R, V)$, $\mathcal{M}' = (W', R', V')$ des modèles et $x_0 \in W$, $x'_0 \in W'$ des mondes possibles. Nous dirons que $(\mathcal{M}, x_0)$ et $(\mathcal{M}', x'_0)$ sont bisimilaires, notation $(\mathcal{M}, x_0) \leftrightarrows (\mathcal{M}', x'_0)$, lorsqu'il existe une relation binaire Z entre W et W' telle que $x_0 Z x'_0$ et pour chaque $x \in W$ et pour chaque $x' \in W'$, si xZx' alors $V(x) = V'(x')$ et pour tout $i \in AG$,

– s'il existe $y \in W$ tel que xR_iy alors il existe $y' \in W'$ tel que $x'R'_iy'$ et yZy',
– s'il existe $y' \in W'$ tel que $x'R'_iy'$ alors il existe $y \in W$ tel que xR_iy et yZy'.

La proposition suivante nous explique à quel point des modèles bisimilaires se ressemblent.

Proposition 1
Soit $\mathcal{M} = (W, R, V)$, $\mathcal{M}' = (W', R', V')$ *des modèles et* $x_0 \in W$, $x'_0 \in W'$ *des mondes possibles. Si* $(\mathcal{M}, x_0) \leftrightarrows (\mathcal{M}', x'_0)$ *alors* $(\mathcal{M}, x_0) \models \phi$ *ssi* $(\mathcal{M}', x'_0) \models \phi$ *pour chaque formule* ϕ.

2.2 Axiomatisation/complétude

L'intuition que nous avons des énoncés "il est cru par l'agent i que ϕ" et "il est admis par l'agent i que ϕ" est telle que nous nous intéressons à la logique LCO des croyances objectives dont les axiomes sont les suivants :

(LCP) les axiomes de la logique classique propositionnelle,

(K) $\Box_i(\phi \to \psi) \to (\Box_i\phi \to \Box_i\psi)$,

(4) $\Box_i\phi \to \Box_i\Box_i\phi$,

(5) $\Diamond_i\phi \to \Box_i\Diamond_i\phi$,

(T_{bool}) $\Box_i\phi \to \phi$ (ϕ formule booléenne).

Les axiomes (4) et (5) expriment le caractère introspectif de la croyance : "lorsque l'agent i croit que ϕ, il croit qu'il croit que ϕ" et "lorsque l'agent i admet que ϕ, il croit qu'il admet que ϕ" pour chaque formule ϕ. Les axiomes (T_{bool}) expriment le caractère objectif de la croyance : "lorsque l'agent i croit que ϕ, ϕ" pour chaque formule booléenne ϕ. Autrement dit, nous considérons que les croyances des agents sur le monde réel sont consistantes avec celui-ci. Les théorèmes de LCO sont toutes les formules déductibles des axiomes en utilisant les règles de déduction suivantes :

(MP) si ϕ est un théorème et $\phi \to \psi$ est un théorème alors ψ est un théorème,

(GD) si ϕ est un théorème alors $\Box_i\phi$ est un théorème.

A la suite de Hommersom *et al.* (11), considérons la classe $\mathcal{C}_0$ des modèles $\mathcal{M} = (W, R, V)$ dans lesquels

– R_i est transitive pour tout agent $i \in AG$: pour tout $x \in W$, pour tout $y \in W$ et pour tout $z \in W$, si xR_iz et zR_iy alors xR_iy,
– R_i est euclidienne pour tout agent $i \in AG$: pour tout $x \in W$, pour tout $y \in W$ et pour tout $z \in W$, si zR_ix et zR_iy alors xR_iy,

– R_i est o-sérielle pour tout agent $i \in AG$: pour tout $x \in W$, il existe $y \in W$ tel que xR_iy et $V(x) = V(y)$.

La proposition suivante établit l'adéquation et la complétude de la logique LCO pour la classe $\mathcal{C}_0$.

Proposition 2
Soit ϕ une formule. Alors ϕ est un théorème de LCO ssi ϕ est valide dans tout modèle de la classe $\mathcal{C}_0$.

Notons que le résultat de la proposition 2 n'apparaît pas dans (11). Considérons la formule $\Diamond_i\top$. Cette formule est valide dans toute structure relationnelle $\mathcal{M} = (W, R, V)$ où R_i est sérielle pour tout agent $i \in AG$: pour tout $x \in W$, il existe $y \in W$ tel que xR_iy. Par conséquent, la logique LCO est une extension de la logique $KD45_{AG}$ obtenue en remplaçant l'axiome (T_{bool}) par l'axiome $\Diamond_i\top$. L'inclusion de $KD45_{AG}$ dans LCO est stricte si $AT \neq \emptyset$. Pour le montrer, associons à toute partie X de AT, la logique LCO_X obtenue en remplaçant l'axiome (T_{bool}) par l'axiome $\Box_i\phi \rightarrow \phi$ pour chaque formule booléenne ϕ basée sur X. Le lecteur montrera sans peine que la fonction $X \longmapsto LCO_X$ est strictement croissante. Qui plus est, $LCO = LCO_{AT}$ et $KD45_{AG} = LCO_\emptyset$.

2.3 Décidabilité/complexité

De la proposition suivante, il ressort que la logique LCO possède la propriété dite du modèle fini.

Proposition 3
Soit ϕ une formule. Si ϕ est satisfaite dans un modèle de la classe $\mathcal{C}_0$ alors ϕ est satisfaite dans un modèle fini de la classe $\mathcal{C}_0$.

