La logique de la connaissance

Éric GILLET

1. Introduction

La connaissance est depuis l'aube de la philosophie un sujet d'étude privilégié. Aujourd'hui cette étude reste centrale dans bien des domaines qui vont de la philosophie (métaphysique, philosophie des sciences, épistémologie, etc.) aux diverses branches de la psychologie. Parallèlement à ces domaines d'investigation, l'avènement de l'informatique a donné des raisons supplémentaires pour réfléchir à la connaissance. La constitution de bases de données contenant plusieurs millions d'informations élémentaires pose le problème de la récupération par les utilisateurs des informations stockées; et à côté du développement de langages d'interrogation qui visent à standardiser les procédures d'accès aux données, apparaît le concept de base de connaissances, où c'est le programme lui-même qui organise le stockage des données et qui se charge de retrouver les informations répondant aux questions de l'utilisateur. La base de connaissances se comporte non plus comme un simple archiviste qui sait où l'information est rangée, mais comme un expert qui connaît son domaine et peut évaluer la pertinence du croisement de certaines informations. La réalisation de tels programmes dépend fortement de l'analyse du fonctionnement de la connaissance et de sa représentation. Par ailleurs, le développement technique et la montée en puissance des ordinateurs amène les ingénieurs à développer des architectures parallèles où différents processeurs se partagent l'exécution de programmes. La synchronisation des tâches nécessite que chaque processeur puisse *connaître* l'état des autres processeurs et leur état d'avancement dans la réalisation de leur

✉ Université de Liège; Séminaire de Logique et d'Épistémologie; 32, Place du 20-Août; B-4000 Liège (Belgique).
Fax : + 32 41 66 57 00

MOTS-CLÉS : Logique épistémique, connaissance de groupe, connaissance faible.

programme. Ici encore la connaissance est un concept central, non plus en tant qu'information à représenter, mais en tant que processus déductif.

Plus généralement, les sciences de l'ordinateur, via les projets de l'intelligence artificielle et du connexionisme, sont confrontées à la nécessité d'implémenter des fonctions qui relèvent de près ou de loin de la connaissance; les fonctions de bas niveau, comme la reconnaissance des formes ou le traitement de signal, sont confrontées à des ambiguïtés que seul le recours à des représentations plus élaborées permet de résoudre, et inversement, des problèmes de haut niveau, comme la prise de décision ou la compréhension des langues naturelles, ne peuvent se résoudre que grâce à de multiples connaissances élémentaires, dites de sens commun. Cette interaction incessante entre les divers niveaux fonctionnels d'un problème rend cruciale l'étude de la connaissance tant du point de vue de la hiérarchisation et de la représentation de l'information que du raisonnement et de la communication. L'approche spécifique de la connaissance par les sciences de l'ordinateur et l'élaboration progressive et pluridisciplinaire des sciences cognitives témoignent de l'importance et de l'actualité de la question de la connaissance.

À partir des années soixante, la théorie de la connaissance s'est trouvée enrichie par l'apport de la logique. Il ne s'agit plus ici de répartir les divers champs d'application de la connaissance, ni d'en découvrir les conditions de possibilité, les fondements, ou encore de construire des applications « intelligentes », mais bien d'étudier la structure déductive propre aux propositions qui contiennent des verbes de type épistémique tels « *savoir* », « *connaître* », « *croire* », « *vérifier* », etc. Avec sa méthodologie propre, la logique aborde la connaissance par la formalisation d'un langage qui comporte des opérateurs de connaissance, et se propose d'en élucider le fonctionnement et d'en débusquer les présupposés.

Dans la suite de ce texte, je me propose de présenter brièvement la sémantique standard que la logique épistémique donne pour la formalisation des opérateurs de connaissance, et d'aborder la problématique particulière de la connaissance dans un groupe d'agents où la connaissance peut être distribuée selon divers schémas, allant de la connaissance commune à la connaissance répartie. Je proposerai pour terminer une analyse de la connaissance faible ou indirecte qui permet de représenter plus finement la connaissance qu'un agent peut avoir à propos de ce que les autres agents connaissent.

2. La formalisation de la connaissance

C'est en 1951 dans *An Essay on Modal Logic* que H. von Wright, le premier, formalise les attitudes propositionnelles de type épistémique comme des *opérateurs modaux*. Il ne s'agit, alors, que d'une utilisation d'une logique modale à la C.I. Lewis où les attitudes propositionnelles sont abordées de façon impersonnelle. En 1962, l'ouvrage de J. Hintikka, *Knowledge and Belief*, présente une analyse des attitudes épistémiques ou doxastiques parallèle à celle de la nécessité ontique ou aléthique. Hintikka tient compte des agents porteurs de l'attitude propositionnelle et leur associe un ensemble des mondes qui représente ce qui est compatible avec ce qu'ils savent. Le sens de l'expression « A sait que p » ($K^a p$) se ramène au sens que prend la proposition p dans les mondes compatibles avec les connaissances de l'agent A. Cette analyse va permettre d'utiliser les outils de la logique des mondes possibles et de les adapter aux besoins de la logique épistémique. La sémantique présentée ici est la sémantique standard des mondes possibles utilisant les structures de Kripke apparues en 1963.

