Logique & Analyse **133-134** (1991), 133-149

LA CONNAISSANCE COMMUNE : UNE SÉMANTIQUE POUR LA LOGIQUE MODALE.

Luc LISMONT* et Philippe MONGIN**

Résumé

L'article introduit la condition de clôture épistémique (une proposition la satisfait si elle est crue en tout monde où elle est vraie) et s'en sert pour redéfinir sémantiquement la connaissance commune. Il montre qu'un système monotone de logique modale épistémique est adéquat et complet pour cette sémantique.

Abstract

This article introduces the condition of belief closure (a proposition is said to be belief closed if it is believed in every world where it is true) and uses it to define a novel semantics of common knowledge. It is shown that a monotonic system of epistemic modal logic is sound and complete with respect to this semantics.

1. *Introduction*

Suivant la définition ordinaire, dont l'idée revient au philosophe D. Lewis dans *Convention* (1969), une proposition est de connaissance commune si elle est vraie, si chaque individu la connaît, chaque individu sait que chacun la connaît, et ainsi de suite à l'infini. On considère banalement que la connaissance porte sur le vrai, tandis que la croyance est susceptible d'être fausse aussi bien que vraie. Si l'on s'en tenait à cette distinction, il faudrait

* G.R.E.Q.E., Ecoles des Hautes Etudes en Sciences Sociales, Marseille.

** Centre National de la Recherche Scientifique (France) et C.O.R.E., Université Catholique de Louvain.Adresse des auteurs : CORE, Voie du Roman Pays 34, B-1348 Louvain-la-Neuve.

dire en toute rigueur que les travaux des logiciens, des théoriciens des jeux ou des informaticiens qui ont formalisé le concept de Lewis, traitent de croyance commune. Car ces travaux analysent généralement la régression infinie du savoir propre et du savoir mutuel en quelque sorte pour elle-même —indépendamment de la vérité ou de la fausseté objective de la proposition initiale.

Barwise (1989, ch. 9) a distingué plusieurs types d'approches techniques du problème. Le type *itéré* formalise directement le concept naïf, tandis que le type *circulaire* condense en une formule de *point fixe* l'infinité de propositions "*a* croit que *b* croit que...". Les deux approches coexistent fréquemment dans une même discipline, voire dans un seul et même article. Par exemple, les théoriciens des jeux se sont penchés sur la connaissance commune à la suite d'Aumann (1976), qui en a proposé une définition dans les deux registres simultanément. Aumann se donne, très classiquement, deux agents qui sont chacun dotés d'une partition d'information sur l'ensemble Ω des états du monde. La partition de *a* manifeste le fait que cet agent identifie certains états entre eux - d'une manière qui lui est propre. Ses croyances ne peuvent donc porter que sur les événements de Ω construits, par opérations booléennes, sur les éléments de la partition. Prenant appui sur cette modélisation rudimentaire de la croyance, Aumann définit comme étant de connaissance commune *tout événement qui contient l'infimum des deux partitions individuelles* (c'est-à-dire leur plus fin grossissement commun). Cette définition est implicitement du type circulaire. Elle est facile à utiliser, mais sans doute peu intuitive. Elle s'avère équivalente à une définition itérative qu'Aumann donne ensuite, et qui transcrit en langage ensembliste chaque proposition : "*a* croit que *b* croit que...". La définition itérative est transparente, mais incommode. Ce contraste des deux définitions est caractéristique de la plupart des recherches menées sur la connaissance commune. De même, quoiqu'elle soit élémentaire dans le cas d'espèce, l'obtention d'une équivalence est un objectif caractéristique de ces travaux.

La dualité des points de vue circulaire et itéré se dégage moins vigoureusement de la théorie des jeux que de la logique épistémique. La raison en est que, dans cette dernière discipline, la distinction interagit le plus souvent avec celle de la syntaxe et de la sémantique. La réserve qu'introduit "le plus souvent" porte sur les logiques infinitaires. Si l'on s'en tient aux conjonctions finies, comme dans les calculs ordinaires du premier ordre, la syntaxe de l'opérateur de connaissance commune C appartiendra nécessairement au type circulaire. Dans la plupart des travaux récents (comme

Fagin, Halpern et Vardi, 1991), elle consiste en un *axiome de point fixe* (PF) et une *règle d'induction* (RI). La sémantique, en revanche, donne lieu à différentes analyses possibles. Quelle que soit la variante finalement retenue, on demandera qu'elle incorpore (mais peut-être seulement à titre de condition nécessaire) le point de vue itératif. Autrement dit, on demandera que, pour tout modèle m et tout monde possible w, si $\langle m, w\rangle$ valide $C\varphi$ alors $\langle m, w\rangle$ valide $\bigwedge_{a\in A} B_a\varphi$, $\bigwedge_{a\in A} B_a(\bigwedge_{a\in A} B_a\varphi)$ et l'infinité dénombrable de conjonctions finies ainsi construites. On appellera *Exigence Sémantique Minimale* (ESM) cette contrainte métathéorique qui s'impose manifestement au modèle retenu.