Décider si ϕ est valide dans tout modèle de la classe $\mathcal{C}_0$ revient à décider si $\neg\phi$ est satisfaite dans un modèle de la classe $\mathcal{C}_0$. Et d'après la proposition précédente, elle est satisfaite dans un modèle de la classe $\mathcal{C}_0$ ssi elle est satisfaite dans un modèle fini de la classe $\mathcal{C}_0$. Par conséquent, le problème de décision suivant est décidable :

– entrée : une formule ϕ,
– sortie : ϕ est-elle satisfaite dans un modèle de la classe $\mathcal{C}_0$?

Plus précisément,

Proposition 4
Le problème de décision ci-dessus est $PSPACE$-complet.

Tester si une formule est satisfaite dans un monde possible relativement à un modèle fini peut certainement être fait en un temps fini. Par conséquent, le problème de décision suivant est décidable :

– entrée : une formule ϕ et un modèle fini $\mathcal{M} = (W, R, V)$ de la classe $\mathcal{C}_0$,
– sortie : ϕ est-elle satisfaite dans un monde possible $x \in W$ relativement à $\mathcal{M}$?

Plus précisément,

Proposition 5
Le problème de décision ci-dessus est P-complet.

Notons que les résultats des propositions 3, 4 et 5 n'apparaissent pas dans (11).

3 Mise à jour

Nous nous posons maintenant la question de la mise à jour des croyances objectives : que se passe-t-il dans un modèle lorsqu'un groupe d'agents apprend qu'une formule booléenne est satisfaite ? Pour répondre à cette question, nous présentons le modèle proposé par Hommersom *et al.* (11), en proposant quelques propriétés nouvelles. Soit $\mathcal{M} = (W, R, V)$ un modèle, ϕ une formule booléenne et $G \subseteq AG$ un groupe d'agents. La mise à jour de $\mathcal{M}$ par ϕ et G est le modèle $MAJ_{\phi,G}(\mathcal{M}) = (W', R', V')$ défini de la façon suivante :

- $W' = W \times \{0, 1\}$,
- $(x, a)R'_i(y, b)$ ssi une des conditions suivantes est vérifiée :
  - $a = 0$, $b = 0$ et xR_iy,
  - $a = 1$, $b = 1$, xR_iy, $(\mathcal{M}, y) \models \phi$ et $i \in G$,
  - $a = 1$, $b = 0$, xR_iy et $i \notin G$.
- $V'(x, a) = V(x)$.

Dans tout cet article, nous considérons seulement des groupes finis d'agents. Pour justifier l'opération MAJ qui associe les modèles à leurs mise à jour par des formules booléennes et des groupes d'agents, examinons les propositions suivantes.

Proposition 6
Soit $\mathcal{M} = (W, R, V)$ un modèle de la classe $\mathcal{C}_0$, ϕ une formule booléenne et $G \subseteq AG$ un groupe d'agents. Pour chaque $x \in W$, si $(\mathcal{M}, x) \models \phi$ alors le sous-modèle de $MAJ_{\phi,G}(\mathcal{M})$ engendré à partir de $(x, 1)$ est un modèle de la classe $\mathcal{C}_0$.

Par conséquent, la classe $\mathcal{C}_0$ est stable par l'opération MAJ.

Proposition 7
Soit $\mathcal{M} = (W, R, V)$ un modèle de la classe $\mathcal{C}_0$, ϕ une formule booléenne et $G \subseteq AG$ un groupe d'agents. Pour chaque $x \in W$ et pour tout $i \in AG$,

- $(MAJ_{\phi,G}(\mathcal{M}), (x, 1)) \models \Box_i\phi$ *lorsque* $i \in G$,
- $(MAJ_{\phi,G}(\mathcal{M}), (x, 1)) \models \Box_i\phi$ *ssi* $(\mathcal{M}, x) \models \Box_i\phi$ *lorsque* $i \notin G$.

Il s'ensuit que les agents du groupe G croient la formule apparaissant dans la mise à jour, et que les autres ne la croient que s'ils la croyaient déjà avant la mise à jour.

Proposition 8
Soit $\mathcal{M} = (W, R, V)$ un modèle de la classe $\mathcal{C}_0$, $x \in W$ un monde possible et $G \subseteq AG$ un groupe d'agents. Alors $(MAJ_{\top,G}(\mathcal{M}), (x, 1)) \iff (\mathcal{M}, x)$.

Proposition 9
Soit $\mathcal{M} = (W, R, V)$ un modèle de la classe $\mathcal{C}_0$, $x \in W$ un monde possible et ϕ une formule booléenne. Alors $(MAJ_{\phi,\emptyset}(\mathcal{M}), (x, 1)) \iff (\mathcal{M}, x)$.

Autrement dit, mettre à jour par la constante booléenne $\top$ ou pour un groupe d'agents vide ne change rien aux croyances des agents.

Proposition 10
Soit $\mathcal{M} = (W, R, V)$ *un modèle de la classe* $\mathcal{C}_0$*,* $x \in W$ *un monde possible,* ϕ, ψ *des formules booléennes et* $G \subseteq AG$ *un groupe d'agents. Si* $(\mathcal{M}, x) \models \phi$ *et* $(\mathcal{M}, x) \models \psi$ *alors* $(MAJ_{\psi,G}(MAJ_{\phi,G}(\mathcal{M})), ((x, 1), 1)) \iff (MAJ_{\phi\wedge\psi,G}(\mathcal{M}), (x, 1))$.