Un modèle $\mathcal{M}$ pour la logique de la connaissance multi-agents est un quadruplet $\langle \mathcal{W}, \mathcal{D}, \mathcal{R}, \mathcal{V} \rangle$: $\mathcal{W}$ est l'ensemble des mondes épistémiques $\{w_0, w_1, w_2, \ldots\}$; $\mathcal{D}$ le domaine des valeurs prises par les propositions, à savoir l'ensemble des valeurs de vérité $\{1, 0\}$; $\mathcal{R}$ est un n-uple $\langle \mathcal{R}^1, \ldots, \mathcal{R}^n \rangle$ où $\mathcal{R}^i$ est la relation d'accessibilité entre mondes associée à l'agent i ; $\mathcal{V}$ est une fonction d'évaluation qui associe une valeur à chaque proposition dans chaque monde épistémique.

La relation sémantique $\models$ qui relie un modèle $\mathcal{M}$ et un monde w aux propositions α qu'ils rendent vraies est définie inductivement. Les atomes propositionnels sont vrais dans un monde quand le modèle les évalue tels :

$$\mathcal{M}, w \models p \Leftrightarrow \mathcal{V}(w, p) = 1,$$

ce qui se lit « le modèle $\mathcal{M}$ vérifie la proposition p dans le monde w si et seulement si la fonction d'évaluation $\mathcal{V}$ associe la valeur 1 à la proposition p dans le monde w ».

Les connecteurs vérifonctionnels sont définis à partir de la négation et de la conjonction; la négation d'une proposition est vraie dans un monde quand le modèle ne la vérifie pas :

$$\mathcal{M}, w \models \neg\alpha \Leftrightarrow \mathcal{M}, w \not\models \alpha;$$

et la conjonction de deux propositions est vraie quand le modèle vérifie les deux propositions :

$$\mathcal{M}, w \models (\alpha \land \beta) \Leftrightarrow \mathcal{M}, w \models \alpha \quad \text{et} \quad \mathcal{M}, w \models \beta.$$

Mais la définition la plus importante concerne les propositions qui affirment la connaissance d'un agent ($K^a p$). Une telle proposition est vérifiée par le modèle dans un monde quand elle est vérifiée dans tous les mondes compatibles avec ce que l'agent sait, et c'est ici qu'interviennent les relations d'accessibilité $\mathcal{R}^i$ entre les mondes. Pour chaque agent, on dispose d'une relation d'accessibilité qui associe à chaque monde les mondes qui lui sont épistémiquement compatibles; dès lors, la définition de la relation $\models$ pour les propositions de la forme $K^i\alpha$ s'énonce comme suit :

$$\mathcal{M}, w \models K^i\alpha \Leftrightarrow \text{pour tout} \quad w' \quad \text{tel que} \quad w\mathcal{R}^i w' : \mathcal{M}, w' \models \alpha,$$

ce qui se lit « le modèle $\mathcal{M}$ vérifie la proposition $K^i\alpha$ dans le monde w si et seulement si le modèle $\mathcal{M}$ vérifie la proposition α dans tous les mondes w' épistémiquement compatibles avec le monde w pour l'agent I ».

En fonction des propriétés de la relation d'accessibilité, les mondes w' vont varier, ce qui va modifier les propriétés de l'opérateur K^i : pour formaliser la connaissance, il suffit que $\mathcal{R}^i$ soit une relation réflexive, ce qui implique que quand on vérifie la connaissance d'un agent dans le monde réel, ce monde réel fait toujours partie des mondes compatibles où l'on vérifie la proposition connue, ce qui suffit à assurer une propriété essentielle de la connaissance, à savoir que ce qui est connu est *a fortiori* vrai. Des contraintes supplémentaires sur la relation $\mathcal{R}^i$, comme la transitivité ou la symétrie, peuvent garantir qu'un agent accède à l'introspection positive : il sait qu'il sait, ou à l'introspection négative : il sait qu'il ne sait pas.

Ces différentes versions sémantiques de la connaissance correspondent à des version axiomatiques qui comportent les axiomes et règles suivants :

— la « *nécessitation épistémique* » est la seule règle spécifique à la connaissance; elle est issue de la logique modale aléthique (d'où son nom) et affirme la rationalité de la connaissance :

$$\vdash \alpha \Rightarrow \vdash K^i\alpha,$$

la connaissance des théorèmes ne dépend pas d'informations factuelles[1];

— « $K^i\alpha \supset \alpha$ » est l'axiome propre de la connaissance[2], il affirme que ce qui est connu est vrai;

— « $K^i(\alpha \supset \beta) \supset (K^i\alpha \supset K^i\beta)$ » est l'axiome de normalité qui permet aux agents de raisonner et de tirer de nouvelles connaissances à partir des anciennes;

[1] L'expression $\vdash \alpha$ se lit « la proposition α est un théorème ». Cette règle est une idéalisation de la structure déductive de la connaissance, elle conduit à une rationnalité excessive connue sous le nom d'omniscience logique. Le lecteur intéressé en trouvera un traitement dans [Gillet 1992].