Les articles de Fagin, Halpern, Moses et Vardi, adoptent généralement l'ESM comme définition de la relation de validation de $C\varphi$. En d'autres termes, l'ESM figure d'emblée comme condition nécessaire et *suffisante* de la sémantique retenue. Le travail précédemment cité, qu'on abrégera en FHV(1991), introduit une notion de modèle originale, celle de structures de connaissance, qui s'avère étroitement liée à la très classique notion de structure de Kripke. Dans ce contexte, les auteurs établissent des résultats de détermination pour la croyance ou la connaissance, individuelle et commune, en renvoyant aux démonstrations antérieures de Fagin et Vardi (1985) et de Halpern et Moses (1988). Leur système le plus fort est une variante de S5 à plusieurs agents; le plus faible est une variante de K, ce qui apparaît rétrospectivement comme inévitable, compte tenu de l'isomorphisme démontré entre structures de connaissance et structures de Kripke.

Suivant une autre méthode sémantique, on définira la relation de validation à l'aide des interprétations déjà fournies pour les opérateurs individuels B_a, $a \in A$. Voici une application au cas particulier commode où la croyance individuelle et la croyance commune satisfont au moins le système K. On définit alors R_C, la relation de Kripke qui sert de contrepartie à C, comme la clôture transitive de $\bigcup_{a\in A} R_a$ — les R_a désignant les contreparties relationnelles des B_a. On vérifie facilement que cette définition respecte l'ESM. La méthode qu'illustre cet exemple peut s'appliquer à des systèmes plus faibles que K . Lismont (1992, 1993) n'en retient que la monotonie de la croyance individuelle et de la croyance commune. Il introduit alors des systèmes de voisinages N_C et N_a, $a \in A$, pour interpréter C et les B_a, respectivement. Le système N_C est défini en fonction des N_a — cette fois-ci, par une construction appropriée, et non par une stipulation élémentaire comme celle qu'autorisait le cas kripkéen. La sémantique de Lismont satisfait l'ESM. Elle débouche sur des théorèmes de détermination, tous relatifs à des systèmes monotones.

Le présent article s'inscrit dans le courant de recherche logique qui vient d'être analysé. Il lui emprunte la syntaxe finitaire, et donc circulaire, de la connaissance commune. En revanche, il innove dans la sémantique. Celle-ci découlera d'une définition proposée pour la validation de $C\varphi$, que l'on peut paraphraser très informellement ainsi : une proposition est une croyance commune en w si elle est impliquée par une proposition qui, d'une part, a la propriété d'être crue par chacun en tout monde où elle est vraie, d'autre part, est crue par chacun en w. On appellera *clôture épistémique* la propriété requise en premier lieu : c'est elle qui porte l'innovation technique des résultats démontrés ici. Elle a déjà été envisagée —mais informellement— dans des travaux qu'inspirait l'approche d'Aumann (Binmore et Brandenburger, 1990, Bacharach, 1992).

Pour énoncer plus précisément la sémantique de cet article, on peut faire tout d'abord l'hypothèse commode : chaque B_a a pour contrepartie une relation de Kripke R_a. On dira alors qu'un sous-ensemble P de mondes possibles est épistémiquement clos si, pour tout $w \in P$, P contient tous les mondes w' qui sont reliés à w par une quelconque des R_a. Mais la notion de clôture épistémique s'exprime aussi bien dans une sémantique de voisinages. Si, donc, on dispose de systèmes de voisinages individuels N_a, on dira que P est épistémiquement clos lorsque, pour tout $w \in P$, P appartient à chacun des voisinages pris en w, i.e., $P \in N_a(w)$ quel que soit a. La définition retenue pour "$C\varphi$ est vraie en w" se lit synthétiquement : il existe un sous-ensemble P de mondes possibles qui (i) est inclus dans l'ensemble de vérité de φ, (ii) est épistémiquement clos (au sens des R_a ou des N_a respectivement), (iii) en w, satisfait une propriété ensembliste dont l'interprétation est que P soit cru par chacun (au sens des R_a ou des N_a respectivement).

Lismont et Mongin (1993) analysent en détail les possibilités offertes par cette définition de la croyance commune. On se limitera ici aux faits les plus généraux. Après avoir énoncé les notions syntaxiques (section 2), on précisera les définitions sémantiques de ce paragraphe dans le contexte des systèmes de voisinages (section 3), que l'on a préféré à celui des relations de Kripke pour sa plus grande généralité. On démontrera un théorème d'adéquation et de complétude pour un système —épistémiquement faible— de croyance *monotone*, individuelle et commune (section 4). On conclura en énonçant, sans les démontrer, quelques résultats, qui permettraient de faire le lien entre le théorème précédent et un théorème de détermination relatif aux structures de Kripke (section 5).

2. *Définitions syntaxiques*

Les systèmes que l'on examinera procèdent de ceux qu'étudie d'ordinaire la logique modale propositionnelle [Chellas (1980) ou Hughes et Cresswell (1984)]. Le vocabulaire en comporte les éléments suivants : un ensemble *VP* de variables propositionnelles (dont la cardinalité est quelconque); les connecteurs logiques $\wedge$, $\vee$, $\neg$, $\rightarrow$, $\leftrightarrow$; enfin, les opérateurs unaires $(B_a)_{a \in A}$, C et E. L'ensemble des individus A est fini. L'opérateur E, que l'on introduit par commodité, peut se lire "tout le monde croit que". On notera Φ l'ensemble des formules bien formées construites à partir de ce vocabulaire.