Deux mises à jour successives sont donc équivalentes à une seule.

Proposition 11
Soit $\mathcal{M} = (W, R, V)$ *un modèle de la classe* $\mathcal{C}_0$*,* $x \in W$ *un monde possible,* ϕ, ψ *des formules booléennes et* $G, H \subseteq AG$ *des groupes d'agents. Si* $(\mathcal{M}, x) \models \phi$ *et* $(\mathcal{M}, x) \models \psi$ *alors* $(MAJ_{\psi,H}(MAJ_{\phi,G}(\mathcal{M})), ((x, 1), 1)) \iff (MAJ_{\phi,G}(MAJ_{\psi,H}(\mathcal{M})), ((x, 1), 1))$.

L'ordre dans une suite de mises à jour est donc sans importance.

Proposition 12
Soit $\mathcal{M} = (W, R, V)$*,* $\mathcal{M}' = (W', R', V')$ *des modèles de la classe* $\mathcal{C}_0$*,* $x \in W, x' \in W'$ *des mondes possibles,* ϕ *une formule booléenne et* $G \subseteq AG$ *un groupe d'agents. Si* $(\mathcal{M}, x) \models \phi$*,* $(\mathcal{M}', x') \models \phi$ *et* $(\mathcal{M}, x) \iff (\mathcal{M}', x')$ *alors* $(MAJ_{\phi,G}(\mathcal{M}), (x, 1)) \iff (MAJ_{\phi,G}(\mathcal{M}'), (x', 1))$.

Deux modèles bisimilaires le restent donc après une même mise à jour. Notons que les résultats des propositions 8, 9, 10 et 12 n'apparaissent pas dans (11).

4 Mise à jour des croyances objectives

4.1 Syntaxe et sémantique

La logique des croyances objectives utilise uniquement les opérateurs $\Box_i$. Par conséquent, elle est loin d'être satisfaisante pour modéliser la mise à jour. Que pouvons-nous faire pour introduire dans LCO de nouveaux opérateurs donnant la possibilité d'analyser la dynamique de la mise à jour des croyances objectives ? En adaptant les travaux de Hommersom *et al.* (11), nous développons une logique de la mise à jour des croyances objectives dont le langage contient, en plus des opérateurs $\Box_i$, des opérateurs de la forme $[\phi, G]$ pour chaque formule booléenne ϕ et pour chaque groupe d'agents $G \subseteq AG$. Nous complétons ensuite cette présentation par des résultats nouveaux d'expressivité et de décidébilité/compexité. $[\phi, G]\psi$ ayant "après que les agents du groupe G aient appris que ϕ, ψ" pour signification, ces opérateurs introduisent une idée de mise à jour. Nous généralisons la relation de satisfaisabilité de la façon suivante :

– $(\mathcal{M}, x) \models [\phi, G]\psi$ ssi $(\mathcal{M}, x) \models \phi$ seulement si $(MAJ_{\phi,G}(\mathcal{M}), (x, 1)) \models \psi$.

La proposition suivante nous explique à quel point des modèles bisimilaires se ressemblent.

Proposition 13
Soit $\mathcal{M} = (W, R, V)$, $\mathcal{M}' = (W', R', V')$ des modèles et $x_0 \in W$, $x_0' \in W'$ des mondes possibles. Si $(\mathcal{M}, x_0) \iff (\mathcal{M}', x_0')$ alors $(\mathcal{M}, x_0) \models \phi$ ssi $(\mathcal{M}', x_0') \models \phi$ pour chaque formule ϕ.

4.2 Axiomatisation/complétude

Soit $LMAJCO$ la logique de la mise à jour des croyances objectives dont les axiomes sont les suivants :

(LCO) les axiomes de la logique des croyances objectives,

$(R0)$ $[\phi, G](\psi \rightarrow \chi) \rightarrow ([\phi, G]\psi \rightarrow [\phi, G]\chi)$,

$(R1)$ $[\phi, G]p \leftrightarrow (\phi \rightarrow p)$,

$(R2)$ $[\phi, G]\bot \leftrightarrow (\phi \rightarrow \bot)$,

$(R3)$ $[\phi, G]\neg\psi \leftrightarrow (\phi \rightarrow \neg[\phi, G]\psi)$,

$(R4)$ $[\phi, G](\psi \vee \chi) \leftrightarrow [\phi, G]\psi \vee [\phi, G]\chi$,

$(R5)$ $[\phi, G]\Box_i\psi \leftrightarrow (\phi \rightarrow \Box_i[\phi, G]\psi)$ lorsque $i \in G$,

$(R6)$ $[\phi, G]\Box_i\psi \leftrightarrow (\phi \rightarrow \Box_i\psi)$ lorsque $i \notin G$.

Les axiomes $(R1)$, $(R2)$, $(R3)$, $(R4)$, $(R5)$ et $(R6)$ ont tous une interprétation simple. Nous le verrons plus bas, leur présence implique que le langage de $LMAJCO$ n'est pas plus expressif que le langage de LCO. Les théorèmes de $LMAJCO$ sont toutes les formules déductibles des axiomes en utilisant les règles de déduction (MP) et (GD) ainsi que la règle de déduction suivante :

$(GMAJ)$ si ψ est un théorème alors $[\phi, G]\psi$ est un théorème.

La proposition suivante établit l'adéquation et la complétude de la logique $LMAJCO$ pour la classe $\mathcal{C}_0$.

