Thèse présentée pour obtenir le grade de

DOCTEUR DE L'ÉCOLE POLYTECHNIQUE

Domaine : **Sciences de l'Homme et de la Société**

Spécialité : **Épistémologie**

par

Monsieur Bruno BACHIMONT

Titre :

Herméneutique matérielle et Artéfacture : des machines qui pensent aux machines qui donnent à penser

Critique du formalisme en intelligence artificielle

soutenue le 24 mai 1996 devant le jury composé de :

Monsieur	Jean-François	Boisvieux	*examinateur*
Monsieur	Jean-Gabriel	Ganascia	*examinateur*
Monsieur	Frédéric	Nef	*rapporteur*
Monsieur	Jean	Petitot	*directeur*
Monsieur	François	Rastier	*examinateur*
Monsieur	Jean-Michel	Salanskis	*rapporteur*
Monsieur	Bernard	Stiegler	*examinateur*

À Isabelle, Marie, et ...

Il m'est agréable de commencer ce mémoire en remerciant les personnes qui, par leur sympathie pour mon projet et leur compétence intellectuelle, retrouveront un peu d'elles-mêmes dans ce travail, même si ses inévitables imperfections ne me sont qu'imputables.

Je n'aurais pu en effet mener à son terme ce travail sans la bienveillance de mon directeur de thèse Jean Petitot, directeur d'études à l'EHESS, qui m'a toujours accordé sa confiance. Il n'a pas ménagé son temps ni sa peine pour m'aider à revoir les premières versions de ce travail. Mais ma dette excède le cadre de cette thèse dans la mesure où c'est dans ses travaux que j'ai appris à mêler réflexions philosophiques et préoccupations scientifiques et techniques. J'aimerai profiter ici de l'occasion pour lui témoigner mon admiration pour son œuvre intellectuelle et sa grande probité.

Jean-Michel Salanskis, professeur de philosophie à l'université de Lille, m'a toujours conseillé avec intelligence et pertinence ; il a été pour moi un guide dans mon itinéraire philosophique et un modèle en incarnant, à mes yeux, une authentique réflexion philosophique dans son ascèse et sa richesse. Je le remercie d'avoir accepté de rapporter sur ce travail.

Frédéric Nef, professeur de philosophie à l'université de Rennes, m'a relu en me faisant profiter de toute son intelligence philosophique et de sa rigueur. Je suis honoré qu'il ait accepté de rapporter sur mon travail et l'en remercie.

Je dois beaucoup à Bernard Stiegler, philosophe et enseignant-chercheur à l'Université de Technologie de Compiègne. Outre son animation de l'équipe d'enseignement et de recherche à laquelle je participe avec bonheur sous sa férule, Bernard m'a fait progresser sur des points critiques de mon travail, notamment grâce à sa compréhension fine et unique des problèmes de la technique. Bernard est un ami cher et je veux profiter de l'occasion offerte ici de lui dire l'admiration que m'inspire son travail et la jubilation intellectuelle que j'éprouve à l'entendre l'expliquer.

J'ai rencontré François Rastier, directeur de recherches à l'Institut de la Langue Française, alors que j'effectuais mon travail de doctorat en intelligence articielle et mon itinéraire intellectuel lui doit beaucoup. C'est à François Rastier que je dois d'avoir compris les enjeux sémiotiques en intelligence artificielle. Je tiens à le remercier tout particulièrement pour la bienveillance avec laquelle il a toujours accueilli mes travaux et la finesse avec laquelle il les a critiqués.

Rapporteur de ma thèse en Intelligence artificielle, Jean-Gabriel Ganascia, professeur d'informatique à l'université Paris 6, avait vigoureusement soutenu mes travaux à l'époque. Il a bien voulu examiner un travail de lecture difficile pour un informaticien non philosophe. Il m'a montré à sa manière et dans les termes qui sont les siens l'importance des enjeux liés aux questions soulevées ici et m'a confirmé la nécessité de poursuivre cet effort. Pour tout cela, je le remercie.