Le système minimal de cet article MC_A consiste, d'une part, en une axiomatisation de la logique propositionnelle, que l'on n'explicitera pas, d'autre part, en schémas d'axiomes et de règles spécifiquement modaux :

(RM$_a$) $$\frac{\varphi \rightarrow \psi}{B_a\varphi \rightarrow B_a\psi} \qquad \text{(pour tout } a \in A\text{)}$$

(Déf. E) $$E\varphi \leftrightarrow \bigwedge_{a \in A} B_a\varphi$$

(RM$_C$) $$\frac{\varphi \rightarrow \psi}{C\varphi \rightarrow C\psi}$$

(PF) $$C\varphi \rightarrow E(C\varphi \wedge \varphi)$$

(RI) $$\frac{\varphi \rightarrow E\varphi}{E\varphi \rightarrow C\varphi}$$

La présence de *règles de monotonie* $(\mathrm{RM}_a)_{a \in A}$ signifie que la logique implicitement définie par les opérateurs B_a, C et la logique de MC_A ont beaucoup en commun. Mais le système autorise deux cas-limites : un agent ne croit en aucun énoncé ; il croit en une contradiction. Ces cas-limites disparaîtraient si l'on enrichissait MC_A de ces deux schémas : pour tout a,

(N$_a$) $$B_a \top$$

$(\mathrm{P}_a) \qquad \neg B_a \perp$

Même ainsi renforcé, le système laisserait subsister cette possibilité : un agent croit en φ et en $\neg\varphi$ sans en tirer les conséquences. Pour l'exclure, il faudrait encore ajouter les schémas : pour tout a,

$(\mathrm{C}_a) \qquad B_a\varphi \wedge B_a\psi \rightarrow B_a(\varphi \wedge \psi)$

La réciproque de (C_a) est déjà impliquée par la monotonie (RM_a).

On a longuement débattu, ces dernières années, du problème difficile de l'omniscience logique (par exemple Stalnaker, 1991 ; Dubucs, 1992 ; Gochet, 1992). La présence d'une règle de monotonie suffit à faire tomber le système MC_A sous une objection de ce type. Pour atténuer l'objection plus considérable encore qu'appelle, dans FHV (1991), l'adoption du système K, Vardi (1986) avait proposé de décrire la croyance individuelle par la seule *règle d'équivalence* : pour tout a,

$$(\mathrm{RE}_a) \qquad \frac{\varphi \leftrightarrow \psi}{B_a\varphi \leftrightarrow B_a\psi}$$

En ce qui concerne les opérateurs B_a tout au moins, cet affaiblissement de (RM_a) ramène exactement au système minimal engendrant une sémantique des voisinages. Malheureusement l'axiomatisation de la croyance commune sous (RE_a) pose de nombreux problèmes techniques. Les difficultés que l'on rencontre alors ne se paient certainement pas d'un gain conceptuel décisif. Car (RE_a) préserve des formes gênantes d'omniscience logique ; par exemple, un agent qui connaît au moins une tautologie connaîtrait ipso facto *toutes* les tautologies du système.

Le schéma (PF) indique en substance que la croyance commune implique la croyance partagée, ainsi que la croyance partagée en le fait qu'il y ait croyance commune : cette propriété explique que l'on parle d'*axiome du point fixe*. La règle (RI) indique que, si un énoncé est analytiquement de croyance partagée, il est aussi de croyance commune. L'appellation *règle d'induction* peut se justifier heuristiquement ainsi : en appliquant $(\mathrm{RM}_a)_{a\in A}$ et (Déf. E) on voit facilement que E est monotone :

$$(\mathrm{RM}_E) \qquad \frac{\varphi \rightarrow \psi}{E\varphi \rightarrow E\psi}$$

Une récurrence simple montre alors que :

$$\frac{\varphi \rightarrow \psi}{E^k\varphi \rightarrow E^k\psi} \qquad \text{(pour tout } k > 0\text{)}$$

(le symbole E^k désigne $E\overset{k\,fois}{\ldots}E$). Intuitivement parlant, (RI) permet de résumer cette infinité dénombrable d'inférences en une seule ou, si l'on veut, de "passer à la limite". La présence de (RM_C) vient du fait que, contrairement à E, C n'hérite plus automatiquement de la monotonie des B_a. Il est techniquement très utile, et il paraît conceptuellement défendable, de prêter à la croyance commune la structure minimale que l'on attribue aux opérateurs individuels.

Parmi tous les autres axiomes que considère d'habitude la logique modale épistémique, on mettra à part les deux schémas suivants : pour tout a,

$$(\mathrm{T}_a) \qquad B_a\varphi \rightarrow \varphi$$

$$(\mathrm{D}_a) \qquad B_a\varphi \rightarrow \neg B_a \neg\varphi$$

L'axiome de vérité (T_a), qui vaut aussi pour l'opérateur C à cause de (PF) et (RM_a), fait passer des systèmes de la croyance à ceux de la connaissance, strictement entendue. Le schéma (D_a) est un affaiblissement de (T_a) et est approprié à une logique de la croyance pour des agents cohérents.