[2] Le symbole $\supset$ désigne l'implication matérielle, l'expression $\alpha \supset \beta$ se lit « si α alors β ».

— l'introspection positive et négative que l'on introduit facultativement dans le système sont rendues par les axiomes

$$«K^i\alpha \supset K^iK^i\alpha»$$

et $$«\neg K^i\alpha \supset K^i\neg K^i\alpha».$$

Cet outillage formel suffit à rendre compte de la structure déductive de la connaissance d'un agent; le résultat du croisement de la connaissance de plusieurs agents dans un groupe repose sur cette structure élémentaire et demande l'introduction de quelques concepts supplémentaires.

3. La connaissance dans un groupe

Imaginons un groupe constitué de deux agents A et B ; on veut pouvoir formaliser la connaissance de chacun d'eux et son évolution via l'acquisition et l'échange d'informations. Supposons qu'au départ chacun des agents ne connaît qu'un seul fait : l'agent A connaît «p» et l'agent B, «q». L'agent A imagine donc que dans toutes les variantes du monde compatibles avec ce qu'il sait, «p» est vrai et «q» tantôt vrai, tantôt faux; l'agent B, lui, imagine que «q» est toujours vrai et que «p» est tantôt vrai, tantôt faux. Un modèle possible de cette situation contient trois mondes et deux relations d'accessibilité réflexives : le monde réel w_0 (accessible par A et B) rend vrais «p» et «q»; le monde w_1 est accessible par A et rend «p» vrai et «q» faux; le monde w_2 est accessible par B et rend «q» vrai et «p» faux.

Dans cette situation, le modèle satisfait les propositions «p», «q», «K^ap» et «K^bq», ainsi que les propositions «$\neg K^aq$» et «$\neg K^bp$». Qu'en est-il des propositions concernant la connaissance d'un agent à propos des connaissances de l'autre agent ? Pour y répondre, il faut reprendre la sémantique et décomposer les conditions de vérité de ces énoncés. La proposition «K^aK^bq» est vraie dans w_0 si et seulement si pour tout w' tel que $w_0\mathcal{R}^aw'$: $\mathcal{M}, w' \models K^bq$, et «$K^bq$» est vrai dans chacun des w' si et seulement si pour tout w'' tel que $w'\mathcal{R}^bw''$: $\mathcal{M}, w'' \models q$: ce qui revient à vérifier «q» dans tous les mondes w'' tel que $w_0\mathcal{R}^aw'$ et $w'\mathcal{R}^bw''$. Les mondes w' sont les mondes accessibles par A à partir de w_0, à savoir, w_0 et w_1 ; les mondes accessibles par B à partir de w_0 sont w_0 et w_2 et à partir de w_1, B n'accède qu'à w_1. Or ce monde falsifie «q» et donc «K^aK^bq» est rendu faux par le modèle dans le monde réel w_0. Symétriquement, «K^bK^ap» est rendu faux, mais par contre le modèle vérifie les propositions qui affirment que les agents ne savent pas que l'autre sait.

Si l'agent A apprend que B sait «q», il élimine les mondes qui sont devenus incompatibles avec ses nouvelles connaissances : dès lors le monde w_1 cesse d'être compté parmi ses variantes possibles et le modèle vérifie en w_0, la proposition « $K^a K^b q$ ». Par ailleurs, le modèle révisé vérifie également la proposition « $K^a q$ » car l'agent A en apprenant que B connaît « q » a également appris que « q » est vrai : un modèle de la connaissance répercute les informations nouvelles sur l'ensemble de la structure et permet d'en déduire toutes les conséquences.

Il est également possible de parler et de raisonner à propos de la connaissance du groupe dans son ensemble et non plus seulement de la connaissance de chacun de ses membres, ce qui constitue l'objet des deux prochaines sections

3.1. Connaissance de tous et connaissance commune

Imaginons la situation posée par le *puzzle* intitulé « le problème des trois sages » introduit par McCarthy et cité dans [Konolige 1986, 58]. Un roi réunit ses trois conseillers afin de déterminer lequel est le plus sage. Il leur propose de deviner si la tache qu'il leur peint sur le front est blanche ou noire, chacun voit les deux autres et ne peut ni parler ni utiliser de miroir. Le roi leur peint à chacun une tache blanche sur le front, il les réunit, et leur annonce qu'au moins l'un d'entre eux a une tache blanche. Au bout d'un certain temps le premier sage, perplexe, avoue son ignorance. Le second ne tarde pas à déclarer forfait et le troisième, triomphant, annonce que sa tache est blanche. La question est de savoir comment il a procédé.

La question est intrigante, et pourtant la logique épistémique donnée plus haut, bien qu'élémentaire, suffit à la résoudre. J'épargnerai au lecteur la preuve formelle de la solution pour me consacrer à l'explicitation de la démarche suivie par le troisième sage. La troisième sage imagine qu'il a une tache noire et raisonne à la place du deuxième sage : imaginons que moi, le deuxième sage, j'aie également une tache noire, donc le premier sage doit voir deux taches noires et savoir, grâce à la déclaration du roi, qu'il a une tache blanche, mais comme il avoue son ignorance, ce n'est pas le cas et donc je n'ai pas une tache noire, d'où je sais que j'ai une tache blanche. Mais comme le deuxième sage déclare ignorer la couleur de sa tache, le troisième sait que son hypothèse n'est pas la bonne, et donc, qu'il a une tache blanche.

