

Patrice ENJALBERT*

Sur l'utilisation de la logique en sciences cognitives : questions de méthode

*Les mathématiques qu'il nous faut
construire sont celles de l'esprit humain.*

Georges Boole, *The mathematical
analysis of logic*, 1847.

*We offer this book to all who have
bumped up against inadequacies in the
received view of logic.*

Jon Barwise et John Perry,
Situations and attitudes, 1986.

Cet article se propose de discuter quelques problèmes de méthode relatifs à l'usage de la logique en sciences cognitives. En premier lieu nous rappelons un certain nombre de caractéristiques de la logique formelle, liées à ses origines au sein de l'univers mathématique et qui imposent une attitude prudente lorsqu'il est question de modéliser la cognition. En second lieu nous insistons sur une bonne exploitation de l'ensemble des méthodes logiques et plaidons pour une ouverture sur la calculabilité et sur la théorie de la preuve, en incluant leurs prolongements en informatique théorique. Ces considérations sont appuyées sur une étude de la sémantique des modalités. Finalement nous nous interrogeons sur les conditions d'un dépassement du paradigme logique au profit de théories mathématiques encore à construire.

Mots-Clés : logique, modélisation cognitive, théorie de la preuve, théorie du calcul, sémantique des modalités.

On using logics in cognitive sciences : methodological issues. This paper intends to discuss some methodological problems relative to the use

* GREYC, Université de Caen et CNRS, URA 1526, 14032 Caen Cedex,
Patrice.Enjalbert@info.unicaen.fr.

of logic in cognitive sciences. First we recall a number of characteristics formal logics inherits from its origins in the mathematical universe and which prescribe a cautious attitude when cognitive modelling is concerned. Secondly we insist on taking profit of the various branches of logic and especially of concepts and techniques related to computation and proof theory, including their developments in theoretical computer science. These considerations are supported by a study of the semantics of modalities in natural language. Finally we investigate the conditions of an overtaking of the present logical paradigm in profit of mathematical theories still to be elaborated.

Key Words : Logics, Cognitive Modelling, Proof Theory, Computation Theory, Semantics of Modalities.

INTRODUCTION

Cet article se propose de discuter quelques problèmes de méthode relatifs à l'usage de la logique en sciences cognitives. Précisons d'emblée que nous entendons ici *logique* dans le sens de *logique formelle* ou *mathématique*, ou encore *symbolique*. Le lecteur pourra très provisoirement identifier la logique du premier ordre classique (ou calcul des prédicats) ou quelque extension non classique de son goût, mais précisons également dès à présent que nous serons amenés à élargir très sensiblement la perspective, et dans plusieurs directions. Nous sommes conscients que l'histoire de la logique ne commence pas avec Boole ou Frege, qu'il existe une tradition philosophique bien antérieure, courant d'Aristote à Port-Royal en passant par les médiévaux, et qu'une «logique» non formaliste est sans doute possible encore aujourd'hui. Mais notre réflexion concerne l'usage, répandu notamment sous l'influence de l'I.A. et de certaines traditions philosophiques, d'un certain type de méthodes formelles dans l'analyse de la cognition.

La logique formelle est au départ et avant tout une théorie mathématique — ou peut-être plus justement une *branche* des mathématiques, recouvrant un *ensemble de théories* particulières, fortement intercorrélées, développées pour poser correctement et résoudre des problèmes de divers types, soit internes aux mathématiques elles-mêmes (fondements, problèmes d'algèbre ou d'analyse...) soit externes : problèmes d'ordre informatique (théorie de la programmation, automatisation du raisonnement...) philosophique (il existe une forte tradition logicienne dans la philosophie anglo-saxonne) linguistique (Montague, Kamp, Barwise...) sans oublier de nombreux

«croisements». Nous parlerons alors de logique appliquée. Il n'est pas inutile de noter que la grande majorité des travaux, ceux qui ont donné les résultats les plus profonds et conduit aux avancées techniques les plus importantes, sont certainement du côté des premiers, c'est-à-dire internes à l'univers des mathématiques. Nous ne voulons pas signifier par là que les motivations externes n'ont jamais posé de problèmes féconds aux logiciens, ni qu'il existe une muraille infranchissable entre logique «pure» et «appliquée», mais uniquement prendre conscience d'un état de fait du développement de la discipline, et insister sur son caractère de théorie mathématique.

C'est bien sûr de logique appliquée dont il est question ici c'est-à-dire, si l'on accepte les quelques réflexions qui précèdent, de *modélisation mathématique*. Un ensemble de questions se posent alors. Sur le rapport de modélisation d'abord : implicites véhiculés par le modèle, champ d'application, limites de sa pertinence... Sur le choix des méthodes et théories utilisées ensuite, le modélisateur devant avoir le souci de rechercher les techniques les mieux appropriées à son objet et le cas échéant de les développer dans un sens adéquat ou d'en forger de nouvelles. Ce sont ces questions que nous voulons discuter dans cet article.

Nous commencerons par un rapide aperçu d'ensemble de la logique mathématique, base des discussions ultérieures. Nous nous poserons alors la question de la pertinence de cet appareil théorique en regard de l'étude de la cognition humaine. Un tel recul critique nous apparaît d'une grande importance, compte tenu d'un l'héritage logiciste prégnant, en provenance de certaines écoles philosophiques et linguistiques ou induit par les contacts étroits entretenus entre certaines branches des sciences cognitives et l'I.A. Tendances qui nourrissent elles-mêmes un rejet «en bloc» de la part d'autres écoles, sans d'ailleurs que ces dernières fassent toujours preuve du meilleur discernement ni de la plus grande cohérence¹. À l'encontre de tentations assimilationnistes, nous insisterons sur l'hétérogénéité qui existe selon nous entre la construction mathématique logicienne et la cognition humaine : ce qui pose un problème méthodologique majeur pour lequel nous proposerons une possible attitude.

¹ Ainsi, comme le note D. Andler (1995), les techniques de construction de *modèles mentaux* présentées par le psychologue «anti-logicien» P. Johnson-Laird (1983) sont toutes pétries de notions logiques. En fait elles s'apparentent étroitement à la très classique *méthode des tableaux*.

La troisième partie concerne le choix des techniques logiques à mettre en œuvre : en nous appuyant sur des travaux personnels en sémantique linguistique, nous plaiderons pour un recentrage sur la notion de calcul dans son acception la plus large (pour nous clairement intégrée à la logique) et sur la théorie de la preuve qui lui est intimement liée, plutôt que sur la sémantique ensembliste et la notion de modèle très majoritairement invoquées en la matière. Nous illustrerons à ce propos l'attitude prudente que nous proposons d'adopter en termes de modélisation cognitive. Finalement, nous nous interrogerons sur les conditions d'un dépassement du paradigme logique tel que nous le connaissons aujourd'hui, la «received view of logic» de l'épigraphe, au profit de théories mathématiques qui restent pour nous à élaborer (Nous disons bien dépassement, ce qui suppose une intégration de l'acquis).

Un ultime «avertissement» est de rigueur. Nous ne pensons pas faire œuvre d'épistémologie, ce qui aurait nécessité des études historiques et comparatives, appuyées sur une bibliographie systématique². Notre travail est plutôt une réflexion méthodologique, à partir d'une pratique théorique au croisement de l'informatique théorique, de l'IA et du traitement automatiques des langues. Nous voudrions analyser les difficultés rencontrées dans nos propres recherches ou repérées dans des travaux similaires, de manière à pouvoir les résoudre ou les contourner.

1. LA LOGIQUE DES LOGICIENS

Grosso modo, on peut dire que la logique mathématique s'est développée selon trois grands axes :

— *Théorie des modèles*

Modèle est à prendre ici au sens technique, comme le concept principal de la sémantique des langages logiques. La théorie des modèles (classique) s'intéresse à décrire les modèles des théories axiomatiques, dans le but d'en tirer des informations sur certaines classes de *structures mathématiques* (algébriques, topologiques, etc.)