Enfin, je tiens tout particulièrement à remercier Jean-François Boisvieux, professeur en médecine à l'université de Paris 6-Pitié-Salpêtrière, directeur du laboratoire au sein duquel j'exerce mes activités professionnelles en informatique médicale. Il m'a en effet permis de développer des problématiques qui, sans sa bienveillance et sa compréhension, n'auraient certainement pas pu aboutir.

Il est difficile de citer toutes les personnes dont l'amitié et la compétence m'ont été précieuses. Cependant, il m'est impossible de ne pas mentionner ici Yves-Marie Visetti pour lui témoigner mon admiration pour son intelligence profonde et exigeante et mon amitié. Yves-Marie a l'art de me faire progresser sur des points cruciaux par sa capacité, au détour d'une conversation, à me faire mieux comprendre ce que j'avais écrit ou voulu dire. J'aimerai également remercier l'équipe réunie autour de Bernard Stiegler, constituant l'« École de Compiègne » qu'il a fondée. Je pense surtout à Charles Lenay, John Stewart et Véronique Havelange.

Le travail sur les ontologies constituant la première partie de ce mémoire renvoie à de nombreuses discussions avec mes collègues de l'intelligence artificielle et de la linguistique. Je pense à Gilles Kassel, Manuel Zacklad, Didier Bourigault, Annes Condamines et les autres partenaires du groupe « Terminologie et IA » auquel j'ai le plaisir de participer. Je n'aurais pas pu développer l'ontologie médicale sous-tendant la réflexion de la première partie sans l'aide de Pierre Dujols, professeur d'informatique médicale à Montpellier. J'aimerai lui réaffirmer ici mon amitié et mon admiration pour ses valeurs humaines et son intelligence.

Enfin, je voudrais associer à ces remerciements les membres du Service d'Informatique Médicale de l'Assistance Publique - Hôpitaux de Paris. Je voudrais remercier tout spécialement ici Jacques Bouaud et Jean Charlet, collègues mais avant tout amis, qui m'ont toujours soutenu dans les périodes de doute et d'incertitude, Pierre Zweigenbaum, Brigitte Séroussi, Vincent Morice, Vincent Brunie, Sawsan, Frédéric, Béatrice, Anne-Claude, et tous ceux que je ne peux citer auxquels ces remerciements sont également adressés. Je fais néanmoins une mention spéciale pour Alain Fauconnet, ingénieur système du SIM, qui cache sous cette appelation un ami serviable et impressionant de compétences.

Bien qu'il y ait de multiples occasions de dire à ses proches combien leur assistance et leur compréhension furent précieuses, il m'est agréable de conclure ces remerciements en leur adressant mes plus affecteuses pensées.

Paris, le 24 mai 1996,

Bruno Bachimont

Table des matières

Première Partie : Artéfacture et ontologies en intelligence artificielle

Chapitre 2 : Ontologie et intelligence artificielle

Chapitre 3 : Ontologie et philosophie

Chapitre 4 : Ontologie et artéfacture

Deuxième Partie : Artéfacture et herméneutique matérielle

Chapitre 5 : Sens et épistémologie transcendantale

Table des figures

Chapitre premier

Introduction
Intelligence artificielle, artéfacture et herméneutique matérielle

Le travail qu'on va lire est une « critique du formalisme en intelligence artificielle » (IA). Dans son projet de construire des systèmes à base de connaissances (SBC) effectuant des tâches ou résolvant des problèmes dans des domaines de spécialité, l'IA doit déterminer quelles connaissances sont nécessaires pour traiter ces problèmes ou tâches et définir comment les rendre exploitables informatiquement, c'est-à-dire, les opérationnaliser. Le principal problème que rencontre l'IA pour réaliser ce projet et qui mobilise l'essentiel de l'effort méthodologique, est que la plupart des connaissances dont on dispose sont exprimées en langue naturelle. Le problème que l'IA doit alors affronter dans la conception de SBC est de définir la nature de ces connaissances pour déterminer la manière de les opérationnaliser. Or, dans ce cadre, l'IA [1] trouve dans le formalisme [2] à la fois une conception philosophique de la connaissance et une théorie scientifique des systèmes informatiques qui lui permettent de formuler un programme méthodologique de recherche pour la conception de SBC.