La difficulté de ce casse-tête logique réside dans la double indirection qui consiste à se mettre dans la peau du deuxième sage qui se met dans la peau du premier. Et c'est précisément pour résoudre les difficultés causées par les raisonnements multi-agents que le recours à la logique de la connaissance s'avère indispensable. Mais cette conclusion n'est pas la seule leçon que l'on puisse tirer de ce *puzzle*.

En effet, la déclaration du roi selon laquelle au moins un des sages a une tache blanche semble superflue : avant cette déclaration, chaque sage, qui voit les autres, sait que au moins deux sages (et *a fortiori* un) ont une tache blanche sur le front; et pourtant sans cette déclaration du roi, le troisième sage n'aurait pas pu effectuer sa déduction. Quand le roi énonce ce qui peut sembler trivial, les sages apprennent plus que la proposition « au moins un sage a une tache blanche », les conditions de l'énonciation donnent à chacun l'information selon laquelle « chacun sait que au moins un sage a une tache blanche », plus l'information « chacun sait que chacun sait que au moins un sage a une tache blanche », et ainsi de suite. C'est la double itération de l'opérateur de connaissance qui permet au troisième sage de raisonner indirectement à partir de la connaissance des autres.

Quand on prend en considération un groupe d'agents, on envisage à côté de l'opérateur de connaissance individuelle K^i, les opérateurs qui expriment la connaissance du groupe. Dans l'exemple des trois sages, on peut en isoler deux : le premier, l'opérateur T, exprime la connaissance que tous les agents possèdent (dans l'exemple précédent le fait que deux sages ont une tache blanche), et le second, l'opérateur C, exprime la connaissance commune, celle qui est partagée par tous les membres du groupe et qui affirme de surcroît que cette connaissance même est partagée par tous, et ainsi de suite.

La connaissance de tous se formalise aisément en prenant l'équivalence entre cet opérateur et la conjonction de la connaissance des n agents :

$$T\alpha \underset{\text{déf}}{=} (K^a\alpha \wedge K^b\alpha \wedge \ldots \wedge K^n\alpha)$$

La sémantique de cet opérateur se construit aisément en imposant que la proposition « α » soit vraie dans tous les mondes accessibles, pour chaque agent : il suffit pour cela d'utiliser la relation d'accessibilité $\mathcal{R}^T$ équivalente à l'union des relations d'accessibilité de chaque agent du groupe : $\mathcal{R}^T = \cup_{i=1}^{n}\mathcal{R}^i$.

La formalisation de la connaissance commune offre une difficulté supplémentaire : en effet la connaissance commune est une abréviation de la séquence infinie $T\alpha \wedge TT\alpha \wedge TTT\alpha \wedge \ldots$ Du point de vue sémantique, le problème se résout assez aisément, il suffit de consulter tous les mondes accessibles, par $\mathcal{R}^T$, puis pour chacun d'eux de consulter chacun des mondes accessibles, et ainsi de suite. La construction d'une relation d'accessibilité $\mathcal{R}^C$ pour la connaissance commune est la simple extension de $\mathcal{R}^T$ dont on prend la clôture transitive, c'est-à-dire tous les mondes accessibles par la répétition de la relation $\mathcal{R}^T$. Axiomatiquement, une définition en une ligne n'est pas possible, il faut recourir à un système d'axiomes qui règle le fonctionnement global de ce nouvel opérateur, comme on l'a fait pour l'opérateur de connaissance individuelle[3].

Ces trois opérateurs de connaissance sont reliés entre eux par la déductibilité : la connaissance commune est la plus forte : elle implique la connaissance de tous, qui à son tour implique la connaissance de chaque agent. Mais dans un groupe, il n'y a pas que des opérateurs de connaissance déductivement plus forts que la connaissance individuelle, il est utile de considérer les propriétés logiques de certains aspects de la connaissance partagée par l'ensemble d'un groupe.

3.2. Connaissance de quelqu'un et connaissance répartie

Un réseau d'ordinateurs se partage l'accomplissement d'une vaste tâche qui requiert de nombreuses étapes intermédiaires hiérarchisées : chaque ordinateur ignore la façon dont les tâches sont réparties, cette répartition dépend de la disponibilité de chaque machine et de la résolution des tâches intermédiaires. Pourtant un ordinateur doit de temps à autre connaître l'état de réalisation des tâches subordonnées à la sienne : il va demander si quelqu'un connaît l'information dont il a besoin. Les messages les plus économiques sont ceux qui se limitent aux informations essentielles, et l'identification des agents qui détiennent la connaissance n'est pas toujours pertinente : il sera souvent suffisant de raisonner à partir d'informations minimales comme l'est celle qui affirme qu'un membre du groupe sait qu'une tâche intermédiaire est accomplie. On voit sans peine que la proposition « quelqu'un sait que p », bien que moins forte que la proposition « l'agent A sait que p », est cependant susceptible d'entrer dans un processus inférentiel fécond. Le premier opérateur de connaissance de groupe plus faible que la connaissance individuelle est celui qui formalise la connaissance de quelqu'un notée Q.