Proposition 14
Soit ϕ une formule. Alors ϕ est un théorème de $LMAJCO$ ssi ϕ est valide dans tout modèle de la classe $\mathcal{C}_0$.

Soit τ la traduction des formules du langage de $LMAJCO$ en des formules du langage de LCO définie de la façon suivante :
- $\tau([\phi_1, G_1] \ldots [\phi_n, G_n], p) = \phi_1 \wedge \ldots \wedge \phi_n \rightarrow p$,
- $\tau([\phi_1, G_1] \ldots [\phi_n, G_n], \bot) = \phi_1 \wedge \ldots \wedge \phi_n \rightarrow \bot$,
- $\tau([\phi_1, G_1] \ldots [\phi_n, G_n], \neg\phi) = \phi_1 \wedge \ldots \wedge \phi_n \rightarrow \neg\tau([\phi_1, G_1] \ldots [\phi_n, G_n], \phi)$,
- $\tau([\phi_1, G_1] \ldots [\phi_n, G_n], \phi \vee \psi) = \tau([\phi_1, G_1] \ldots [\phi_n, G_n], \phi) \vee \tau([\phi_1, G_1] \ldots [\phi_n, G_n], \psi)$,
- $\tau([\phi_1, G_1] \ldots [\phi_n, G_n], \Box_i\phi) = \phi_1 \wedge \ldots \wedge \phi_n \rightarrow \Box_i\tau(\mu, \phi)$ où μ est la séquence obtenue de $[\phi_1, G_1] \ldots [\phi_n, G_n]$ en éliminant les $[\phi, G]$ qui ne concernent pas i,
- $\tau([\phi_1, G_1] \ldots [\phi_n, G_n], [\phi, G]\psi) = \tau([\phi_1, G_1] \ldots [\phi_n, G_n][\phi, G], \psi)$.

Nous laissons au lecteur le soin de montrer que
- $\mid \tau([\phi_1, G_1] \ldots [\phi_n, G_n], \psi) \mid \leq (\mid \phi_1 \mid + \ldots + \mid \phi_n \mid + \mid \psi \mid) \times \mid \psi \mid$,

- $\tau([\phi_1, G_1] \dots [\phi_n, G_n], \psi) \leftrightarrow [\phi_1, G_1] \dots [\phi_n, G_n]\psi$ est valide dans tout modèle de la classe $\mathcal{C}_0$.

Par conséquent,
- $\mid \tau(\epsilon, \psi) \mid \leq \mid \psi \mid^2$,
- $\tau(\epsilon, \psi) \leftrightarrow \psi$ est valide dans tout modèle de la classe $\mathcal{C}_0$.

Le langage de $LMAJCO$ est-il plus expressif que le langage de LCO ? La traduction ci-dessus nous autorise à dire que non : toute formule de $LMAJCO$ est équivalente à une formule de LCO.

4.3 Décidabilité/complexité

De la proposition suivante, il ressort que la logique $LMAJCO$ possède la propriété dite du modèle fini.

Proposition 15
Soit ϕ une formule. Si ϕ est satisfaite dans un modèle de la classe $\mathcal{C}_0$ alors ϕ est satisfaite dans un modèle fini de la classe $\mathcal{C}_0$.

Par conséquent, le problème de décision suivant est décidable :
- entrée : une formule ϕ,
- sortie : ϕ est-elle satisfaite dans un modèle de la classe $\mathcal{C}_0$?

Plus précisément,

Proposition 16
Le problème de décision ci-dessus est $PSPACE$-complet.

Quant au problème de décision suivant, il est également décidable :
- entrée : une formule ϕ et un modèle fini $\mathcal{M} = (W, R, V)$ de la classe $\mathcal{C}_0$,
- sortie : ϕ est-elle satisfaite dans un monde possible $x \in W$ relativement à $\mathcal{M}$?

Plus précisément,

Proposition 17
Le problème de décision ci-dessus est P-complet.

5 Mise à jour arbitraire des croyances objectives

5.1 Syntaxe et sémantique

La logique de la mise à jour des croyances objectives utilise uniquement les opérateurs $\Box_i$ et $[\phi, G]$. Par conséquent, elle est loin d'être satisfaisante pour modéliser la mise à jour arbitraire. Que pouvons-nous faire pour introduire dans $LMAJCO$ de nouveaux opérateurs donnant la possibilité d'analyser la dynamique de la mise à jour arbitraire des croyances objectives ? En nous inspirant des travaux de Balbiani *et al.* (3), nous développons une logique de la mise à jour arbitraire des croyances objectives dont le langage contient, en plus des opérateurs $\Box_i$ et $[\phi, G]$, des opérateurs de la forme $[G]$ pour chaque groupe d'agents $G \subseteq AG$. $[G]\psi$ ayant "après que les agents du groupe G

aient appris quoi que ce soit, ψ" pour signification, ces opérateurs introduisent une idée de mise à jour arbitraire. Nous généralisons la relation de satisfaisabilité de la façon suivante :

– $(\mathcal{M}, x) \models [G]\psi$ ssi pour toute formule booléenne ϕ, $(\mathcal{M}, x) \models [\phi, G]\psi$.

Proposition 18
Les formules suivantes sont valides dans tout modèle de la classe $\mathcal{C}_0$:

(T) $[G]\psi \rightarrow \psi$,

(4) $[G]\psi \rightarrow [G][G]\psi$,

(CR) $\langle G\rangle[H]\psi \rightarrow [H]\langle G\rangle\psi$.