Le système formé par (RM_a), (N_a) et (C_a) est le K-*système pour l'agent* a, c'est-à-dire le plus faible qui permette d'interpréter B_a par une relation de Kripke. Dans un K-système pour a, (D_a) est équivalent à (P_a) et donc permet d'exclure un des cas-limites précédemment envisagés : l'agent croit en une contradiction.

On notera $\vdash$ la relation d'inférence de MC_A. On désignera les renforcements du système monotone minimal par MC_A+(N_A), MC_A+(P_A), et à l'avenant.

3. *Structures de voisinages et clôture épistémique.*

L'opérateur C n'a pas de contrepartie ensembliste directe dans la définition que l'on va donner des structures : celles-ci seront du type ordinairement considéré en logique multi-modale. Plus précisément, *une structure de voisinages monotone* est un $(|A|+2)$-uple :

$$m = \langle W, (N_a)_{a \in A}, v \rangle$$

dans lequel

. W est un ensemble non vide, appelé ensemble de mondes possibles ;
. pour chaque a, N_a est une application $W \to P(P(W))$ —$P(.)$ désignant l'ensemble des parties— telle que, pour tout $w \in W$, $N_a(w)$ soit fermé par sur-ensembles, i.e., si $P \in N_a(w)$ et $P \subseteq P' \subseteq W$, alors $P' \in N_a(w)$;
. v est une application $W \times VP \to \{0, 1\}$, appelée valuation.

Pour alléger les définitions qui suivent, on introduira également l'application N_a :

$$\forall w \in W, N_E(w) = \bigcap_{a \in A} N_a(w)$$

La validation sémantique se définit de la manière ordinaire, sauf pour les formules $C\psi$. On aura donc, pour tout $m = \langle W, (N_a)_{a \in A}, v \rangle$:

. si $\varphi \in VP$, $\langle m, w \rangle \models \varphi$ ssi $v(w, \varphi) = 1$;
. si $\varphi = \neg\psi$, $\langle m, w \rangle \models \varphi$ ssi $\langle m, w \rangle \not\models \psi$;
. si $\varphi = \psi_1 \wedge \psi_2$, $\langle m, w \rangle \models \varphi$ ssi $\langle m, w \rangle \models \psi_1$ et $\langle m, w \rangle \models \psi_2$;

et à l'avenant pour les autres clauses propositionnelles;

. si $\varphi = B_a\psi$, $\langle m, w \rangle \models \varphi$ ssi $\|\psi\|^m \in N_a(w)$;
. si $\varphi = E\psi$, $\langle m, w \rangle \models \varphi$ ssi $\|\psi\|^m \in N_E(w)$;

La notation $\|\psi\|^m$ renvoie à l'ensemble de vérité de ψ, c'est-à-dire à

$$\left\{ w' \in W \,\middle|\, \langle m, w' \rangle \models \psi \right\}$$

Le plus souvent, on omettra l'indice suscrit.

La première section, à caractère heuristique, faisait un usage libre du mot "proposition". Dans ce qui suit, on s'en tiendra à l'usage ordinaire en logique, suivant lequel une proposition est l'ensemble de vérité d'une formule. Pour désigner les sous-ensembles de W plus généralement, on parlera d'*événements*.

Les fonctions de voisinage N_a ou N_E sont particulièrement faciles à interpréter lorsqu'on a en vue des applications épistémiques de la logique modale. Elles se voient directement comme des systèmes de croyance pour a et pour le collectif E ("tout le monde"), respectivement. A chaque w elles associent les propositions crues en w. [Ceci ne veut pas dire, naturellement, que $N_a(w)$ et $N_E(E)$ ne contiennent que des propositions : on y trouvera inévitablement aussi des événements non propositionnels.] Les structures de Kripke ne se prêtent pas à une lecture aussi transparente. Pour attribuer un caractère épistémique aux R_a, il faut y voir des relations d'accessibilité *subjectives* ("a regarde w' comme accessible depuis w"), ce qui est obscur. Mongin (1993) recommande en conséquence de privilégier la sémantique des voisinages en logique modale épistémique (y compris lorsque les axiomes incluent un K-système pour chaque agent).

Il reste à introduire la sémantique de C. Soit $P \subseteq W$; on dira que P est *épistémiquement clos* (e.c.) si :

$$\forall w \in P, P \in N_E(w).$$

Si P est une proposition, $P = \|\psi\|^m$ pour un $\psi \in \Phi$, et la clôture épistémique stipule que ψ est crue en tout point où elle est vraie. La définition perd sa lisibilité (si ce n'est analogiquement) lorsque P est un événement non propositionnel. On pose enfin :

. Si $\varphi = C\psi$, $\langle m, w\rangle \models \varphi$ ssi $\exists P \in N_E(w)$ t.q. $P \subseteq \|\psi\|^m$ et P est e.c.

On notera l'ensemble des structures de voisinages monotones par M^N. Comme de coutume, $m \models \varphi$ et $M \models \varphi$ signifient respectivement : [pour tout $w \in W$, $\langle m, w\rangle \models \varphi$] et [pour tout m dans la classe M, $m \models \varphi$].