Quand on affirme « $Q\alpha$ », on signifie que quelqu'un du groupe (constitué de n agents) connaît la proposition « p », ce qui peut se paraphraser par l'expression « l'agent A connaît p, ou l'agent B connaît p, ..., ou l'agent N connaît p ». Cette paraphrase suffit à donner une définition rigoureuse de la connaissance d'au moins une personne dans un groupe ou connaissance de quelqu'un :

$$Q\alpha \underset{\text{déf}}{=} (K^a\alpha \vee K^b\alpha \vee \ldots \vee K^n\alpha).$$

La sémantique de cet opérateur se construit aisément : la définition de la connaissance de quelqu'un comme étant la disjonction de la connaissance de chaque agent revient à imposer que la proposition « α » soit vraie dans tous les

[3] Le lecteur intéressé peut se référer à [Lejoly 1987] ou à [Halpern & Moses (1986)].

mondes accessibles d'au moins un agent : dès lors, la définition de la relation $\models$ pour les propositions de la forme « $Q\alpha$ » s'énonce comme suit :

$$\mathcal{M}, w \models Q\alpha \iff \text{il y a un agent tel que, pour tout } w' \text{ tel que } w\mathcal{R}^i w' : \mathcal{M}, w' \models \alpha .$$

Il est à remarquer que cette conception de la connaissance n'exploite pas totalement les ressources et les informations disponibles dans le groupe : supposons qu'un agent A sache que ou « p » ou « q » (« $K^a p \vee q$ ») est vrai, et qu'indépendamment, un agent B sache que « p » est faux. (« $K^b \neg p$ »). Le groupe constitué par ces deux agents dispose des informations suffisantes pour conclure par syllogisme disjonctif que si « $p \vee q$ » est vrai et si « p » est faux, alors « q » doit être vrai. Et pourtant, ni l'agent A ni l'agent B ne savent que « q » est vrai, et donc la proposition « Qq » est fausse. Cette connaissance qui exploite toutes les ressources du groupe est appelée connaissance répartie ou connaissance implicite du groupe et notée « I ». Pour qu'elle soit effective, il faut que les membres du groupe échangent l'information dont ils disposent et en déduisent les conséquences.

La connaissance répartie est le premier type de connaissance de groupe qui soit réellement spécifique au groupe, c'est-à-dire qui ne soit pas la connaissance d'un ou de plusieurs agents. Ici, par la mise en commun de toutes ses ressources, le groupe constitue une seule entité qui se comporte vis-à-vis de la connaissance comme un individu. La sémantique de la connaissance répartie doit donc s'exprimer de manière similaire à celle d'un individu. L'observation du rôle que jouent les mondes compatibles révèle que ces mondes représentent des variantes épistémiquement possibles à ce que les agents savent. La mise en commun des ressources de chacun va éliminer parmi les variantes individuellement possibles celles qui sont globalement incompatibles avec la connaissance implicite du groupe, pour ne conserver que les variantes sur lesquelles il y a consensus. Ce qui suffit pour construire un sémantique pour l'opérateur I en prenant comme relation d'accessibilité pour le groupe l'intersection de toutes les relations d'accessibilité des agents : $\mathcal{R}^I = \cap_{i=1}^n \mathcal{R}^i$; ce qui donne la sémantique suivante :

$$\mathcal{M}, w \models I\alpha \Leftrightarrow \text{pour tout } w' \text{ tel que } w\mathcal{R}^I w' : \mathcal{M}, w' \models \alpha .$$

Comme pour la connaissance commune, l'axiomatisation de la connaissance répartie demande plus qu'une simple définition : il faut se redonner l'ensemble des axiomes de la connaissance en les interprétant en terme de connaissance implicite de groupe et ajouter l'axiome « $K^i\alpha \supset I\alpha$ » qui exprime le rapport entre la connaissance individuelle et la connaissance répartie : dès que quelqu'un connaît une proposition, le groupe la connaît. La connaissance d'un agent, la connaissance de quelqu'un et la connaissance implicite du groupe entretiennent des relations de déductibilité, ce qui, ajouté aux relations précédentes, donne les relations suivantes entre les cinq types de connaissance :

$$C\alpha \Rightarrow T\alpha \Rightarrow K^i\alpha \Rightarrow Q\alpha \Rightarrow I\alpha .$$

Cette hiérarchie permet la formalisation de la majorité des problèmes faisant intervenir de la connaissance dans un groupe, et en particulier de traiter élégamment les échanges d'informations impliqués, en informatique, par la constitution des systèmes distribués (voir [Lejoly 1987]). Néanmoins, ces divers modes de connaissance relèvent tous du même niveau épistémique : qu'elle soit commune ou individuelle, répartie ou de tous, la connaissance formalisée par la logique épistémique ne fait pas intervenir, dans sa sémantique, le processus de justification qui permet la distinction entre la croyance, la proposition vraie et la connaissance, ou entre la connaissance de première main et la connaissance par ouï-dire. Weingartner [Weingartner 1982] a tenté de remédier à cette lacune en introduisant un opérateur de connaissance faible, qui peut intervenir pour nuancer la connaissance qu'un agent possède à propos de la connaissance d'un autre agent.