Une approche formaliste

En affirmant en effet que toute connaissance est une connaissance uniquement dans la mesure où elle est formelle [3], c'est-à-dire dans la mesure où elle correspond à une structure logique, le formalisme affirme que l'IA, pour dégager des connaissances exprimées en langue le savoir qu'elles recèlent, doit les représenter formellement par des structures syntaxiques. Il s'agit de rapporter d'une part les expressions symboliques informelles de la langue à des formules logiques consistant en des assemblages syntaxiques de symboles ininterprétés. D'autre part, il s'agit de rapporter les raisonnements mobilisant les expressions linguistiques et leur contenu à des inférences ou déductions entre les formules logiques qui les formalisent : on obtient ainsi un système formel, c'est-à-dire un ensemble de symboles ou signes ininterprétés (alphabet) pour lesquels on dispose d'une part de règles permet-

1. Nous justifions et détaillons cette caractérisation de l'IA dans la section 1-2.

2. Nous explicitons dans la section 1-3 ce que nous entendons exactement par « formalisme ».

3. Ceci est une thèse *épistémologique* et non *psychologique* sur la connaissance. Elle ne se prononce pas sur la nature mentale d'une connaissance, mais sur le rapport épistémologique liant une connaissance à ce dont elle est la connaissance.

tant de construire des formules (syntaxe formelle) et d'autre part de règles permettant de construire des preuves, inférences, etc., entre ces formules. Ce qu'affirme le formalisme, c'est que le système formel obtenu, dégageant le savoir sur le monde contenu dans les expressions linguistiques, est un modèle[4] du domaine. La formalisation des connaissances modélise en effet le domaine dans la mesure où les représentations formelles *dénotent*, font *référence*, aux objets du domaine. Alors que les connaissances exprimées en langue reposent sur leur interprétation linguistique pour véhiculer un savoir sur le monde, leur formalisation véhicule un savoir en dénotant le monde : en d'autres termes, après avoir rapporté les expressions aux représentations, les raisonnements aux inférences, le formalisme rapporte l'interprétation linguistique à la dénotation. Le formalisme est alors un *principe de modélisation* fondé sur un postulat épistémologique sur la nature de la connaissance, à savoir sa formalité.

Nous avons rappelé qu'un système formel se définit comme étant un ensemble de règles permettant de manipuler des assemblages de symboles ou signes ininterprétés : « ininterprété » signifie que cette manipulation s'effectue de manière mécanique ou machinale sans faire appel à une quelconque intuition ou attention, ou quelque marque d'intelligence, au sens où, idéalement, une machine pourrait le faire. Or, une machine capable de manipuler des symboles ininterprétés ou signes formels existe, c'est la machine universelle de Turing, c'est-à-dire le concept mathématique de ce qu'est un ordinateur. Autrement dit, à partir moment où une représentation de connaissance est un assemblage syntaxique de symboles, que tout raisonnement mobilisant cette connaissance se ramène à appliquer une règle d'une manière formelle, sans faire appel à la signification des symboles, un ordinateur peut la manipuler et rendre *automatique* le raisonnement. Elle est interprétée par l'ordinateur comme une instruction qu'il exécute dans la mesure où la présence d'une telle représentation commande l'application des règles d'inférences. Par conséquent, la notion même de formalisation, consistant à transcrire dans un formalisme ininterprété les connaissances exprimées linguistiquement, permet d'opérationnaliser le savoir contenu dans ces connaissances et conduit ainsi, couplé au formalisme comme principe de modélisation, à la réalisation de l'objectif que l'IA s'est fixé : obtenir un système qui résout un problème en opérationnalisant les connaissances dont on dispose alors qu'elles ne s'expriment que linguistiquement. En d'autres termes, l'IA construit des systèmes formels automatiques modélisant les problèmes à traiter et opérationnalisant par conséquent leur résolution (Haugeland, 1981). Le formalisme est alors un *principe d'effectivité* fondé sur la réalité scientifique et technique de l'informatique (voir la figure 1.1 qui résume ces propos).