On terminera cette présentation des concepts sémantiques en vérifiant que la définition proposée pour la validation de $C\psi$ satisfait bien l'ESM.

Proposition 1. Quels que soient $m \in M^N$, $w \in W$, $k > 0$ et $\varphi \in \Phi$,

$$\langle m, w\rangle \models C\varphi \Rightarrow \langle m, w\rangle \models E^k\varphi.$$

On pourra démontrer syntaxiquement ce résultat, une fois qu'on aura établi l'adéquation du système MC_A pour la sémantique proposée (cf. théorème 2) : il suffit alors de dériver $\vdash C\varphi \rightarrow E^k\varphi$, ce qui est immédiat. Mais la démonstration purement sémantique est éclairante. Elle dépend notamment d'un lemme (également utilisé dans la section 4) qui met en évidence le lien entre clôture épistémique et ESM.

Lemme 1. Si $P \subseteq \|\varphi\|$ et si P est e.c., alors $P \subseteq \|C\varphi\|$.

Preuve : Par définition $\|C\varphi\| = \{w \in W \mid \exists P' \in N_E(w), P \subseteq \|\varphi\|$ et P' est e.c.$\}$. Soit $w \in P$. Par la clôture épistémique de P, $P \in N_E(w)$. En considérant l'autre hypothèse sur P, on voit que $w \in \|C\varphi\|$. ■

Preuve de la proposition 1 : Par récurrence sur k. Par définition, si $\langle m, w\rangle \models C\varphi$ alors $\langle m, w\rangle \models E\varphi$. On suppose ensuite que $\|C\varphi\| \subseteq \|E^k\varphi\|$. Soit $w \in W$ tel que $\langle m, w\rangle \models C\varphi$. Il existe $P \in N_E(w)$, $P \subseteq \|\varphi\|$ et P est e.c. Le lemme 1 indique que $P \subseteq \|C\varphi\|$. Donc, par l'hypothèse de récurrence: $P \subseteq \|E^k\varphi\|$ et $\|E^k\varphi\| \in N_E(w)$, ce qui donne $\langle m, w\rangle \models E^{k+1}\varphi$. ■

4. *Détermination de la croyance monotone (individuelle et commune)*

Théorème 2. $\vdash \varphi$ ssi $M^N \models \varphi$.

Preuve de l'adéquation : De manière évidente, (RM_a), (Déf.E) et (RM_C) sont valides. Pour vérifier la validité de (PF), on se donne m dans M_N et w dans W tels que $\langle m, w\rangle \models C\varphi$. Il existe donc $P \in N_E(w)$ tel que $P \subseteq \|\varphi\|$ et P est e.c. On veut montrer que $\|C\varphi \wedge \varphi\|$, qui est égal à $\|C\varphi\| \cap \|\varphi\|$, se trouve dans $N_E(w)$. Ceci découle de $P \subseteq \|C\varphi\|$ (puisque $N_E(w)$ est fermé par sur-ensembles), ce qui résulte de l'hypothèse $P \subseteq \|\varphi\|$ (en vertu du lemme 1). Il reste à vérifier que la règle (RI) est valide. On suppose donc que, pour tout m dans M^N et tout w dans W,

$$w \in \|\varphi\| \Rightarrow \|\varphi\| \in N_E(w)$$

On doit prouver que, quels que soient m et w,

$$\|\varphi\| \in N_E(w) \Rightarrow \exists P \subseteq \|\varphi\| \text{ t.q. } P \in N_E(w) \text{ et } P \text{ est e.c.}$$

Soit donc $\|\varphi\| \in N_E(w)$: on vérifie facilement que $\|\varphi\|$ a toutes les propriétés requises pour P. ■

La démonstration de complétude résultera des lemmes 2 à 5. Pour établir l'implication :

$$M_N \models \psi \Rightarrow \vdash \psi$$

on *fixera* la formule considérée ψ et on adaptera la construction habituelle des ensembles de formules maximaux consistants. Ce type de démonstration —formule par formule— n'est pas rare en logique modale. Boolos (1979) et Cresswell (1983), par exemple, l'ont utilisé pour résoudre d'autres problèmes que ceux créés par l'opérateur C. On ne rappellera pas ici les définitions et propriétés des ensembles maximaux consistants, pour lesquelles le lecteur est renvoyé à Chellas (1980, ch. 2) ou Hughes et Cresswell (1983, ch. 2).

Soit $\varphi \in \Phi$. On définit par induction la profondeur de φ, que l'on notera pf(φ) :

. si $\varphi \in VP$, pf(φ) = 0 ;
. si $\varphi = \neg\psi$, pf(φ) = pf(ψ) ;
. si $\varphi = \psi_1 \wedge \psi_2$, pf($\varphi$) = max(pf($\psi_1$), pf($\psi_2$)), et de manière identique pour les autres connecteurs binaires ;
. si $\varphi = B_a\psi$, pf(φ) = pf(ψ)+1, et de manière identique pour les autres opérateurs épistémiques.

On peut maintenant introduire $\Phi[\psi]$, l'ensemble des formules de Φ qui satisfont les deux conditions suivantes :

(i) elles sont construites à partir de l'ensemble $VP[\psi]$ des variables propositionnelles ayant une occurrence dans ψ ;
(ii) elles ont une profondeur au plus égale à pf(ψ)+1.