3.3. La connaissance faible (ou indirecte)

Weingartner fait la distinction entre la conception classique de la connaissance, qu'il appelle la connaissance *forte*, et la connaissance *faible* qui est celle d'un agent qui sait qu'un autre agent sait une chose que lui ne connaît pas. Les logiciens récusent généralement cette dernière forme de connaissance en doutant qu'il s'agisse à proprement parler de connaissance. Weingartner s'en défend et écrit : « This is adequate to real discourse : there are situations where we want to say that someone (A) who has just a knowledge of someone's else (B) knowing p (but *not* himself *enough knowledge*[4] to know that p) that he knows that B knows that p » [Weingartner 1982, 255]. La connaissance *faible* est une forme épistémologiquement bien fondée de connaissance que Weingartner réserve aux agents dont le niveau de connaissance est limité, ce qui est confirmé par un exemple qui met en scène un père et son enfant (« or in general ... a person A whose knowledge is on a lower level as compared to another person B »). La faiblesse de la connaissance ne l'est que relativement à « p » et non à « B sait que p »; de « A sait que B sait que p », on peut inférer « B sait que p » : il s'agit donc bien de connaissance au sens habituel, seule l'impossibilité d'en tirer « A sait que p » la distingue de la connaissance forte.

---

[4] C'est nous qui soulignons.

Nous avons montré ailleurs ([Gochet & Gillet 1991] et [Gillet 1992]) que l'analyse de Weingartner est insuffisante : ce n'est pas le niveau de connaissance qui est concerné par la distinction, mais plutôt la procédure qui permet de connaître : quand le laborantin sait que la solution est acide, il le sait car il a eu recours au papier tournesol; quand le chef du laboratoire sait que le laborantin sait que la solution est acide, il le sait grâce à la consultation du registre journalier.

Dans un groupe, que ce soit une communauté scientifique, culturelle ou un réseau d'ordinateurs, la connaissance faible est probablement le cas le plus fréquent. L'auditeur, qui sait que le musicien connaît la partition, n'est pas censé connaître lui-même la partition, il lui suffit de voir le musicien l'interpréter de mémoire; de la même façon, la base de connaissance qui sait que tel programme sait résoudre une équation n'est pas tenue de la résoudre elle-même. Dans l'acception faible de connaître, on ne présuppose pas que la justification de la connaissance de A implique la capacité de justifier « p », même si la réalisation du fait que B sache « p » implique que « p » soit vrai. À l'expression de connaissance *faible*, je préfère celle de connaissance *indirecte*, qui laisse deviner l'origine de la distinction sans supposer de différence épistémologique fondamentale.

La formalisation de la connaissance indirecte nécessite une généralisation du système multi-agents de la connaissance qui fasse place à la distinction entre ce qui justifie la connaissance forte ou directe (à savoir le recours au monde) et ce qui justifie la connaissance indirecte (à savoir le recours à des informations sur le monde). Je présenterai ici la version sémantique d'une logique qui permet une représentation satisfaisante de la connaissance indirecte, je renvoie le lecteur intéressé par l'axiomatique correspondante et une preuve de complétude à [Gillet 1992].

La thèse qui différencie les connaissances directe et indirecte est la thèse de la transparence « $K^aK^b\alpha \supset K^a\alpha$ » qui affirme la capacité de l'agent A à connaître l'objet de la connaissance de l'agent B. Cette thèse n'est jamais falsifiée par les modèles de la connaissance présentés plus haut. Pour en comprendre la raison, je vais en décomposer le processus de vérification : il se déroule selon la procédure suivante :

(1) $\mathcal{M}, w \models K^aK^b\alpha \supset K^a\alpha \Leftrightarrow \mathcal{M}, w \models K^aK^b\alpha \Rightarrow \mathcal{M}, w \models K^a\alpha$,

(2) $\mathcal{M}, w \models K^aK^b\alpha \Leftrightarrow$ pour tout w' tel que $w\mathcal{R}^a w' : \mathcal{M}, w' \models K^b\alpha$,

(3) $\mathcal{M}, w' \models K^b\alpha \Leftrightarrow$ pour tout w'' tel que $w'\mathcal{R}^b w'' : \mathcal{M}, w'' \models \alpha$,

(4) $\mathcal{M}, w \models K^a\alpha \Leftrightarrow$ pour tout w' tel que $w\mathcal{R}^a w' : \mathcal{M}, w' \models \alpha$.

Pour falsifier la transparence, il faut falsifier « $K^a\alpha$ » tout en vérifiant « $K^aK^b\alpha$ ». Considérons que « $K^aK^b\alpha$ » est vrai, et donc que pour tout w' et pour tout w'' tels que $w\mathcal{R}^a w'$ et $w'\mathcal{R}^b w''$, on a bien $\mathcal{M}, w'' \models \alpha$. Comme $\mathcal{R}^b$ est une relation réflexive, on a pour tout monde w' : $w'\mathcal{R}^b w'$, d'où le monde w' fait partie des w'' qui sont utilisés à la ligne (3) et qui vérifient α. Dès lors la ligne (4) est vérifiée, et il est impossible, dans ce système, de falsifier la thèse de la transparence.