² Nous ne participerons pas non plus à la guerre des *ismes* : représentationnalisme, cognitivisme, connexionnisme, enactivisme... Nous sommes bien conscients de ce que les questions abordées ici recourent des débats entre ces différentes «écoles». Mais lesdits débats reposent trop souvent sur des caricatures (qui se reconnaîtra dans le martin-pêcheur de Varela ?) et les grands paradigmes étant maintenant bien établis, c'est à l'épreuve des résultats que l'on évaluera la pertinence et la contribution de chacun.

Des résultats originaux peuvent être ainsi obtenus ou un nouvel éclairage sur des résultats précédemment obtenus par des méthodes propres aux différents domaines, ou encore la construction de nouvelles structures (telle l'analyse non standard, qui formalise une notion d'infiniment petit). Dans l'orbite des logiques non classiques, toute la théorie de la correspondance en logique modale³ ressort d'une théorie des modèles (Van Benthem, 1984) et la recherche d'une sémantique extensionnelle est un sujet actuel en logique non monotone.

— *Théorie de la démonstration*

On s'intéresse à la notion de *preuve*. Différentes présentations peuvent être données : déduction naturelle, calcul des séquents... On étudiera par exemple la complexité des preuves d'un même théorème selon différents systèmes déductifs ou langages, ou leur «géométrie» (régularités, équivalences...). Trouver une notion satisfaisante de preuve pour les logiques non monotones demeure un problème ouvert.

— *Théorie de la calculabilité*

À l'origine on trouve la recherche d'une notion intrinsèque de fonction calculable, avec divers modèles : machines de Turing (MT), fonctions récursives, λ -calcul, logique combinatoire... puis l'étude de ces différents modèles et de leurs propriétés. La notion de décidabilité en découle, avec l'étude de méthodes permettant de prouver la décidabilité ou indécidabilité de théories mathématiques. La théorie de la complexité abstraite utilise la notion de MT et ses diverses variantes comme modèle de calcul universel pour évaluer les ressources nécessaires à la résolution d'un problème donné.

Il conviendrait d'ajouter à cet inventaire la théorie des ensembles, mais nous sommes là dans une problématique de fondement des mathématiques décidément trop éloignée des applications en général et de l'étude de la cognition en particulier. Nous voudrions par contre insister sur *l'unité* de ces trois grandes problématiques. Les liens sont multiples. En premier lieu, la notion de langage formel : présumée par les concepts de système déductif et de modèle, elle est nécessaire à la définition des machines de Turing. Les notions de *calcul* et de *preuve* (formelle) sont étroitement liées : une preuve est le résultat d'un calcul, chaque règle d'inférence étant une relation calculable entre

³ Il s'agit de la correspondance entre axiomes modaux et propriétés des modèles de Kripke.

formules ; inversement, tout calcul peut se ramener à une preuve⁴. Dans l'étude d'un système logique on se posera des problèmes de consistance et de complétude, relativement à une classe de modèles donnés. Un type de problème générique consiste à se demander si telle théorie mathématique, formalisée par tel système d'axiomes, est décidable et on pourra à l'occasion utiliser pour ce faire des techniques de théorie des modèles. Etc. Cette unité est une caractéristique essentielle de la logique, qui en fait plus qu'une collection de techniques, un véritable corps théorique de référence.

Qu'en conclure?

En premier lieu nous voulions insister sur la richesse, la diversité, de cet appareillage mathématique, non réductible à la logique du premier ordre, ni à la notion de système logique, de règle d'inférence ou d'axiomatique avec les inévitables «théorèmes de consistance et de complétude» couramment mis en avant dans les travaux logiques en I.A. Soulignons la diversité des systèmes logiques créés par les logiciens : ordre supérieur, logique intuitionniste, modale, non monotone, théorie des types, λ -calcul... Autant de matière pour le modélisateur, autant de cordes à son arc qu'il se doit d'exploiter au mieux. En particulier, l'ensemble n'existe pas sans théorie de la calculabilité, partie intégrante au premier chef de l'édifice — comme en témoignent de nombreux ouvrages généraux de logique (Shoenfield, 1967 ; Barwise, 1988 ; Delahaye, 1986 etc.) — et nous pensons qu'un certain nombre d'idées et de techniques pourraient être recherchées de ce côté : ce point sera illustré dans la section 3.

Mais, aussi riche soit-elle, la logique demeure *une* branche particulière des mathématiques, avec des préoccupations que nous avons tenté de cerner très rapidement, et des techniques spécifiques développées à ce propos (très rapidement de nouveau : plutôt du côté des mathématiques discrètes et de l'algèbre). En prendre conscience doit nous conduire à relativiser son usage en matière de modélisation cognitive : après tout, que peut-il y avoir de commun entre les préoccupations des mathématiciens (fondements, développement de l'algèbre ou de méthodes effectives pour les ordinateurs...) et les cognitivistes ? Quelle pertinence pour ce type de méthode ? Les

⁴ Évaluer $f(a)$, où f est une fonction entière, peut être défini comme la preuve d'une formule $f(a) = b$.

logiques non classiques constituent-elles une adaptation adéquate ? C'est ce que nous nous proposons de discuter maintenant.⁵

2. LOGIQUE ET MODELISATION COGNITIVE

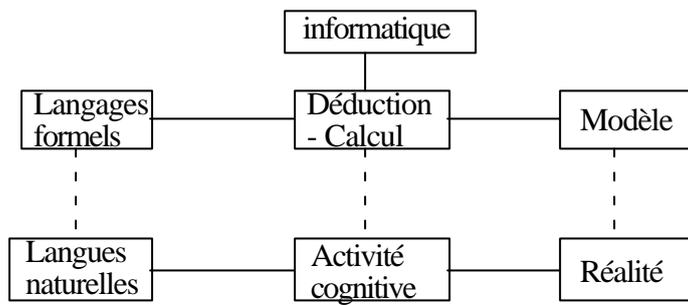
2.1. Le miroir logique

Il est tout à fait remarquable de retrouver dans l'appareil théorique de la logique et dans les grandes lignes de son organisation interne, telle que nous venons de l'esquisser rapidement, un correspondant apparemment fidèle à quelques grandes «instances» à prendre en compte dans l'étude de la pensée : les langages formels pour les langues naturelles ; l'ensemble déduction-calcul pour l'activité cognitive ; le modèle pour la réalité. Insistons brièvement sur ce dernier rapprochement qui sera peut-être moins bien perçu du lecteur. En théorie des modèles classique, nous venons de le rappeler, un modèle est une structure mathématique (ensemble muni de relations et d'opérations : un groupe, un corps, les nombres entiers naturels, les réels...). Ces structures et les objets qui les composent sont bien les *êtres mathématiques* par excellence⁶ les entités qui constituent *l'univers mathématique* depuis l'avènement de la théorie des ensembles. Si l'on prend en compte le prolongement technologique «en acte» à la théorie du calcul que représente l'informatique, nous parvenons au tableau suivant⁷ :

⁵ Nous posons ces questions à propos de la logique, tenant pour acquis que chacun doit balayer devant sa porte. Mais elles peuvent être adressées de la même manière à d'autres types de modèles (mathématiques du continu, variétés différentielles...) qui ont eux aussi une histoire et qui, à ce qu'il nous semble, ne s'enracinent pas non plus dans l'étude de la cognition.

⁶ Tout débat philosophique sur l'existence de ces êtres (platonisme mathématique) écarté.

⁷ On observera qu'il n'y a pas place pour la perception dans ce schéma. Ce n'est pas un hasard (la perception suppose un sujet, évacué de toute théorie mathématique) et certes pas sans conséquences (un type important d'activité cognitive, influant sur la structuration la cognition dans son ensemble, est écarté d'emblée).