Une critique sous forme d'un double constat

La possibilité de critiquer un tel programme méthodologique ne peut concerner le formalisme comme principe d'effectivité dans la mesure où celui-ci ne correspond pas à un *choix* méthodologique mais à une *propriété* de l'approche adoptée : formaliser les connaissances donne *par définition* un système effectif et calculable de représentations. De plus, une telle propriété est une bonne propriété dans la mesure où, dans les termes où nous avons formulé le projet de l'IA, l'objectif est de construire des systèmes opérationnels, en particulier des machines concrètes. D'une certaine manière, le principal attrait du formalisme comme principe de modélisation réside dans le fait qu'il s'accompagne nécessairement

4. Nous précisons dans la section 1-3.1 ce que nous entendons par les termes de « modèle », « formalisation », « axiomatisation » ou « mathématisation », etc.

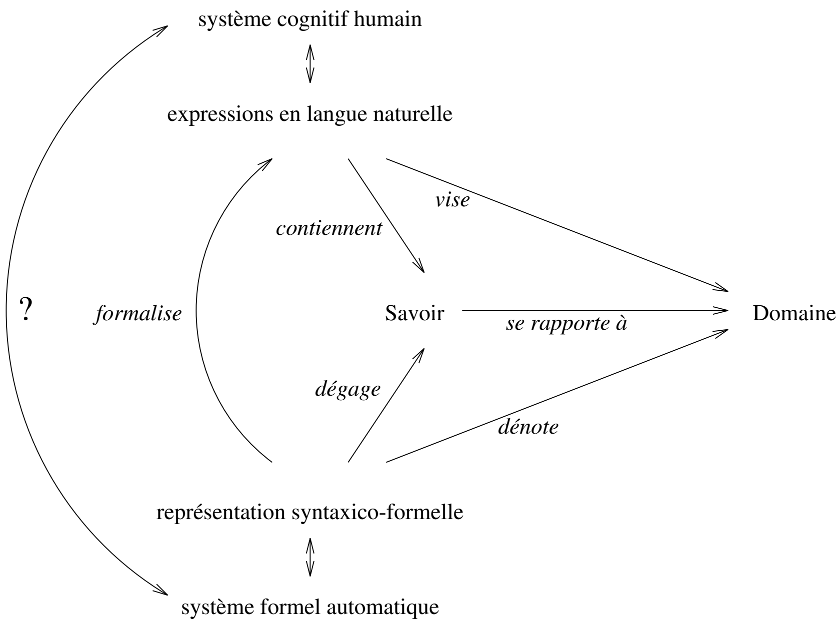


FIG. 1.1 – *L'approche formaliste en* IA*: des connaissances exprimées en langue naturelle vise une réalité dans le cadre de laquelle un problème doit être traité; elles visent cette réalité en explicitant un savoir à son égard; la formalisation dégage ce savoir et produit des représentations syntaxico-formelles qui constituent dès lors un modèle de la réalité visée par les connaissances linguistiques. Ce modèle, étant syntaxico-formel, est effectif et implémentable sur un ordinateur: son exécution opérationnalise les connaissances et permet de résoudre le problème posé.*

du formalisme comme principe d'effectivité. Néanmoins, ces deux aspects du formalisme ne sont pas équivalents, et l'effectivité des formalisations effectuées en IA ne saurait par elle-même assurer le fait qu'elles soient des modèles. Autrement dit, le principe de modélisation doit être valide par lui-même indépendamment du fait qu'il s'accompagne d'effectivité ou non. Par conséquent, s'il existe une critique à exercer quant au formalisme en IA, c'est à ce niveau qu'il faut mener l'interrogation.