Remarquons d'abord que les formules de la forme $[G \cup H]\psi \rightarrow [G][H]\psi$ ne sont pas toutes valides dans tout modèle de la classe $\mathcal{C}_0$. Par exemple, la formule $[\{i, j\}](\Box_i p \rightarrow \Box_i\Box_j p) \rightarrow [\{i\}][\{j\}](\Box_i p \rightarrow \Box_i\Box_j p)$ n'est pas valide dans le modèle présenté sur la figure 1 (qui est de la classe $\mathcal{C}_0$). Remarquons par ailleurs que les formules de la forme $[G]\langle H\rangle\psi \rightarrow \langle H\rangle[G]\psi$ ne sont pas toutes valides dans tout modèle de la classe $\mathcal{C}_0$. Par exemple, la formule $[\{i\}]\langle\{j\}\rangle(\Box_i p \oplus \Box_j p) \rightarrow \langle\{j\}\rangle[\{i\}](\Box_i p \oplus \Box_j p)$ — dans laquelle l'opérateur $\oplus$ dénote la disjonction exclusive — n'est pas valide dans ce même modèle.

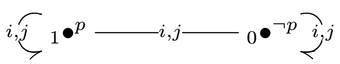

FIG. 1 –

La proposition suivante nous explique à quel point des modèles bisimilaires se ressemblent.

Proposition 19
Soit $\mathcal{M} = (W, R, V)$, $\mathcal{M}' = (W', R', V')$ des modèles et $x_0 \in W$, $x_0' \in W'$ des mondes possibles. Si $(\mathcal{M}, x_0) \Longleftrightarrow (\mathcal{M}', x_0')$ alors $(\mathcal{M}, x_0) \models \phi$ ssi $(\mathcal{M}', x_0') \models \phi$ pour chaque formule ϕ.

5.2 Axiomatisation/complétude

Soit $LMAJACO$ la logique de la mise à jour arbitraire des croyances objectives dont les axiomes sont les suivants :

$(LMAJCO)$ les axiomes de la logique de la mise à jour des croyances objectives,

$(S0)$ $[G](\psi \rightarrow \chi) \rightarrow ([G]\psi \rightarrow [G]\chi)$,

$(S1)$ $[G]\psi \rightarrow [\phi, G]\psi$.

L'axiome $(S1)$ a une interprétation simple. Toutefois, sa seule présence ne suffit pas à rendre $LMAJACO$ complet pour $\mathcal{C}_0$. Pour assurer la complétude de la logique $LMAJACO$ pour la classe $\mathcal{C}_0$, il nous a été nécessaire d'utiliser les règles de déduction suivantes :

$(GMAJA)$ si ψ est un théorème alors $[G]\psi$ est un théorème,

(X) si pour toute formule booléenne ϕ, $\varphi([\phi, G]\psi)$ est un théorème alors $\varphi([G]\psi)$ est un théorème.

Dans la règle de déduction (X), φ représente une forme admissible, i.e. un élément de l'ensemble obtenu à partir d'un symbole dénoté $\sharp$ en itérant un nombre indéfini de fois les opérations suivantes :
– $\sharp$ est une forme admissible,
– si φ est une forme admissible alors pour toute formule ϕ, $(\varphi \vee \phi)$ est une forme admissible,
– si φ est une forme admissible alors pour tout agent $i \in AG$, $\Box_i\varphi$ est une forme admissible.

Remarquons que $\sharp$ apparaît exactement une fois dans toute forme admissible. Pour toute forme admissible φ et pour toute formule ϕ, $\varphi(\phi)$ dénote la formule obtenue de φ en remplaçant l'unique occurrence de $\sharp$ dans φ par ϕ. La proposition suivante établit l'adéquation et la complétude de la logique $LMAJACO$ pour la classe $\mathcal{C}_0$.

Proposition 20
Soit ϕ une formule. Alors ϕ est un théorème de $LMAJACO$ ssi ϕ est valide dans tout modèle de la classe $\mathcal{C}_0$.

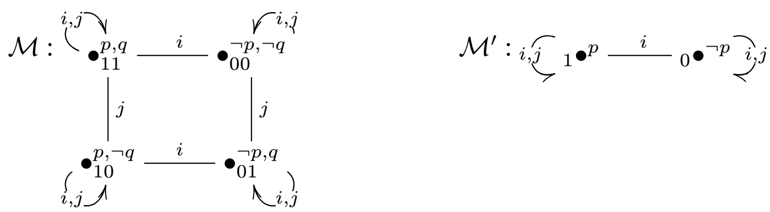

FIG. 2 –

Le langage de $LMAJACO$ est-il plus expressif que le langage de $LMAJCO$? Pour répondre à cette question, il suffit de considérer la formule $\langle\{i,j\}\rangle(\Box_i p \wedge \neg\Box_j\Box_i p)$ et les modèles $\mathcal{M}$ et $\mathcal{M}'$ (de la classe $\mathcal{C}_0$) présentés sur la figure 2. Nous laissons au lecteur le soin de montrer que $(\mathcal{M}, 11) \models \langle\{i,j\}\rangle(\Box_i p \wedge \neg\Box_j\Box_i p)$ et $(\mathcal{M}', 1) \not\models \langle\{i,j\}\rangle(\Box_i p \wedge \neg\Box_j\Box_i p)$. Si la formule $\langle\{i,j\}\rangle(\Box_i p \wedge \neg\Box_j\Box_i p)$ était équivalente à une formule de $LMAJCO$, elle serait satisfaite de la même manière dans les modèles bisimilaires relativement au langage restreint à p. Puisque $(\mathcal{M}, 11) \Longleftrightarrow (\mathcal{M}', 1)$ relativement au langage restreint à p, alors la formule $\langle\{i,j\}\rangle(\Box_i p \wedge \neg\Box_j\Box_i p)$ n'est équivalente à aucune formule de $LMAJCO$. Par conséquent, le langage de $LMAJACO$ est plus expressif que le langage de $LMAJCO$.