S'intéressant à la philosophie de l'informatique, Gérard Chazal (1995) parle joliment de *miroir automate* pour désigner cette situation où nous pouvons reconnaître *d'une certaine manière* nos capacités et notre activité cognitives dans des machines issues de notre propre et ingénieuse création. Nous proposerions volontiers de parler de *miroir logique* pour désigner ce parallèle plus large, appréhendé du point de vue théorique plus que technologique. Certes, personne ne prétendra que l'état actuel de la logique (ou de l'informatique) fournit un reflet très fidèle, mais il nous semble que *l'effet miroir* est à la source même de l'intérêt que les sciences cognitives, à la suite de la philosophie, portent à l'édifice logique. C'est lui que nous voudrions questionner.

2.2. Réflexions critiques sur le logicisme en sciences cognitives

Il sera bon de se souvenir ici des origines de la logique mathématique. Héritière de la philosophie, elle s'en est très clairement détachée en se mathématisant, d'un double point de vue : par ses méthodes — cela doit être assez clair — mais aussi par son objet, tout entier tourné vers les mathématiques elles-mêmes, avec pour fonction essentielle de résoudre la crise des *fondements* qui éclate au début du siècle à la suite notamment des travaux de Cantor. Il fallait à la fois une *notation* (plus qu'un langage) permettant de décrire les théories mathématiques et poser convenablement les problèmes rencontrés et une notion de *dédution* mathématiquement bien définie, ce qui à son tour nécessite d'élaborer une notion claire, intrinsèque, de *calculabilité*. On voit comme on s'écarte de la modélisation du fonctionnement de *l'esprit* envisagée par G. Boole (cf. notre épigraphe) même dans sa part la plus rationnelle. Examinons alors les trois termes du parallèle logique/cognition.

Langages formels et langues naturelles

Dès l'origine, la logique mathématique suppose la définition d'une idéographie totalement affranchie du langage parlé. Définir une telle notation est un des premiers actes de celui qui est considéré comme le véritable fondateur de la logique symbolique moderne, G. Frege. Il

faudrait également évoquer ici la problématique des «langages logiquement parfaits» développée par Russell (auquel nous devons la notation logique aujourd'hui utilisée) caractérisés par la non ambiguïté et la compositionnalité du sens. C'est dire si ces logiciens-là étaient «insatisfaits» des langues telles qu'on les parle⁸ ! Une problématique que l'on retrouve aujourd'hui en informatique avec les spécifications formelles, dont on souhaite qu'elles décrivent les fonctionnalités d'un système informatique sans l'ambiguïté des textes en langue, fût-elle technique.

De là découlent les nombreuses «différences» maintes fois répertoriées entre les deux types de «langages»⁹ : règles de clôture dans la définition des premiers et caractère ouvert, évolutif, des seconds ; phénomène de l'ambiguïté, bannie des langages logiques alors qu'elle semble inhérente aux langues humaines ; absence de sujet énonciateur ou de dimension dialogique dans les langages logiques, etc. (voir par exemple Grize, 1984 et 1990).

Opérations cognitives, déduction, calcul

Discutons d'abord le cas de la déduction et de sa formalisation par règles d'inférences, fondées sur une sémantique ensembliste. En premier lieu ce que l'on peut modéliser de cette manière, de l'ordre du *raisonnement*, ne représente qu'une petite partie de la cognition humaine. Plus encore : la logique formelle constitue une sorte de quintessence de la rationalité, ou plutôt d'ailleurs de la rationalité scientifique, voire mathématique¹⁰. Et encore ne représente-t-on ainsi que l'un des mouvements de l'activité du mathématicien, la preuve d'un théorème, à l'exclusion du processus de recherche ; il est clair également que ce dernier utilise des représentations mentales qui ne sont certes pas *que* propositionnelles, et par conséquent difficilement

⁸ Ces travaux fondateurs achevaient le programme de Leibniz visant à élaborer une *lingua characteristica universalis*, inspirée par l'algèbre et en possédant la rigueur, base d'un *calculus ratiocinator* : où l'on trouve donc déjà le lien étroit entre *langage formel* et *calcul*.

⁹ Il y a quelque chose de pervers dans l'appellation même de *langage formel*, dans la mesure où ce vocable semble indiquer une parenté avec le langage humain. Cela ne pose aucun problème au mathématicien qui a l'habitude d'employer dans un sens qui lui est entièrement propre des termes du langage courant. Mais tout se passe comme si un certain nombre de philosophes, linguistes, informaticiens se laissaient abuser par — soyons un peu provocateur — cette pure *homonymie*...

¹⁰ Pour la logique classique ; nous revenons plus loin sur la modélisation du «raisonnement de sens commun» par les logiques non classiques

réductibles à des formules, et que les pas d'inférence notés dans une démonstration mathématique sont infiniment plus complexes et chargés de sous-entendus que les pas d'inférence, extrêmement élémentaires, dans un système formel. Bref, s'il est question de représenter l'activité mathématique c'est tout au plus celle d'un «mathématicien idéal abstrait», ayant peu à voir avec des mathématiciens en chair et en esprit. Tout cela n'a rien de surprenant si l'on veut bien se rappeler la démarche historique qui a conduit à la naissance de la logique formelle.

Par ailleurs, un certain nombre d'expériences de psychologie cognitive (telles que la tâche de Wason) semblent indiquer un hiatus sans appel entre la pensée commune et la logique : performances désastreuses sur des tâches triviales du point de vue logique ; fonctionnement constructif et concret opposé à la notion de règles de déduction formelle (Johnson-Laird, 1983).¹¹

La notion de calcul paraît plus libérale, dans la mesure où nous ne sommes pas astreint au format très particulier des règles de déduction, et où nous sommes affranchis de «l'exigence» de validité sémantique des règles dans une sémantique ensembliste. Elle a pour cela la faveur de certains chercheurs en I.A., *via* divers paradigmes de programmation. Mais un trait fondamental demeure : l'assimilation cognition/calcul suppose la réduction des opérations cognitives à une manipulation de symboles¹². Derrière se profile l'hypothèse d'un «mentaux», un «langage de l'esprit», qui plus est assimilé à un langage formel. Ce qui pose divers types de problèmes bien connus. D'abord,

¹¹ D. Andler (1995) doute de la pertinence de ces expériences, qui ne remettent pas en cause selon lui l'aptitude du sujet humain «logiquement naïf» à effectuer des raisonnements logiques élémentaires. Sa discussion des procédures expérimentales en question est tout à fait intéressante, mais nous ne sommes pas convaincu de ses conclusions. Brièvement : ce n'est pas parce que des sujets ordinaires peuvent *comprendre* certaines règles de déduction (ils répondent correctement à une question du type : «vous savez que P implique Q. Maintenant qu'en déduisez-vous?») que celles-ci constituent les étapes élémentaires de son raisonnement courant. Nous sommes également capables de comprendre des opérations arithmétiques, sans que ces dernières soient vraiment typiques de nos capacités en terme de *résolution de problèmes*.

¹² Autre terme «piégé» : un symbole dans ce contexte est une pure graphie, ou son équivalent électronique. Peu de chose à voir avec le symbole comme signe, au sens de Peirce, qui doit renvoyer à une interprétation. En logique, on manipule des «symboles» indépendamment de toute signification qui puisse leur être attribuée par ailleurs, c'est là un trait essentiel de la formalisation. Le double usage de ce terme est-il anodin?

d'après ce que l'on peut savoir du système neuronal, le manque de fondement naturaliste. Reste alors la possibilité, à un niveau plus abstrait, de modélisation des «représentations mentales» : l'assouplissement par rapport aux représentations propositionnelles est-il suffisant ? Cela n'est pas évident (pensons par exemple à l'imagerie mentale ou autres impressions ou simulations sensorielles, ou même aux représentations manipulées par les mathématiciens, comme évoqué plus haut).