Pour formaliser la connaissance indirecte, il faut distinguer entre les mondes qui représentent ce qui est compatible avec ce qu'un agent est capable de vérifier directement par lui-même, et les mondes qui sont compatibles avec l'information dont il dispose. La réflexivité de $\mathcal{R}^b$ implique que tous les mondes utilisés par B pour vérifier «α» sont également utilisés par A pour vérifier «$K^b\alpha$», or il est possible que les mondes utilisés par B soient insuffisants pour représenter toutes les options compatibles avec ce qui est directement vérifiable par A. La difficulté est d'introduire, dans la vérification effectuée par A, des mondes qui correspondent à l'information indirecte qu'il possède sans altérer la réflexivité de sa relation d'accessibilité, laquelle garantit que l'opérateur K^i représente bien la connaissance, et non une croyance ou une forme dégénérée de connaissance.

C'est pourquoi j'ai généralisé le modèle multi-agents de la connaissance pour y introduire deux types de mondes épistémiquement compatibles : les mondes *publics* qui sont accessibles par tous les agents, et les mondes *privés* réservés à l'usage d'un agent. Les mondes publics sont les mondes compatibles avec les informations disponibles, et les mondes privés sont ceux qui représentent ce qu'un agent peut vérifier directement. Dès lors, le modèle $\mathcal{M}$ comportera un ensemble de mondes $\mathcal{W}^i$ et une relation $\mathcal{R}^i$ par agent; on y trouvera un monde distingué w_0 qui représente le monde réel et qui appartiendra à tous les $\mathcal{W}^i$. La sémantique de la connaissance indirecte ($\mathcal{M}, w \models K^i\alpha$) y est définie différemment selon que w est un monde public ou privé : quand w est un monde public, on a $w \in \mathcal{W}^i$ et c'est la définition standard qui est appliquée, quand w est un monde privé, la sémantique est :

$$\text{si } w \notin \mathcal{W}^i : \mathcal{M}, w \models K^i\alpha \Leftrightarrow \mathcal{M}, w_0 \models K^i\alpha,$$

ce qui revient, pour un monde où l'agent ne dispose pas d'éléments pour vérifier une information, à s'aligner sur les informations disponibles dans le monde réel.

Nous avons maintenant un système qui donne la possibilité de formaliser la connaissance indirecte d'une façon assez satisfaisante puisqu'il s'agit d'une version plus générale du système classique. Mais la question qui se pose alors est celle de la représentation de la connaissance directe, ou plus précisément, la cohabitation des deux types de connaissance. Il est, en effet, peu souhaitable d'avoir à choisir entre les deux versions, même si l'une d'entre elles est formellement plus attrayante, puisque plus générale. Le seul gain pragmatiquement rentable est celui offert par un système d'expressivité supérieure qui permette à la fois la représentation des deux types de connaissance et l'évaluation de leurs interactions.

3.4. La connaissance forte

La plus grande souplesse dans la représentation de la connaissance indirecte laisse deviner la possibilité de formaliser la connaissance forte à partir du système de connaissance faible. Il devrait suffire de rendre possible l'inférence exprimée par la thèse de la transparence « $\mathbb{K}^a K^b \alpha \supset \mathbb{K}^a \alpha$ ». Celle-ci affirme que quand un agent sait *fortement* ($\mathbb{K}^a$) qu'un autre agent connaît une proposition « α », il connaît fortement « α ». Le terme « fortement » ne doit pas nous faire perdre de vue qu'il ne s'agit pas là d'une connaissance d'un niveau supérieur, mais au contraire, d'une connaissance qui ne fait pas la distinction entre la vérification directe et la vérification par le biais d'informations indirectes. Quand un agent connaît fortement, deux cas sont possibles : soit sa connaissance dépend d'une procédure de vérification directe, et dans ce cas, les connaissances directe et indirecte coïncident; soit sa connaissance repose sur des informations obtenues indirectement par l'intermédiaire d'un autre agent, ou par le recoupement d'informations de plusieurs autres agents. Ce dernier cas nous ramène à la connaissance répartie, ou plus exactement, à une combinaison de la connaissance faible et de la connaissance répartie. Ce qui permet de définir la connaissance forte comme un cas dégénéré (du point de vue de la distinction des procédures de vérification et de justification) de la connaissance faible. Je propose donc la définition suivante qui reprend l'alternative proposée plus haut : la connaissance forte est la connaissance qui utilise comme dans sa procédure de justification, soit des ressources propres, soit des ressources répartie dans le groupe, ce qui s'énonce formellement de la manière suivante :

$$\mathbb{K}^i \alpha \underset{\text{déf}}{=} K^i(\alpha \vee G\alpha)\,.$$

Je ne ferai pas ici la preuve que cette définition permet de retrouver très exactement les propriétés que la version classique de la connaissance multi-agents attribue à l'opérateur de connaissance forte. Je me bornerai à indiquer pourquoi elle vérifie la transparence. La thèse de la transparence pour la connaissance forte (« $\mathbb{K}^a K^b \alpha \supset \mathbb{K}^a \alpha$ ») dépend de la relation d'inférence qui existe entre la connaissance de l'agent B et la connaissance répartie sur le groupe. Si B connaît « α », la connaissance de la proposition « α » est accessible au groupe dont fait partie B (« $K^b \alpha \supset G\alpha$ »). Passer de cette relation à la transparence n'est plus alors qu'un jeu formel.