5.3 Décidabilité/complexité

Nous ignorons si la logique $LMAJACO$ possède la propriété dite du modèle fini. Nous ignorons également si le problème de décision suivant est décidable :
– entrée : une formule ϕ,

– sortie : ϕ est-elle satisfaite dans un modèle de la classe $\mathcal{C}_0$?

Par contre, nous savons que le problème de décision suivant l'est :

– entrée : une formule ϕ et un modèle fini $\mathcal{M} = (W, R, V)$ de la classe $\mathcal{C}_0$,

– sortie : ϕ est-elle satisfaite dans un monde possible $x \in W$ relativement à $\mathcal{M}$?

Plus précisément,

Proposition 21
Le problème de décision ci-dessus est $PSPACE$-complet.

6 Conclusion

La logique dynamique épistémique présuppose que le dialogue est un programme cognitif qui met à jour l'état épistémique des agents qui y participent. Son développement récent trouve son origine dans les travaux de Plaza (12) qui définit la première logique des annonces publiques. Dans les travaux de Balbiani *et al.* (3), Baltag et Moss (4) et Gerbrandy (8), des actions plus élaborées que les annonces publiques sont considérées : annonces arbitraires, annonces privées et mise à jour de croyances. D'autres travaux, notamment ceux de Dechesne et Wang (5) et Hommersom *et al.* (11), montrent comment certaines variantes de la logique dynamique épistémique peuvent servir à la modélisation formelle et à la vérification des protocoles cryptographiques. Dans cet article, nous avons enrichi la variante élaborée par Hommersom *et al.* — et qui consiste principalement à ajouter au langage de la logique classique propositionnelle les opérateurs $\Box_i$ ("il est cru par l'agent i que") et $[\phi, G]$ ("après que les agents du groupe G aient appris que ϕ") — de l'opérateur $[G]$ ("après que les agents du groupe G aient appris quoi que ce soit"). De notre système modal, nous avons établi l'axiomatisation complète et proposé l'étude de la décidabilité et de la complexité. Les actions de mise à jour des croyances objectives dont nous venons de parler consistent à faire effectuer par un agent omniscient extérieur à un groupe d'agents des annonces objectives à destination des agents de ce groupe. Un prolongement possible de notre travail concerne, à la manière d'Agotnes et van Ditmarsch (1), l'intégration d'opérateurs de la logique des annonces publiques à notre langage. Que deviennent l'axiomatisation/complétude et la décidabilité/ complexité dans ce contexte ? Un autre prolongement possible de notre travail concerne, à la manière de Dechesne et Wang (5), l'utilisation de notre langage pour la modélisation formelle et la vérification des protocoles cryptographiques. Comment modéliser formellement et vérifier la confidentialité et l'intégrité des communications dans ce contexte ?

Références

[1] Agotnes, T., van Ditmarsch, H. : Coalitions and announcements. In : Autonomous Agents and Multiagent Systems. ACM (2008) 673–680.

[2] Agotnes, T., Balbiani, P., van Ditmarsch, H., Seban, P. : Group announcement logic. En préparation. Voir la version préliminaire sur http ://www.irit.fr/-Publications-

[3] Balbiani, P., Baltag, A., van Ditmarsch, H., Herzig, A., Hoshi, T., de Lima, T. : What can we achieve by arbitrary announcements ? A dynamic take on Fitch's knowability. In : Theoretical Aspects of Rationality and Knowledge. ACM (2007) 42–51.
[4] Baltag, A., Moss, L. : Logics for epistemic programs. Synthese **139** (2004) 165–224.
[5] Dechesne, F., Wang, Y. : Dynamic epistemic verification of security protocols : framework and case study. In : Logic, Rationality and Interaction. College Publications (2007).
[6] Van Ditmarsch, H., van der Hoek, W., Kooi, B. : Dynamic Epistemic Logic. Springer (2007).
[7] French, T., van Ditmarsch, H. : Undecidability for arbitrary public announcement logic. A paraître.
[8] Gerbrandy, J. : The surprise examination in dynamic epistemic logic. Synthese **155** (2007) 21–33.
[9] Grädel, E., Otto, M. : On logics with two variables. Theoretical Computer Science **224** (1999) 73–113.
[10] Halpern, J., Moses, Y. : A guide to completeness and complexity for modal logics of knowledge and belief. Artificial Intelligence **54** (1992) 319–379.
[11] Hommersom, A., Meyer, J.-J., de Vink, E. : Update semantics of security protocols. Synthese **142** (2004) 229–267.
[12] Plaza, J. : Logics of public communications. Synthese **158** (2007) 165–179.

Annexe

Démonstration de la proposition 1 : Par induction sur ϕ.

Démonstration de la proposition 2 : Nous laissons au lecteur le soin de montrer que ϕ est un théorème de LCO seulement si ϕ est valide dans tout modèle de la classe $\mathcal{C}_0$. L'argument habituel basé sur le modèle canonique de LCO permet de montrer que ϕ est un théorème de LCO si ϕ est valide dans tout modèle de la classe $\mathcal{C}_0$.