Modèles et réalité

Les sciences de la cognition ont sans doute besoin de développer une conception adéquate de «la réalité». Mais l'assimilation modèle/réalité pose divers problèmes des plus sérieux. En premier lieu le paradigme s'accorde parfaitement avec le réalisme, c'est-à-dire de l'idée d'une «réalité» dotée d'une structure propre, en soi — par opposition à un point de vue constructif, selon lequel les notions, les concepts (et la catégorisation des objets selon ces concepts, etc.) sont élaborés par des sujets, tributaires de représentations culturellement déterminées. Un modèle au sens logique est la *donnée* d'un ensemble et de relations et fonctions définies sur cet ensemble, une correspondance étant établie entre le vocabulaire du langage logique et ces relations et fonctions. C'est parfait pour les mathématiques puisque les «structures» étudiées préexistent au langage logique qui les décrit, qui n'est en fait qu'une *notation*. La transposition à la cognition humaine est pour le moins hasardeuse ; ce dont nous avons besoin c'est d'étudier la *constitution* de représentations du monde dans l'activité cognitive.¹³

Ensuite, dans le parallèle modèle/réalité, le premier est une mathématisation de la seconde, et l'on peut s'interroger sur les présupposés et la richesse de l'appareillage mathématique mis en œuvre. 1) La notion de modèle présuppose des entités, des objets, parfaitement différenciés et individualisés, ce qui est parfaitement problématique par exemple s'il est question de substances, d'étendue, de milieux continus¹⁴ ou encore d'entités abstraites (l'eau versée dans le vase, le macadam devant la porte qui fond au soleil, la beauté sans rivale de la comtesse...). 2) Aucun physicien n'accepterait de travailler avec des notions aussi *élémentaires* que celles de relation et de

¹³ Dans notre modèle (cf. section 3), aucun «modèle-réalité» donné *a priori* de l'activité des agents, mais des «systèmes conceptuels» propres à chaque agent.

¹⁴ Voir (Smith et Casatti 1993) pour une critique d'une physique naïve à base ensembliste.

fonction «pures», ni de recourir à l'axiomatisation formelle pour définir des notions plus élaborées ! On objectera peut-être que les mathématiques peuvent être codées ou traduites en langage logique : certes, mais ceci pour des besoins de fondement, ou peut-être d'automatisation, mais ce n'est qu'une transcription formalisatrice, qui ne doit pas cacher le travail d'élaboration du mathématicien appliqué. Pourquoi en serait-il autrement lorsque l'on étudie la manière dont un sujet cognitif se représente le monde qui l'entoure ?

Enfin, *l'unité de l'édifice logique* n'est pas sans conséquence. Assimiler les langues humaines aux langages formels suppose implicitement l'assimilation des «procédures de traitement», soit respectivement la cognition et les MT (ou tout autre modèle équivalent). Les règles de la déduction naturelle classique sont étroitement liées à la théorie des ensembles, dont elles sont en quelque sorte la contrepartie effective et dont nous venons de mettre en cause, sinon la pertinence, du moins l'universalité en termes de cognition. Les méthodes informatiques relèvent du même mode de pensée (il existe des sémantiques extensionnelles des langages de programmation), donc également les langages de représentation de connaissance de l'I.A. (dans lesquels par exemple les concepts deviennent des objets) : mes collègues formalisateurs «anti-logiciens» en sont-ils conscients ?

2.3. Logiques non classiques

On pourra objecter que le tableau qui vient d'être dressé concerne une vision très classique de la logique, sans tenir compte du travail réalisé pour modéliser le raisonnement dit «de sens commun», et notamment de l'apport des logiques non classiques, dont c'est (ou peut être) un objectif. Nous discuterons ici très brièvement la problématique du raisonnement «incertain», portant sur des données imprécises ou incomplètes. Le bilan doit selon nous être nuancé.

D'un côté, ces travaux peuvent faire état de résultats cognitivement pertinents, nous évoquerons deux exemples pour illustrer de cette position. Le premier concerne la *logique non monotone* que l'on peut présenter de manière synthétique comme résultant d'un affaiblissement de la relation de conséquence logique classique : comment faisons-nous, à partir d'un ensemble d'hypothèses A_1, \dots, A_n pour *sauter* à la conclusion B, alors même que l'ajout d'informations complémentaires pourraient conduire à remettre cette conclusion en cause ? On note $A_1 \dot{\sim} \dots \dot{\sim} A_n / \sim B$ cette relation de conséquence *révisable*. On capte ainsi par exemple des notions de typicité (*oiseau / \sim vole* mais *oiseau \dot{\sim} autruche / \sim \neg vole*), de persistance temporelle (si on me dit que le

gardien a allumé son phare à 19h 40 et l'a éteint le lendemain à 6h 50, je conclus qu'à minuit il était allumé, alors que si une panne de l'alimentation électrique est mentionnée de 23h à 1h, cette conclusion est rejetée), etc. Diverses méthodes ont été proposées pour fonder logiquement ce type d'inférence, et certains résultats sont saisissants. Ainsi Kraus, Lehmann, Magidor (1990) traitent l'exemple suivant (dit diagramme de Nixon) : si les républicains sont typiquement non pacifistes et les quakers pacifistes (*républicain* / \sim \neg *pacifiste*, *quaker* / \sim *pacifiste*), que conclure si on est républicain et pacifiste? si on est républicain et ouvrier? etc. Le système en question produit des réponses d'une grande plausibilité : ni *républicain* \dot{Y} *quaker* / \sim \neg *pacifiste*, ni *républicain* \dot{Y} *quaker* / \sim *pacifiste* (il n'y a aucune information pour trancher dans un sens ou un autre) mais *républicain* \dot{Y} *ouvrier* / \sim \neg *pacifiste* (*a priori*, le fait d'être ouvrier est sans influence), et *républicain* / \sim \neg *quaker* (les républicains typiques ne sont pas quakers) etc. Ici nous voyons à l'œuvre une modification des règles de la *déduction naturelle* classique, qui rend compte d'un trait important des procédures de raisonnement courant.¹⁵

Notre second exemple concernera la logique floue (Bouchon-Meunier, 1993), qui apporte (entre autres !) une réponse à la formalisation de notions aux frontières mal déterminées, extrêmement courantes en langue (*grand, cher, la trentaine...*) avec un traitement de modifieurs linguistiques tels que *très, peu, un peu*, etc. Rappelons l'idée, très simple dans ses prémisses, qui est de considérer un domaine de valeurs (tailles, prix, ages...) et pour chaque prédicat flou A, une fonction caractéristique f_A allant continûment de 0 à 1, par exemple pour $A = \text{la trentaine}$: $f_A(30) = 1$, $f_A(20) = 0$, $f_A(28) = 0,8$, etc. Les modifieurs ont pour fonction de modifier cette fonction caractéristique. Certes, la représentation de situations concrètes n'est pas simple et se coule plus ou moins bien dans ce moule, mais la modélisation existe et, semble-t-il est d'une certaine manière validée par un certain nombre de réalisations de type système expert.

On pourrait multiplier les exemples. Cependant, échappons nous aux grandes caractéristiques de la logique relevées plus haut? Nous ne le pensons pas. Nous restons dans le cadre de représentations

¹⁵ Ce qui laisse à penser que la déduction naturelle classique n'est pas absolument la base du raisonnement «naturel». Étant donnés les liens intimes qu'elle entretient avec la théorie des ensembles, nous pouvons être conduits à remettre en cause l'ontologie ensembliste — nouvel exemple de conséquences de l'unité de la logique, et peut-être source des difficultés rencontrées dans la définition de sémantiques extensionnelles pour les logiques non monotones.

propositionnelles et il s'agit toujours de rendre compte de la rationalité sous une forme qui est simplement assouplie. Notons à cet égard que les travaux en logique non classique ont un caractère *normatif* : comment raisonner rigoureusement, compte tenu des informations dont on dispose, en restant conscient de leur incertitude et des limites des conclusions qui seront tirées? Il est bien clair également que les développements mathématiques fondant ces systèmes logiques ne sont nullement établis par modélisation de processus mentaux (cf. par exemple la notion de *mondes préférés* dans Kraus et al., 1990). Finalement, le principe du raisonnement formel, c'est-à-dire sur des formes syntaxiques n'est nullement remis en cause.