L'ensemble $\Phi[\psi]^-$ est défini par la condition (i) et la variante suivante de (ii) :

(ii') elles ont une profondeur au plus égale à pf(ψ).

Soit $C(MC_A)$ la classe des maximaux consistants du système MC_A, et la relation d'équivalence sur $C(MC_A)$:

$$\forall \Gamma, \Delta \in C(MC_A), \Gamma \stackrel{\psi}{\equiv} \Delta \text{ ssi } \Gamma \cap \Phi[\psi] = \Delta \cap \Phi[\psi]$$

Pour chaque classe d'équivalence $[\Gamma]^{\psi}$, on notera Γ^{ψ} son intersection commune. Dans Γ^{ψ}, il y aura des formules de profondeur quelconque. Ce sont toutes celles qui sont déductibles de $\Gamma \cap \Phi[\psi]$. On aura évidemment $\Gamma \stackrel{\psi}{\equiv} \Delta$ ssi $\Gamma^{\psi} = \Delta^{\psi}$. On notera I^{ψ}, l'ensemble des classes d'équivalence. Le lemme suivant montre que chaque partie de I^{ψ} peut être caractérisée par une formule de $\Phi[\psi]$:

Lemme 2. Si $P \subseteq I^{\psi}$, on peut trouver $\varphi_0 \in \Phi[\psi]$ telle que :

$$P = [\varphi_0]^{\psi} \stackrel{\text{déf}}{=} \{[\Gamma]^{\psi} \mid \varphi_0 \in \Gamma\}$$

Schéma de la preuve : Il est clair que $\Phi[\psi]$ n'a qu'un nombre fini de formules à l'équivalence logique près. Par conséquent, quel que soit le maximal consistant Γ, l'ensemble $\Gamma \cap \Phi[\psi]$ a un nombre fini de formules à l'équivalence logique près et $[\Gamma]^{\psi}$ peut donc être caractérisé par une formule $\varphi^{\Gamma} \in \Phi[\psi]$. Du fait de la clôture des maximaux consistants pour l'équivalence logique, il n'y a qu'un nombre fini de classes dans I^{ψ} et, par conséquent, $\varphi_0 = \bigvee_{[\Gamma]^{\psi} \in P} \varphi^{\Gamma}$ est une formule bien définie. Elle est dans $\Phi[\psi]$. On vérifie facilement qu'elle possède la propriété souhaitée. ■

Les ensembles $[\varphi]^{\psi}$ ont une propriété techniquement utile :

Lemme 3. $\forall \varphi_1, \varphi_2 \in \Phi[\psi]$, si $[\varphi_1]^{\psi} \subseteq [\varphi_2]^{\psi}$, alors $\vdash \varphi_1 \rightarrow \varphi_2$.

Schéma de la preuve : On se donne un maximal consistant Γ. En utilisant l'hypothèse et l'appartenance de φ_1 à $\Phi[\psi]$, on démontre que $\varphi_1 \in \Gamma \Rightarrow \varphi_2 \in \Gamma$. ■

Pour chaque formule ψ, on définit un *modèle ψ-canonique* comme un $m \in M^N$ tel que:

. $W = I^{\psi}$;
. quels que soient $a \in A$ et $[\Gamma]^{\psi} \in I^{\psi}$

$$N_a([\Gamma]^{\psi}) = \{P \subseteq I^{\psi} \mid \exists \varphi \in \Phi[\psi] \text{ t.q. } [\varphi]^{\psi} \subseteq P \text{ et } B_a\varphi \in \Gamma^{\psi}\};$$

. quel que soit $[\Gamma]^{\psi} \in I^{\psi}$,

$$N_E([\Gamma]^{\psi}) = \bigcap_{a \in A} N_a([\Gamma]^{\psi}) ;$$

. quels que soient $\varphi \in VP$ et $[\Gamma]^{\psi} \in I^{\psi}$,

$$v([\Gamma^{\psi}], \varphi) = 1 \text{ ssi } \varphi \in \Gamma^{\psi}.$$

Il est évident que les $N_a([\Gamma]^{\psi})$ sont fermés par sur-ensembles.

Les deux lemmes suivants montrent que le rôle joué par les modèles ψ-canoniques dans cette construction est analogue à celui des modèles canoniques usuels en logique modale.

Lemme 4. Quel que soit le maximal consistant Γ :

(i) $\forall\varphi \in \Phi[\psi]^{-}, B_a\varphi \in \Gamma \Rightarrow [\varphi]^{\psi} \in N_a([\Gamma]^{\psi})$;
(ii) $\forall\varphi \in \Phi[\psi], [\varphi]^{\psi} \in N_a([\Gamma]^{\psi}) \Rightarrow B_a\varphi \in \Gamma$.