Je terminerai en revenant sur le renversement quelque peu paradoxal qui fait de la connaissance (dite) forte un cas affaibli de la connaissance (dite) faible.

Déductivement la connaissance faible est la plus forte : si un agent connaît indirectement une certaine proposition « α » (« $K^i \alpha$ »), *a fortiori*, il connaît indirectement la disjonction formée par cette proposition et n'importe quelle autre, et donc, par exemple, il connaît indirectement la disjonction formée par cette proposition et la proposition « $G\alpha$ », ce qui constitue la définition de la proposition « $\mathbb{K}^i\alpha$ ». Et ceci établit l'inférence « $K^i\alpha \Rightarrow \mathbb{K}^i\alpha$ ».

Par contre, quand on regarde les rapports entre les connaissances directe et indirecte du point de vue de leur extension, on voit tout de suite que si l'inférence « $K^i\alpha \Rightarrow \mathbb{K}^i\alpha$ » est valide, c'est que « $\mathbb{K}^i\alpha$ » ne peut pas être faux quand « $K^i\alpha$ » est vrai (la réciproque étant possible), et donc, que la connaissance forte est vraie de plus de propositions que la connaissance faible. Si la connaissance forte est vraie d'un plus grand nombre de propositions, ça ne peut être que parce que ses conditions de vérité sont plus lâches, et les conditions de vérité d'une proposition exprimant la connaissance reflètent les conditions épistémologiques associées à cette connaissance. Qualifier la connaissance indirecte (et donc plus étendue) de faible, cesse d'être paradoxal si on veut bien admettre que s'agissant de connaissance, la raison épistémologique précède la raison logique.

4. Conclusion

L'étude de la connaissance que propose la logique épistémique ne se borne pas à établir une liste des inférences valides. La logique épistémique se présente en première analyse comme une science qui permet de repérer et d'examiner les critères formels qui permettent de construire des inférences pour les contextes de connaissance. Ce travail d'analyse conduit à la répartition des différents concepts selon des critères purement déductifs : on distingue les modes de connaissance propres aux individus et aux groupes, et on formalise leurs interactions. Mais l'apport de la logique ne s'arrête pas là. Une deuxième tâche consiste à offrir des outils pour calculer et résoudre les questions concernant la déduction en contexte de connaissance qui échappent très vite à l'intuition et au raisonnement informel. Les progrès de la logique sont, en cette matière, impressionnants : d'une part les performances des méthodes de preuve augmentent continuellement, et d'autre part, les mesures théoriques de faisabilité et de complexité donnent un éclairage inestimable sur les problèmes à résoudre. Mais l'ambition de la logique va plus loin : l'investigation de la dimension logique des concepts liés à la connaissance n'apporte pas que des réponses aux questions logiques, ce sont les concepts eux-mêmes qui s'en trouvent éclairés par la mise en évidence de propriétés structurelles révélées par le niveau d'abstraction et de généralité imposé par la démarche formelle, ainsi que par l'exposition systématique des propriétés essentielles sur lesquelles repose tout l'édifice formel qui modélise les concepts étudiés.

À travers l'examen de la notion de connaissance dans un groupe, j'espère avoir pu contribuer quelque peu au dévoilement de la connaissance.

Bibliographie

GILLET É., 1992, *Contributions à la logique de la connaissance. Omniscience logique et connaissance faible*, thèse de doctorat en philosophie présentée à l'Université de Liège.

GOCHET P. & GILLET É., 1991, "On Professor Weingartner's Contribution to Epistemic Logic", in *Advances in Scientific Philosophy*, Schurz G. & Dorn G.J.W. ed., Poznan Studies in the Philosophy of the Sciences and the Humanities, 24, Rodopi, Amsterdam, pp. 97–115.

HALPERN J. & MOSES Y., 1986, "A Guide to the Modal Logics of Knowledge and Belief", preliminary draft, in *Proc. of the Conference on Theoretical Aspects of Reasonning about Knowledge*, pp. 480–490.

HINTIKKA Jaako, 1962, *Knowledge and Belief*, Ithaca, Cornell U.P.

KONOLIGE Kurt, 1986, *A Deduction Model of Belief*, London, Pitman, Los Altos, Morgan Kaufmann.

KRIPKE Saul, 1963, "Semantical analysis of modal logic I. Normal modal propositionnal calculi", *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*, 9, pp. 67–96.

LEJOLY Philippe, 1987, *La logique de la connaissance et les systèmes distribués*, Mémoire de licence en informatique, Liège.

WEINGARTNER P., 1982, "Conditions of the Rationality for the Concepts Belief, Knowledge and Assumption", *Dialectica*, 36, pp. 243–263.

VON WRIGHT G.H., 1951, *An Essay on Modal Logic*, Amsterdam, North Holland.