2.4. Premières conclusions

Nous observons donc que la logique formelle moderne n'a pas été conçue dans un but de modélisation cognitive, mais pour de toutes autres tâches, internes aux mathématiques¹⁶ et qu'il en résulte un certain nombre de traits qui posent problème à quiconque tente de l'utiliser à d'autres fins. Tenter d'observer notre activité cognitive dans le *miroir logique*, serait comme se mirer dans un portrait peint par Fernand Léger : on aura du mal à s'y observer, soi-même, en toute objectivité et réalisme... Même dans les approches non classiques, développées précisément pour se rapprocher du raisonnement «naturel» et non «mathématique», un certain nombre de caractéristiques essentielles à l'idée même de logique formelle demeurent. Souligner cette hétérogénéité nous a paru d'une grande importance, précisément parce que la logique est une théorie remarquablement riche, et que le parallèle logique-cognition en trois points relevés plus haut est extrêmement attractif.

Mais cela ne signifie nullement que les méthodes logiques ne soient d'aucune pertinence. Si les développements qui précèdent ont été principalement et unilatéralement critiques, c'est que notre but est ici de cerner les limites d'une approche purement logicienne, de manière à faire émerger de nouvelles pistes de travail, et que nous avons fui un impossible et ennuyeux «bilan». Mais nous venons d'évoquer quelques résultats, à notre sens probants, en logique non classique. Bien d'autres auraient pu être alignés et la prochaine partie présentera de nouvelles propositions, issues de nos propres travaux : le portrait par F. Léger, l'esquisse de Picasso, peuvent nous apprendre beaucoup sur certains traits de la personnalité du modèle ou de la nature humaine. Nous sommes alors devant une sorte de paradoxe, dont la solution peut être envisagée dans trois directions complémentaires.

1) Précautions méthodologiques. Toute modélisation de type logique (dans notre sens «large», incluant la théorie du calcul et ses prolongements informatiques) ne peut d'abord être que partielle, concernant un aspect très spécifique de la cognition, et même du raisonnement. En second lieu, il convient de bien repérer le *biais* dû aux caractéristiques intrinsèques du formalisme logique évoquées plus haut. Une manière prudente de penser le rapport d'un système logique à son correspondant cognitif se situe alors sur le mode de la *métaphore* : une analogie partielle peut être relevée et posséder une vertu

¹⁶ Il faut rappeler ici l'antipsychologisme de Frege.

heuristique précieuse mais elle ne doit pas être confondue avec une authentique théorisation de l'objet cognitif.

2) À l'intérieur de ces limites, nous pouvons choisir quelle branche de la logique utiliser, quelles notions et méthodes sont les plus pertinentes pour le phénomène visé : ce choix fait partie du travail du modélisateur qui doit faire preuve en la matière de souplesse et d'imagination.

3) Envisager les conditions d'un dépassement du paradigme logique sous sa forme actuelle au profit de méthodes mathématiques qui intégreraient ce qui s'y trouve le plus pertinent, abandonnant ce qui est moins adéquat, et ouvrant possiblement sur d'autres paradigmes et techniques.

3. SEMANTIQUE ET CALCUL

Cette troisième partie concerne le choix des techniques logiques à mettre en œuvre, en illustration des points 1) et 2) ci-dessus. En nous appuyant sur des travaux personnels, nous voudrions y plaider pour un centrage sur la notion de calcul dans son acception la plus large (pour nous clairement intégrée à la logique) et corrélativement sur la notion de preuve, plutôt que sur la sémantique ensembliste. Nous illustrerons à ce propos l'attitude prudente que nous proposons d'adopter en matière de modélisation cognitive. La question abordée concerne la sémantique des modalités¹⁷.

3.1. Sémantique des modalités

La question de la sémantique des modalités reste mal résolue. On connaît les deux grandes techniques utilisées dans un cadre apparenté à la théorie des modèles de la logique du premier ordre. Sémantique des *mondes possibles* d'une part : *l'agent A croit P* si P est vrai dans toutes les alternatives épistémiques que A peut concevoir, c'est à dire tous les *états des choses*, les *mondes possibles* compatibles avec ses croyances et connaissances (bien sûr incomplètes). Et approches dites *syntaxiques* de l'autre : *A croit P* est interprété comme une relation

¹⁷ Épistémiques (*croire que, savoir que/qui/quand/où...*), intentionnelles (*vouloir, souhaiter...*) et plus généralement des attitudes propositionnelles (incluant en outre des expressions telles que *dire que, voir, entendre...*). Dans la suite, sauf précision contraire, *modalité* signifiera *modalité épistémique*. L'étude des modalités a ceci d'intéressant que, bien que très partielle et spécifique dans le vaste univers de la sémantique linguistique, elle oblige à prendre une position nette sur la question du sens, notamment dans son rapport à la référence.

entre A et l'énoncé P, c'est à dire le texte même de cet énoncé, ou mieux : un terme qui en constitue la syntaxe abstraite et le dénote. Il n'est pas possible de discuter sérieusement en quelques lignes ces théories, mais reprenons simplement, pour situer notre propos, les principaux problèmes posés à toute théorie des modalités et les difficultés majeures rencontrées par les deux approches citées :

— rendre compte des différentes notions et expressions langagières, exprimant divers degrés ou nuances de savoir ou de croyance (explicite ou implicite, opposition *savoir/croire*, locutions *savoir qui/quand/où...* et non seulement *que* etc.). En particulier, une pierre d'achoppement de la formalisation modale est *l'omniscience logique*, c'est à dire la clôture (à un degré plus ou moins accentué selon les théories) des croyances d'un individu par conséquence logique. On voit mal une sémantique réaliste et nuancée reposer sur une telle hypothèse. Ajoutons que la plausibilité cognitive de la notion modale de monde possible nous semble extrêmement douteuse. Et notons que la sémantique des locutions *savoir qui/quand/où...* n'a reçu finalement que peu d'attention, et les solutions classiques, basées sur une analyse en termes de référence, ne sont guère satisfaisantes (Enjalbert et Stéfani, 1990).

— rendre compte de la référence en présence de modalité (problème répertorié comme *opacité référentielle*). Pour reprendre un exemple canonique, quelle est la référence du groupe nominal *le nombre des planètes* dans l'énoncé : «Hegel croyait que le nombre de planètes est plus grand que 5»? Cet énoncé est faux dans son acception la plus naturelle¹⁸, mais si l'expression en question réfère à la «réalité», soit 9, elle devrait être juste. En gros, les théories citées permettent très bien de bloquer l'inférence qui conduirait de «Hegel croyait que 9 est plus grand que 5» à l'énoncé cité, mais l'interprétation de la solution n'est pas évidente. Par exemple la solution modale interprète l'opacité comme multiplicité de référence d'un même terme dans les différents mondes possibles, ce qui pose, outre le problème de la plausibilité des mondes possibles, celui de l'identification d'objets entre mondes et nous paraît peu exploitable dans une analyse linguistique.

D'où, nous a-t-il semblé, la nécessité de repenser la modélisation dans un tout autre cadre que nous allons maintenant esquisser. Le lecteur pourra se référer pour plus de détail à (Beysade et al., 1995 et 1996 ; Beysade, 1994).

3.2. Une société d'agents cognitifs

¹⁸ Du temps de Hegel on ne dénombrait que 5 planètes.

Nous partons d'une idée fort simple : dire qu'un agent *A* croit que la proposition *P* est vraie, c'est exprimer non un état des mondes possibles ou une relation entre un agent et un énoncé, mais une propriété du système cognitif de *A*. Cette manière de voir s'impose d'autant plus que l'on considère une classe d'expressions plus large que les traditionnels *savoir/croire que*, par exemple : *A sait comment réaliser telle tâche, A doit se douter que P, A a tous les moyens de savoir P* etc. C'est bien des capacités cognitives dont il est question en toute généralité. Cette remarque nous conduit clairement à mettre au premier plan une notion d'agent cognitif, avec une modélisation minimum de son activité.