Démonstration de la proposition 3 : L'argument habituel basé sur la filtration permet de montrer que si ϕ est satisfaite dans un modèle de la classe $\mathcal{C}_0$ alors ϕ est satisfaite dans un modèle fini de la classe $\mathcal{C}_0$.

Démonstration de la proposition 4 : Le problème de décision ci-dessus est $PSPACE$-difficile parce que l'argument développé par Halpern et Moses (10) pour montrer que le problème de la satisfaisabilité des formules de la logique $KD45_{AG}$ est $PSPACE$-difficile s'adapte plutôt facilement. Quant à son appartenance à $PSPACE$, l'argument développé par Halpern et Moses (10) pour montrer que le problème de la satisfaisabilité des formules de la logique $KD45_{AG}$ est dans $PSPACE$ permet de la montrer.

Démonstration de la proposition 5 : Le problème de décision ci-dessus est P-difficile parce que l'argument développé par Grädel and Otto (9) pour montrer que le problème de la satisfaisabilité des formules dans un monde possible relativement à un modèle de la logique K est P-difficile s'adapte plutôt facilement. Quant à son appartenance à P, l'argument développé par Grädel and Otto (9) pour montrer que le problème de la satisfaisabilité des formules dans un monde possible relativement à un modèle de la logique K est dans P permet de la montrer.

Démonstration de la proposition 6 : Simple vérification.
Démonstration de la proposition 7 : Simple vérification.
Démonstration de la proposition 8 : Soit Z la relation binaire entre $W \times \{0, 1\}$ et W définie de la façon suivante :

– $(u, a)Zv$ ssi $u = v$.

Nous laissons au lecteur le soin de montrer que Z est une bisimulation entre le sous-modèle de $MAJ_{\top,G}(\mathcal{M})$ engendré à partir de $(x, 1)$ et le sous-modèle de $\mathcal{M}$ engendré à partir de x.
Démonstration de la proposition 9 : Soit Z la relation binaire entre $W \times \{0, 1\}$ et W définie de la façon suivante :

– $(u, a)Zv$ ssi $u = v$.

Nous laissons au lecteur le soin de montrer que Z est une bisimulation entre le sous-modèle de $MAJ_{\phi,\emptyset}(\mathcal{M})$ engendré à partir de $(x, 1)$ et le sous-modèle de $\mathcal{M}$ engendré à partir de x.
Démonstration de la proposition 10 : Soit Z la relation binaire entre $(W \times \{0, 1\}) \times \{0, 1\}$ et $W \times \{0, 1\}$ définie de la façon suivante :

– $((u, a), b)Z(v, c)$ ssi $u = v$, $a = c$ et $b = c$.

Nous laissons au lecteur le soin de montrer que Z est une bisimulation entre le sous-modèle de $MAJ_{\psi,G}(MAJ_{\phi,G}(\mathcal{M}))$ engendré à partir de $((x, 1), 1)$ et le sous-modèle de $MAJ_{\phi\wedge\psi,G}(\mathcal{M})$ engendré à partir de $(x, 1)$.
Démonstration de la proposition 11 : Soit Z la relation binaire sur $(W \times \{0, 1\}) \times \{0, 1\}$ définie de la façon suivante :

– $((u, a), b)Z((v, c), d)$ ssi $u = v$, $a = d$ et $b = c$.

Nous laissons au lecteur le soin de montrer que Z est une bisimulation entre le sous-modèle de $MAJ_{\psi,H}(MAJ_{\phi,G}(\mathcal{M}))$ engendré à partir de $((x, 1), 1)$ et le sous-modèle de $MAJ_{\phi,G}(MAJ_{\psi,H}(\mathcal{M}))$ engendré à partir de $((x, 1), 1)$.
Démonstration de la proposition 12 : Soit Z une bisimulation entre le sous-modèle de $\mathcal{M}$ engendré à partir de x et le sous-modèle de $\mathcal{M}'$ engendré à partir de x'. Soit Z^{maj} la relation binaire entre $W \times \{0, 1\}$ et $W' \times \{0, 1\}$ définie de la façon suivante :

– $(u, a)Z^{maj}(u', a')$ ssi uZu' et $a = a'$.

Nous laissons au lecteur le soin de montrer que Z^{maj} est une bisimulation entre le sous-modèle de $MAJ_{\phi,G}(\mathcal{M})$ engendré à partir de $(x, 1)$ et le sous-modèle de $MAJ_{\phi}(\mathcal{M}')$ engendré à partir de $(x', 1)$.
Démonstration de la proposition 13 : Par induction sur ϕ.
Démonstration de la proposition 14 : Nous laissons au lecteur le soin de montrer que ϕ est un théorème de $LMAJCO$ seulement si ϕ est valide dans tout modèle de la classe $\mathcal{C}_0$. L'argument habituel basé sur le modèle canonique de $LMAJCO$ permet de montrer que ϕ est un théorème de $LMAJCO$ si ϕ est valide dans tout modèle de la classe $\mathcal{C}_0$.
Démonstration de la proposition 15 : D'après la proposition 3, la traduction τ des formules du langage de $LMAJCO$ en des formules du langage de LCO définie à la section 4.2 nous autorise à dire que le problème de décision ci-dessus est $PSPACE$-complet.
Démonstration de la proposition 16 : D'après la proposition 4, la traduction τ des formules du langage de $LMAJCO$ en des formules du langage de LCO définie à la