Preuve : (i) est immédiat par définition de $N_a([\Gamma]^{\psi})$. Pour prouver (ii), supposons que $[\varphi]^{\psi} \in N_a([\Gamma]^{\psi})$, i.e., qu'il existe $B_a\varphi_0 \in \Gamma^{\psi}$ t.q. $[\varphi_0]^{\psi} \subseteq [\varphi]^{\psi}$. Par le lemme 3, $\vdash \varphi_0 \rightarrow \varphi$, et, par (RM_a), $\vdash B_a\varphi_0 \rightarrow B_a\varphi$, d'où $B_a\varphi \in \Gamma$. ■

Lemme 5. Quels que soient le maximal consistant Γ et la formule $\psi' \in \Phi[\psi]^{-}$,

$$\psi' \in \Gamma \text{ ssi } \langle m, [\Gamma]^{\psi}\rangle \models \psi'$$

Schéma de la preuve : Par induction sur la complexité des formules de $\Phi[\psi]^{-}$. L'hypothèse d'induction peut se traduire par l'égalité $[\psi']^{\psi} = \|\psi'\|^{\psi}$ pour toute formule ψ' de complexité inférieure à celle d'une formule donnée. Les seuls cas non triviaux sont $\psi' = B_a\varphi$ (qui résulte du lemme 4) et $\psi' = C\varphi$. Pour ce dernier cas, il s'agit de montrer que :

(*) $C\varphi \in \Gamma$ ssi $\exists P \in N_E([\Gamma]^{\psi})$ t.q. $P \subseteq [\varphi]^{\psi}$ et P est e.c.

En utilisant notamment (PF) on établit le résultat intermédiaire :

(**) $C\varphi \in \Delta^{\psi} \Rightarrow [C\varphi \wedge \varphi]^{\psi} \in N_E([\Delta]^{\psi})$, pour tout maximal consistant Δ.

L'implication de gauche à droite dans (*) se montre alors en posant $P = [C\varphi \wedge \varphi]^{\psi}$. Quant à l'implication de droite à gauche, on introduit l'ensemble

$$Q = \{[\Delta]^{\psi} \in I^{\psi} \mid \varphi \in \Delta \text{ et } \exists P' \text{ t.q. } P' \subseteq [\varphi]^{\psi},\ P' \in N_E([\Delta]^{\psi}) \text{ et } P' \text{ est e.c.}\},$$

qui est caractérisé par une formule $\varphi_0 \in \Phi[\psi]$ d'après le lemme 2. Le reste de la démonstration consiste alors à montrer que $E\varphi_0 \in \Gamma$, puis à vérifier les théorèmes : $\vdash \varphi_0 \rightarrow \varphi$ et $\vdash \varphi_0 \rightarrow E\varphi_0$, qui, en présence de ($RM_C$) et (RI), impliquent la conclusion désirée. ∎

Fin de la preuve de complétude : Le lemme 5 joue le rôle de "lemme canonique" dans la démonstration ordinaire, qu'il suffit alors d'adapter (ch. Chellas, 1980, ch. 2). ∎

5. *Renforcements du système de la croyance monotone. Conclusions.*

On se contentera ici d'énoncer, sans les démontrer, des résultats qui concernent les schémas supplémentaires de la section 2.

Théorème 3. Soit les classes de structures de voisinages monotones qui, pour tout $w \in W$ et tout $a \in A$, vérifient respectivement : (i) $N_a(w) \neq \varnothing$; (ii) $\varnothing \notin N_a(w)$; (iii) $P, P' \in N_a(w) \Rightarrow P \cap P' \in N_a(w)$; (iv) $\forall P \in N_a(w)$, $w \in P$; (v) $P \in N_a(w) \Rightarrow P^c \notin N_a(w)$; (vi) $N_a(w)$ est augmenté $[\cap N_a(w) \in N_a(w)]$. La classe (i) est déterminée par $MC_A+(N_A)$, (ii) par $MC_A+(P_A)$, (iii) par $MC_A+(C_A)$, (iv) par $MC_A+(T_A)$, (v) par $MC_A+(D_A)$, (vi) par KC_A $[= MC_A+(N_A)+(C_A)]$.

Le schéma de la démonstration, qui est laissée au lecteur, consiste, pour chacune des classes, à constater l'adéquation de l'axiome et/ou de la règle correspondante, et à montrer que le modèle ψ-canonique pour le système considéré vérifie bien la propriété sémantique indiquée.

On voit notamment que la propriété de réflexivité (iv) permet de rendre au plan sémantique la différence, tout d'abord syntaxiquement exprimée, entre connaissance et croyance commune. La détermination de la propriété d'augmentation (vi) revêt une importance particulière parce qu'elle livre indirectement un théorème d'adéquation et de complétude *relatif aux structures de Kripke*. Voici quel est le schéma de cette dérivation. On introduit

M^K la classe des structures de Kripke multi-modales, c'est-à-dire des $(|A|+2)$-uples :

$$m = \langle W, (R_a)_{a \in A}, v\rangle$$

tels que $W \neq \varnothing$, R_a est une relation binaire sur W et v est la fonction de valuation. La seule clause originale de validation concernera le cas des $\varphi = C\psi$. Pour analyser ce cas, on commence par redéfinir ainsi la propriété de clôture épistémique : $P \subseteq W$ est e.c. ssi

$$\forall w \in P, \bigcup_{a \in A} R_a[w] \subseteq P,$$

où $R_a[w]$ est $\{w' \in W \mid w'R_a w\}$, i.e., l'ensemble des mondes a-accessibles depuis w.
On pose alors :

$$\langle m, w\rangle \models C\psi \text{ ssi } \exists P \text{ t.q. } \bigcup_{a \in A} R_a[w] \subseteq P,\ P \subseteq \|\psi\|^m \text{ et } P \text{ est e.c.}$$

On peut vérifier que (i) ces deux définitions retranscrivent dans le contexte kripkéen les mêmes intuitions épistémiques, exactement, que les définitions adaptées aux structures de voisinages, (ii) elle reflètent le morphisme bien connu entre les structures de Kripke et les structures de voisinages augmentées (Chellas, 1980, p. 221). Il suffit alors d'exploiter ce morphisme de structures pour conclure que M^K est déterminé par KC_A.