En second lieu, il faut se poser la question : *qui* formule tel jugement sur les croyances et capacités de *A*. Dans les approches purement logiciennes susmentionnées, c'est un «observateur extérieur» exprimant dans un langage logique (modal ou du premier ordre), censé avoir un «sens» objectif c'est-à-dire relatif à «la réalité» ou à un modèle qui en tient lieu. En vérité, ce ne peut être qu'un autre agent *B alter ego* de *A*, qui ne possède lui-même qu'une vision imparfaite et fondamentalement incomplète de cette «réalité» — dont d'ailleurs l'activité de *A*, ses croyances et compétences, sont partie intégrantes¹⁹. Nous voilà donc devant une *société d'agents*. Et les modalités épistémiques servent à un agent *B* à formuler ce que l'on appelle couramment en informatique une *spécification* (partielle) de l'activité d'un second, *A*.

Notre point de vue général de la sémantique des langues est homogène avec ces premières idées. Il postule (ou observe !) que le langage est une activité cognitive socialement régulée²⁰. Autrement dit, les constructions d'une langue n'ont pas de sens *en soi*, mais sont l'objet d'une interprétation, toujours individuelle *hic et nunc*. Dire cela implique que le sens — quoi que l'on entende par là — n'est pas à rechercher directement dans une réalité extérieure, mais dans l'univers conceptuel de chaque agent.

Le danger serait là de se diriger vers un solipsisme généralisé : plus de vérité autre que subjective, plus de sens au langage que spécifique à chacun. Une telle position serait difficile à défendre : il faut une vérité et un sens qui transcendent l'interprétation et le jugement individuels.

¹⁹ D'un point de vue linguistique, il s'agit de réintroduire le *locuteur* comme terme de l'analyse.

²⁰ Et régulatrice des activités cognitives de ses locuteurs. Mais cette question n'est pas abordée ici.

Notre thèse est qu'ils proviennent de l'activité sociale, des échanges entre agents, et des parentés structurelles que cette activité et communauté d'expérience induit sur le fonctionnement de chacun. Vient alors dans le modèle la notion cruciale de *communauté épistémique*, dont la notion de communauté linguistique est un exemple primordial : il s'agit d'un ensemble d'agents dont les structures cognitives sont (relativement à un domaine de connaissance donné) suffisamment proches pour que la construction du sens ou la prise en compte d'informations s'opèrent de manière similaire chez les uns et les autres. C'est alors en quelque sorte l'arbitrage de la communauté qui fixe la norme aléthique et sémantique. Et c'est cet «arbitrage» qui joue dans notre proposition le rôle de l'inatteignable «réalité».

3.3. Le modèle

Il nous semble donc que l'idée d'une *société d'agents cognitifs* doit être au centre d'une analyse des modalités épistémiques. Par rapport aux approches basées sur la théorie des modèles et ses extensions, il est clair que nous y perdons en abstraction. Mais devant les difficultés récurrentes rencontrées par cette approche, il est plus que légitime d'envisager d'autres voies et de revenir à une analyse concrète de l'objet d'étude. Encore convient-il de se donner un modèle des agents suffisamment abstrait pour autoriser l'étude théorique. Nos agents sont des processus logiques, abstraction formalisante d'agents rationnels. Chaque agent A est doté d'un langage de représentation de connaissances dit *interne* L_A , que nous choisissons d'ordre supérieur, basé (dans la version actuelle) sur le lambda calcul simplement typé de Church (Andrews, 1986). Un agent peut effectuer des preuves ou des calculs. Pour représenter (et ultérieurement implémenter) cette activité, il est commode de le doter d'un métalangage du même type, dans lequel il peut formuler des *jugements*. Un jugement est donc différent d'une proposition. Des exemples de jugements sont :

(*true P*), pour « P est vrai», (*proof Arg P*) pour « Arg est une preuve de P » ou «un argument pour P », (*eval t u*) pour «l'évaluation du lambda terme t produit u ».

Cette technique est aujourd'hui utilisée pour spécifier des systèmes logiques arbitraires et réaliser des démonstrateurs de théorèmes génériques. Nous pouvons par exemple considérer que les agents sont des processus Λ -Prolog (extension de Prolog en logique d'ordre supérieur, cf. Nadathur et Miller, 1990), dans lesquelles la dérivation de jugements tels que ceux présentés ci-dessus peut être programmée. Ainsi, une règle d'inférence d'introduction du et peut être codée :

(proof et_intro(Arg₁,Arg₂) A_1 A_2) :— (proof Arg₁ A_1), (proof Arg₂ A_2).

où *et_intro* est un *foncteur* constructeur de preuves, ce qui peut se lire : *et_intro*(Arg₁,Arg₂) est une preuve de A_1 A_2 si pour $i=1,2$ Arg _{i} est une preuve de A_i .

D'autre part les agents communiquent par envoi de messages, termes d'un langage du même type dit *langage de communication* L_C , qui doit clairement être distinct des divers L_A : en effet, ces derniers sont spécifiques de l'activité de chaque agent, alors que L_C doit (dans sa forme) être partagé. Un agent A interprète les expressions de L_C en

direction de ses propres concepts. Techniquement, cela veut dire qu'il peut porter un nouveau type de jugement :

(*interp T C*) : le terme T de L_C est interprété comme le concept C

3.4. Sémantique des modalités

Donnons maintenant les grandes lignes de l'interprétation des modalités dans ce modèle.

a. Différentes formes de croyances

Nous aurons chez chaque agent une contrepartie de ces différents jugements permettant à chacun de spécifier l'activité de ses congénères. L'opérateur principal est noté p ; $(p A Arg P)$ a pour signification souhaitée : l'agent A est susceptible d'accepter Arg comme preuve de P , c'est-à-dire : peut dériver le jugement (*proof Arg P*). La notion de *croyance implicite* (pour des agents rationnels) A *croit P* sera interprétée comme «A est susceptible d'avoir une preuve de P» et codée : $\$Arg (p A Arg P)$. Une notion de *croyance explicite* fera en outre référence à une présence de P en mémoire de A etc. Enfin, *savoir où est X* c'est connaître une *propriété* de X qui puisse constituer, dans le contexte courant, une localisation pertinente²¹ : ce qui s'exprime aisément et naturellement en logique d'ordre supérieur.

b. L'alternance savoir/croire

Le savoir est traditionnellement présenté comme croyance vraie. Une variante plus fine, défendue par exemple par A. Bonomi (1992) dans une problématique cartésianiste (suivant Ayer et Gettier) suppose en outre que l'agent ait une justification correcte de la proposition sue. (Le *savoir* comme une *croyance justifiée*). Poursuivant cette idée, nous dirons *qu'un agent A sait P* si la preuve qu'il en possède ou est susceptible d'en posséder (savoir implicite) peut être acceptée par la *communauté* des agents compétents sur la question P — ce qui peut parfaitement être codé dans notre modèle. On voit là le type de «réalisme» que nous proposons de substituer à la référence directe au monde réel (ou au modèle, au sens logique, qui en tient lieu) : celui d'une société élaborant la connaissance et de communautés de compétences reconnues, dont les assertions prennent de ce fait le statut de «vérités objectives».

²¹ Ainsi pour «où est Caen», ce peut être aussi bien "en Normandie" que "à l'ouest de Paris, à deux heures de train" — plus plausibles que le couple latitude-longitude auquel nous conduiraient les approches référentielles modales ou syntaxiques.

c. *Référence et opacité*

La référence n'est pas pour nous une relation entre expressions langagières et objets du «monde» (ou du modèle logique), supposé donné indépendamment d'un agent particulier, mais une relation entre concepts élaborés par l'agent. Nous distinguons deux classes de concepts, *complexes* et *primitifs*. Par exemple : (*mère Œdipe*), (+ 2 3) désignent des concepts complexes, quelque chose comme : «l'individu en relation de mère à fils avec l'individu Œdipe» et «le résultat de l'addition de 2 à 3». Les concepts primitifs correspondants sont respectivement «l'individu Œdipe en tant que tel» et «le nombre entier 5». Un opérateur d'évaluation permet le passage des concepts complexes aux concepts primitifs. Par exemple (*croit A (+ 2 3) 5*) stipule que A peut évaluer (+ 2 3) en 5, et (*croit A (mère Œdipe) Jocaste*) stipule que la référence, en quelque sorte *calculée*, de (*mère Œdipe*) est *Jocaste*.