section 4.2 nous autorise à dire que le problème de décision ci-dessus est $PSPACE$-complet.
Démonstration de la proposition 17 : D'après la proposition 5, la traduction τ des formules du langage de $LMAJCO$ en des formules du langage de LCO définie à la section 4.2 nous autorise à dire que le problème de décision ci-dessus est P-complet.
Démonstration de la proposition 18 : Soit $\mathcal{M} = (W, R, V)$ un modèle, $x \in W$ un monde possible et ψ une formule.
(T) : Supposons que $(\mathcal{M}, x) \models [G]\psi$ et $(\mathcal{M}, x) \not\models \psi$. Donc, $(\mathcal{M}, x) \models [\top, G]\psi$. Par conséquent, $(MAJ_{\top,G}(\mathcal{M}), (x, 1)) \models \psi$. D'après la proposition 8, le sous-modèle de $MAJ_{\top,G}(\mathcal{M})$ engendré à partir de $(x, 1)$ et le sous-modèle de $\mathcal{M}$ engendré à partir de x sont bisimilaires : une contradiction.
(4) : Supposons que $(\mathcal{M}, x) \models [G]\psi$ et $(\mathcal{M}, x) \not\models [G][G]\psi$. Donc, il existe une formule booléenne ϕ_1 telle que $(\mathcal{M}, x) \not\models [\phi_1, G][G]\psi$. Par conséquent, $(\mathcal{M}, x) \models \phi_1$ et $(MAJ_{\phi_1,G}(\mathcal{M}), (x, 1)) \not\models [G]\psi$. Donc, il existe une formule booléenne ϕ_2 telle que $(MAJ_{\phi_1,G}(\mathcal{M}), (x, 1)) \not\models [\phi_2, G]\psi$. Par conséquent, $(MAJ_{\phi_1,G}(\mathcal{M}), (x, 1)) \models \phi_2$ et $(MAJ_{\phi_2,G}(MAJ_{\phi_1,G}(\mathcal{M})), ((x, 1), 1)) \not\models \psi$. D'après la proposition 10, le sous-modèle de $MAJ_{\phi_1,G}(MAJ_{\phi_2,G}(\mathcal{M}))$ engendré à partir de $((x, 1), 1)$ et le sous-modèle de $MAJ_{\phi_1\wedge\phi_2,G}(\mathcal{M})$ engendré à partir de $(x, 1)$ sont bisimilaires : une contradiction.
(CR) : Supposons que $(\mathcal{M}, x) \models \langle G\rangle[H]\psi$ et $(\mathcal{M}, x) \not\models [H]\langle G\rangle\psi$. Donc, il existe une formule booléenne ϕ_1 telle que $(\mathcal{M}, x) \models \langle\phi_1, G\rangle[H]\psi$ et il existe une formule booléenne ϕ_2 telle que $(\mathcal{M}, x) \not\models [\phi_2, H]\langle G\rangle\psi$. Par conséquent, $(\mathcal{M}, x) \models \phi_1$, $(MAJ_{\phi_1,G}(\mathcal{M}), (x, 1)) \models [H]\psi$, $(\mathcal{M}, x) \models \phi_2$, $(MAJ_{\phi_2,G}(\mathcal{M}), (x, 1)) \not\models \langle G\rangle\psi$. Donc, $(MAJ_{\phi_1,G}(\mathcal{M}), (x, 1)) \models \phi_2$, $(MAJ_{\phi_2,H}(MAJ_{\phi_1,G}(\mathcal{M})), ((x, 1), 1)) \models \psi$, $(MAJ_{\phi_2,H}(\mathcal{M}), (x, 1)) \models \phi_1$ et $(MAJ_{\phi_1,G}(MAJ_{\phi_2,H}(\mathcal{M})), ((x, 1), 1)) \not\models \psi$. D'après la proposition 11, le sous-modèle de $MAJ_{\phi_2,H}(MAJ_{\phi_1,G}(\mathcal{M}))$ engendré à partir de $((x, 1), 1)$ et le sous-modèle de $MAJ_{\phi_1,G}(MAJ_{\phi_2,H}(\mathcal{M}))$ engendré à partir de $((x, 1), 1)$ sont bisimilaires : une contradiction.Soit $\mathcal{M} = (W, R, V)$ un modèle, $x \in W$ un monde possible et ψ une formule.
Démonstration de la proposition 19 : Par induction sur ϕ.
Démonstration de la proposition 20 : Nous laissons au lecteur le soin de montrer que ϕ est un théorème de $LMAJCO$ seulement si ϕ est valide dans tout modèle de la classe $\mathcal{C}_0$. L'argument habituel basé sur le modèle canonique de $LMAJCO$ permet de montrer que ϕ est un théorème de $LMAJCO$ si ϕ est valide dans tout modèle de la classe $\mathcal{C}_0$.
Démonstration de la proposition 21 : Le problème de décision ci-dessus est $PSPACE$-difficile parce que l'argument développé par Agotnes *et al.* (2) pour montrer que le problème de la satisfaisabilité des formules dans un monde possible relativement à un modèle de la logique GAL est $PSPACE$-difficile s'adapte plutôt facilement. Quant à son appartenance à $PSPACE$, l'argument développé par Agotnes *et al.* (2) pour montrer que le problème de la satisfaisabilité des formules dans un monde possible relativement à un modèle de la logique GAL est dans $PSPACE$ permet de la montrer.