Aucun des faits énumérés dans le théorème 3 ne paraîtra surprenant. Mais c'est peut-être, justement, l'un des mérites de la sémantique présentée dans cet article qu'elle permet de démontrer facilement les théorèmes de détermination qu'on attendait naturellement. L'analyse de la connaissance commune par la clôture épistémique présente un double avantage de simplicité : d'une part elle emploie les structures ordinaires de la logique multi-modale; d'autre part, elle facilite le rapprochement de la syntaxe et de la sémantique, en faisant figurer un élément de circularité dans la définition même de celle-ci. Du point de vue épistémique, le résultat principal de l'article constitue un progrès par rapport à celui de FHV (1991) : il montre en effet qu'on peut axiomatiser la croyance commune à l'aide de (PF) et (RI) sans imposer la totalité du système K à la croyance individuelle. Lismont (1992,1993) avait déjà atteint une conclusion semblable par d'autres moyens sémantiques. La question demeure, toutefois, de savoir comment axiomatiser la connaissance commune sous l'hypothèse épistémique la pl faible qu'autorise la logique modale traditionnelle, c'est-à-dire sous la règle d'équivalence (RE_a) plutôt que sous la règle de monotonie (RM_a).

Université Catholique de Louvain

BIBLIOGRAPHIE

R. Aumann. Agreeing to Disagree. *The Annals of Statistics*, 4:1236-1239, 1976.

M. Bacharach. The Acquisition of Common Knowledge. Ch. 17 in *Knowledge, Belief and Strategic Interaction* (s. la dir. de C. Bicchieri et M.L. Dalla Chiara). Cambridge University Press, Cambridge, 1992.

J. Barwise, *The Situation in Logic*. Center for the Study of Language and Information, Stanford, 1989.

K. Binmore et A. Brandenburger. Common Knowledge and Game Theory. Ch. 4 in K. Binmore, *Essays on the Foundations of Game Theory*. Blackwell, Oxford, 1990.

G. Boolos. *The Unprovability of Consistency*. Cambridge University Press, Cambridge, 1979.

B.F. Chellas. *Modal Logic, an Introduction*. Cambridge University Press, Cambridge, 1980.

M.J. Cresswell. The Completeness of KW and KW1.1. *Logique et Analyse n° 102*, 123-127, 1983.

J. Dubucs. *Omniscience logique et frictions cognitives*. in *Epistémologie et cognition* (s. la dir. de D. Andler et alii). Mardaga, Liège.

R. Fagin, J.Y. Halpern et M.Y. Vardi. A Model-Theoretic Analysis of Knowledge. *Journal of the Association for Computing Machinery*, 38(2):382-428, 1991.

R. Fagin et M.Y. Vardi. An Internal Semantics for Modal Logic. *Proceedings of the 17th A.C.M. Symposium on Theory of Computing*. Providence, 1985.

P. Gochet. Problèmes de logique de la connaissance. *Cahier du département de philosophie n° 92, hors série 07*, Université de Montréal, 1992.

J.Y. Halpern et Y.O. Moses. A Guide to the Modal Logic of Knowledge. *Proceedings of the 9th International Joint Conference on Artificial Intelligence* (IJCAI-85), New York, 1985.

G.E. Hughes et M.J. Cresswell. *A Companion to Modal Logic*. Methuen, London, 1984.

D.K. Lewis. *Convention: A Philosophical Study*. Harvard University Press, 1969.

L. Lismont. *La connaissance commune : approches modales, ensembliste et probabiliste*. Dissertation doctorale, Université Catholique de Louvain, 1992.

L. Lismont. La connaissance commune en logique modale. *Mathematical Logic Quarterly*, 39:115-130, 1993.

L. Lismont et P. Mongin. Belief Closure: A Semantics of Common Knowledge. *Core Discussion Paper* 9339, Université Catholique de Louvain, 1993.

P. Mongin. Some Connections Between Epistemic Logic and the Theory of Nonadditive Probability. In *Patrick Suppes, Scientific Philosopher* (s. la dir. de P. Humphreys). Kluwer, Dordrecht, 1993.

R. Stalnaker. The Problem of Logical Omniscience I. *Synthese*, 89:425-440, 1991.

M.Y. Vardi. On Epistemic Logic and Logical Omniscience. In *Proceedings of the First Conference on Theoretical Aspects of Reasoning about Knowledge* (s. la dir. de J.Y. Halpern), p. 293-305. Morgan Kaufmann Publishes, Los Altos, 1986.