On dispose alors du langage formel permettant de formaliser les mécanismes de la référence dans des contextes modaux. Nous avons notamment pu mener une étude fine de phénomènes *linguistiques* tels que l'ambiguïté référentielle ou encore la projection des présuppositions. Concernant le premier problème, soit par exemple la phrase bien connue «Œdipe voulait épouser sa mère». D'un point de vue linguistique, cet énoncé, comme tout ceux où des descriptions définies apparaissent dans des contextes modaux, est ambigu. Trois lectures peuvent en être faites, que l'on peut paraphraser comme suit :

- (1) Œdipe voulait épouser celle qu'il croyait à tort être sa mère.
- (2) Œdipe voulait épouser celle dont je sais qu'elle était sa mère.
- (3) Œdipe voulait épouser celle qu'il savait être sa mère.

L'ambiguïté réside donc dans la personne qui prend en charge la description définie *sa mère* : le sujet de l'attitude propositionnelle (ici Œdipe), le locuteur, ou les deux. En général, l'interprétation «normale» suppose une prise en charge à la fois du sujet et du locuteur, comme cela apparaît dans la phrase voisine «Jeanne voulait épouser l'instituteur». Dans l'exemple d'Œdipe, c'est l'interprétation (2) qui sera retenue à cause, sans doute, d'un interdit culturel concernant l'inceste.

Dans notre modèle, nous rendons compte de ce type de phénomène en postulant que, dans une situation de communication (i.e., un envoi de messages), les agents interlocuteurs ainsi que ceux mentionnés dans le message forment une *communauté de communication*. Les membres

d'une telle communauté sont supposés partager un ensemble de connaissances, en particulier les références des descriptions définies. En d'autres termes, il est supposé une bonne correspondance entre leurs systèmes cognitifs : une conceptualisation similaire du monde, des connaissances similaires sur les relations entre ces concepts, et des inférences similaires lors de l'interprétation des messages. Ce n'est toutefois qu'une hypothèse, elle peut donc être remise en cause si une autre information la contredit. Ainsi, un agent recevant le message «Jeanne voulait épouser l'instituteur» va inférer, par défaut, que tant l'énonciateur que Jeanne savent que la personne que Jeanne voulait épouser était instituteur ; par contre, dans le cas d'Œdipe, l'inférence sera bloquée parce que contradictoire avec une connaissance préalable concernant l'inceste.

3.5. Conclusion — Discussion

En trois points.

1) Concernant les techniques employées

Nous avons recours de manière essentielle à quelques concepts simples issus de la *théorie de la preuve* pour modéliser l'activité cognitive des agents — et non à la théorie des modèles, modale ou de premier ordre classique. Mais cela ne suffit pas : nous nous situons dans le cadre très général du *lambda-calcul* et nous avons également recours à la notion proprement informatique (versant technologique de la théorie du calcul) de *système multi-agent* pour pouvoir *mettre en scène* les relations entre agents autonomes, modéliser leur communications, et traiter du côté social de l'affaire.

2) Sur la méthode

Les limites du modèle doivent être immédiatement soulignées. Il ne saurait être question pour nous d'assimiler des agents humains à ces agents formels. C'est pourquoi nous parlons volontiers de métaphore logico-informatique. Le recours à la logique induit un biais dans l'analyse : définitions des opérations cognitives en termes de manipulation d'un langage interne formel (L_A) ; séparation des langages de communication et de représentation des connaissances, là où il n'y a dans la cognition humaine que différenciation de différentes fonctions du langage ; réification des concepts, etc.

Nous pensons que notre «métaphore», c'est-à-dire finalement l'analogie entre le langage et la cognition d'agents humains en société d'une part, et le fonctionnement de nos agents logiques, eux aussi «en

société», permet en effet de rendre compte de certains phénomènes linguistiques. Sa vertu est celle de l'abstraction : au lieu de considérer la langue dans sa richesse et la cognition dans sa complexité, elle permet de se concentrer sur un certain type de phénomènes et de proposer une organisation générale en rendant compte. Mais il ne faut pas «pousser» l'expérience plus loin qu'elle ne peut nous conduire. Sa fonction est plutôt heuristique, un «cadrage» général qui devra être ensuite relayée par des études proprement linguistiques — et l'a déjà été pour partie dans (Beyssade, 1994). Notre position pourra paraître ambiguë, au milieu du gué ; nous pensons au contraire qu'elle ne fait que refléter lucidement la *tension* inévitable entre modèles formels de type logique (au sens large) et l'activité humaine qui constitue leur objet. En attendant de posséder des théories mathématiques plus fidèles que le *miroir logique*.

3) Sémantique des langues

Cette discussion devrait être poursuivie sur le terrain général de la sémantique des langues, pour lequel nous défendons un point de vue *constructif*. Comprendre un texte c'est, à partir d'informations portées par la structure linguistique, conjointement à des connaissances sur le domaine concerné et sur les conditions d'énonciation, *construire des représentations* mentales (ou informatiques, s'il est question de traitement automatique)²². Un texte n'a donc pas de sens en soi, mais uniquement pour un interprète et relativement à son système de croyances — que nous appelons son *modèle du monde* (bien évidemment des régulations culturelles assurent une certaine cohérence entre les interprètes). Et l'objet d'étude de la sémantique est précisément de décrire comment se fait la construction du sens. Des méthodes formelles adaptées à ce point de vue doivent évidemment accorder une place centrale à une notion de *processus constructif* — construisant des *représentations* relevant d'un certain *système*. Ce qui nous écarte tant de la sémantique extensionnelle que de la pure déduction. Ces questions sont abordées d'un point de vue sémiotique très général dans (Enjalbert, 1996) et dans (Enjalbert et Victorri, 1994) du point de vue du traitement automatique des langues.

4. CONDITIONS POUR DEPASSEMENT DES METHODES LOGIQUES ACTUELLES

Les deux précédentes parties ont tenté d'une part de montrer une voie insuffisamment explorée, et de l'autre les limites que le travail de

²² C'est la fonction de l'opérateur *interp* dans la métaphore logique multi-agents.

modélisation logique le plus astucieux ne pourra pas franchir. Se pose alors la question d'un dépassement des techniques logiques actuelles. Nous appelons de nos vœux une réflexion de ce type, et allons pour conclure poser quelques jalons dans ce sens.

4.1. Structures et processus

«Dépasser» implique «intégrer les acquis», et donc déterminer, au-delà des formes particulières, les caractéristiques essentielles qui rendent la logique productive et attractive. Le nœud de l'affaire est d'ordre *computationnel* — et non pas du côté de la théorie des modèles, ni même dans le cadre particulier de la déduction, contrairement à ce que les travaux en sémantique formelle ou en logique non classique pourraient induire à penser. Ce qui est remarquable et qui fait le succès de la logique c'est que cette dernière nous fournit à la fois une notion de *structure* et de *processus* opérant sur ces structures, d'une étonnante simplicité. Précisons que nous entendons ici *structure* en un sens général, disons : *ensemble de données interreliées* ; les structures manipulées en logique s'appellent formules (de la logique prédicative), λ -expressions, termes de la logique combinatoire, preuves... reliant divers symboles (prédicats, fonctions, variables, constantes, formules...) selon certaines règles de formation. Les processus sont les règles d'inférences et autres procédures de construction de preuves, la réduction des λ -expressions etc.

Ces structures et processus peuvent être mis en correspondance avec l'idée de *représentations mentales* et de *processus mentaux*. Précisons aussi que si nous employons ici le terme de *représentation*, conformément à l'usage, c'est dans un sens extrêmement neutre. Le terme de *structure* ou de *configuration mentale* nous conviendrait mieux, en évacuant toute idée d'image de «la réalité» supposée pré-structurée : les représentations ou structures dans notre sens sont de pures constructions mentales, qui peuvent être parfaitement imaginaires, délirantes, ou (même !) «en prise» sur la réalité perçue, mais toujours sans la moindre exigence de correspondance étroite avec quelque donnée extérieure prédéfinie. La notion de représentation suscite depuis plusieurs années de nombreux débats et critiques. Certaines sont liées à une philosophie réaliste dont nous venons de nous démarquer. D'autres aux formes logico-informatiques avancées pour en rendre compte, sans toujours prendre de grandes précautions méthodologiques, et nous avons pris position clairement sur ce point dans cet article : les modes de description logico-informatiques posent en effet problème. Mais pouvons-nous nous passer de la notion de

représentation (c'est à dire pour nous de l'idée assemblages d'objets mentaux interreliés) et d'appréhender les (certains) processus mentaux comme consistant à accéder à ces données structurées, les modifier, les construire ? Nous ne le pensons pas.

Sur ce point les modèles alternatifs fondées sur les mathématiques continues et les variétés différentielles posent à notre sens deux types de problèmes (voir par exemple, Fuchs et Victorri, 1992) : 1) la notion *d'espace sémantique* ne permet pas de rendre compte de l'idée de structure, au sens défini plus haut, alors que les notations discrètes (en terme de graphes ou de formules) y excellent ; 2) la notion abstraite de *fonction* considérée n'a pas le caractère opératoire et en particulier constructif dont nous avons besoin, et qui est par contre inhérent à l'idée de fonction calculable.

4.2. Quelques pistes de travail

De ces réflexions découlent quelques orientations.

1) La logique a des origines et une histoire, et son évolution vers les sciences cognitives nécessite la remise en cause de certains traits tels que le primat de la vérité et de la déduction (non réalisée par les logiques non classiques). Mais après tout n'est-ce pas le cas de la *logique* combinatoire et du λ -calcul? La voie est donc ouverte, vers une conception large de la logique comme science mathématique des structures et des processus opérant sur ces structures, selon des règles particulières. Un appareillage théorique, des méthodes mathématiques restent à élaborer pour modéliser les représentations/structures mentales et les processus correspondants. C'est nous semble-t-il la voie que trace un J-B Grize, vers ce qu'il appelle la *logique naturelle* (Grize, 1990). Mais une plus grande imagination technique sera nécessaire pour parvenir à ces «mathématiques de l'esprit humain» annoncées par G. Boole, qui ne soient pas celles de l'esprit du «mathématicien idéal abstrait». Typiquement, il faudra bien poser la question des langages formels et envisager d'autres *structures calculables*, en un sens qui reste à définir.

2) Ce programme suppose de travailler la notion de représentation mentale et de préciser sa place dans une économie générale des sciences cognitives, allant de l'étude du fonctionnement neuronal à la symbolisation, en passant par la perception. *A priori*, cette place se situe du côté du symbolique. J-B Grize précédemment cité situe son objet d'étude autour de l'activité langagière. Une extension vers le sémiotique nous paraît raisonnable. Avec une remarque importante : en fait nous n'avons pas vraiment de notion adéquate du *symbolique*, et

les modèles formels (logiques ou informatiques) sont insatisfaisants, voire caricaturaux lorsque le symbole est conçu comme renvoyant «à la réalité» ou au modèle qui en tient lieu. Les modèles à venir devront précisément nous apprendre à penser le phénomène différemment.

3) Pour autant, on ne peut pas faire l'impasse sur le substrat physiologique de la cognition. Les modèles à élaborer devront tendre à tenir compte de caractéristiques liées au fonctionnement neuronal. La recherche d'une forme de compatibilité peut être éclairante et stimulante, même pour l'étude des processus symboliques les plus abstraits.

4) Finalement, logique ou pas, il conviendrait de réfléchir sur la place d'une modélisation mathématique en sciences cognitives, en tenant compte des expériences dans d'autres domaines des sciences humaines, et qui ne vont pas non plus sans rencontrer de grandes difficultés (économie, sociologie...). En bref : l'objet se prête-t-il à mathématisation ? jusqu'à quel point ? qu'apportent les modèles formels par rapport à d'autres pratiques scientifiques ? Les succès obtenus grâce aux mathématiques dans les sciences physiques ne sont pas nécessairement reproductibles en sciences humaines en général et dans les sciences cognitives en particulier.

Bibliographie

- Andler, D. (1995) Logique, raisonnement et psychologie, *Méthodes logiques pour les sciences cognitives*, J. Dubucs, F. Lepage éd., Hermes, pp. 25-75.
- Andrews, P. (1986) *An introduction to mathematical logic and type theory: to truth through proof*, W. Rheinboldt (éd.), Academic Press.
- Barwise J. éd. (1988), *Handbook of Mathematical Logic*, North-Holland.
- Beyssade C. (1994) *Les modalités épistémiques dans un système multi-agent contribution à l'étude linguistique des modalités et de la présupposition*, Thèse de l'Université de Caen.
- Beyssade C., P. Enjalbert P., Lefevre C. (1995) Cooperating Logical Agents — Application to the analysis of Belief and Knowledge and to the Semantics of Epistemic Modalities in N.L., *2nd World Conf. on the Fundamentals of A.I., WOCFAI '95*.
- Beyssade C., Enjalbert P., Lefevre C. (1996) Sémantique des modalités épistémiques et systèmes multi-agents, in Chazal G., Terrasse M.-N. (éd), *Philosophie du langage et informatique*, Hermès, pp. 147-162.
- Bonomi A. (1992) Persistent truths, *Intellectica* 13-14, pp. 79-103.
- Bouchon-Meunier B. (1993) *La logique floue*, Presses Universitaires de France.
- Chang C.C., Keisler H.J. (1977) *Model Theory*, North-Holland, Amsterdam.

- Chazal G. (1995) *Le miroir automate, Introduction à une philosophie de l'Informatique*, Champ Vallon.
- Delahaye J.-P. (1986), *Outils logiques pour l'intelligence artificielle*, Eyrolles.
- Enjalbert P. (1996) De l'interprétation (sens, structures et processus), *Intellectica*, à paraître.
- Enjalbert P., Stefani J.-B. (1990) «Savoir qui, quand, où... Une approche centrée sur les notions de question et de contexte de résolution de problème», *Actes du 4ème colloque de l'ARC*.
- Enjalbert P., Victorri B. (1994) Du langage au modèle, *Traitement Automatique des Langues*, 35, 1, pp. 37-64.
- Fuchs C., Victorri B. (1992) C. Fuchs, B. Victorri, Construire un espace sémantique pour représenter la polysémie d'un marqueur grammatical : l'exemple de *encore*, *Linguisticae investigationes*, XVI:1, pp. 125-153.
- Grize J.-B. (1984) Langues naturelles et langages formels, *Réseaux*, n°3, pp. 231-286.
- Grize J.-B. (1990) *Logique et Langage*, Orphys.
- Johnson-Laird P. N. (1983) *Mental Models*, Cambridge University Press.
- Kraus S., Lehmann D., Magidor M. (1990) Nonmonotonic reasoning, preferential models and cumulative logic, *Artificial Intelligence*, 44 (1-2), pp. 167-207.
- Nadathur G., Miller D. (1990) Higher-order Horn Clauses, *Journal of the Association for Computing Machinery*, 37, n° 4, pp.777-814.
- Shoenfield J. R. (1967) *Mathematical Logic*, Addison Wesley.
- Smith B., R. Casatti R. (1993) La physique naïve : un essai d'ontologie, *Intellectica* 17, pp. 173-198.
- Van Benthem J. (1984) Correspondance Theory, *Handbook of Mathematical Logic*, D. Gabbay, F. Guenther eds, D. Reidel, pp. 167-248.