La nécessité de critiquer le formalisme comme principe de modélisation et donc comme théorie de la connaissance est née pour nous de notre pratique effective de l'IA dans le cadre de la réalisation de systèmes à base de connaissances dans un contexte médical. Nous attachant à réaliser le programme formaliste de l'IA en nous conformant à son principe de modélisation, nous nous sommes confronté à une difficulté majeure qui compromet voire annule la possibilité même de ce programme. Cependant cette difficulté, dirimante quant aux fondements, doit être examinée à la lumière de deux propriétés essentielles des formalisations effectuées en IA:

La principale difficulté de l'approche formaliste en IA consiste dans le fait que les connaissances exprimées en langue naturelle héritent de la langue la propriété de véhiculer des contenus dont la nature ne correspond pas à des structures syntaxico-formelles; la langue naturelle et sa sémantique sont non compositionnelles et non référentielles: la langue ne décrit pas le monde, ni le dénote. Si la langue permet néanmoins d'exprimer des connaissances sur le monde, c'est que le rapport d'une telle connaissance au monde n'est pas de nature dénotationnelle : il faut alors abandonner la thèse selon laquelle toute connaissance n'est connaissance qu'en tant qu'elle est formelle ;

la formalisation des connaissances exprimées en langue ne permet ni de dégager ni de conserver leur contenu ou le savoir qu'elles véhiculent sur le monde ; en d'autres termes, formaliser des connaissances ne permet pas de constituer un modèle de la réalité visée par ces connaissances au sens où le modèle dénoterait cette réalité ;[5] ; la dénotation n'est pas une relation permettant d'établir un rapport de modélisation entre un système formel et la réalité visée dans une pratique exprimée linguistiquement : *les systèmes formels automatiques ne sont pas des modèles dénotationnels du monde* ;

une première propriété de l'approche formaliste consiste dans le fait que les représentations syntaxico-formelles sont symboliques et peuvent de ce fait adopter la même forme symbolique que les énoncés de la langue naturelle. Le respect lors de la formalisation des connaissances de la structure symbolique de leur expression en langue naturelle confère aux représentations syntaxico-formelles une interprétabilité à laquelle tout locuteur de la langue accède spontanément ; en particulier, la représentation des connaissances constitue un système formel dont la signature fonctionnelle et relationnelle (voir l'encadré 1.1) sont des termes symboliques correspondant à des unités linguistiques de la langue qui sont à ce titre spontanément interprétables par tout spécialiste du domaine traité ; les inférences effectuées par le système formel automatique entre des représentations exprimées dans des termes interprétables en langue possèdent de ce fait un corrélat au niveau de l'expression linguistique des connaissances ; du fait de cette interprétabilité, ils véhiculent des connaissances ;

une seconde propriété de l'approche formaliste est que les symboles, exprimés sous forme de signifiants linguistiques, sont enrôlés dans des processus computationnels dynamiques : les symboles possèdent une sémantique opérationnelle dans la mesure où ils interviennent dans les calculs déterminés par le système formel ; cette productivité du symbole informatique permet à la connaissance à laquelle il renvoie quand il est interprété linguistiquement d'être explicitée par les interprétations des inférences symboliques dont il est une prémisse ; l'effectivité du système formel, le fait qu'il soit automatique, permet d'établir des liens entre les connaissances qui l'interprètent et de les évaluer selon les critères[6] de cohérence, complétude, et systématicité de manière à pouvoir reconnaître dans ce système et son interprétation un *modèle* du domaine ; l'effectivité des systèmes formels permet de dégager, en rationalisant et systématisant l'interprétation linguistique des représentations symboliques, le savoir qu'elles véhiculent ; ce n'est pas la formalisation qui permet de dégager le contenu des connaissances, mais l'effectivité de la formalisation et la signifiance linguistique des représentations symboliques qui permettent de construire un système de représentations interprétées, c'est-à-dire un modèle[7] ;

Alors que les systèmes formels automatiques élaborés en IA ne sont pas modèles dé-

5. Nous revenons sur ces problèmes dans la section 1-5 et au chapitre 2.

6. Ces critères constituent ce que nous appelons un « modèle » ; tout ensemble de connaissances n'est pas obligatoirement un modèle du monde mais doit vérifier pour cela certaines propriétés, à savoir les critères mentionnés ici. La notion de modèle et la définition des critères sont données dans la section 1-3.1.

7. Ces deux dernières propriétés sont détaillées dans la section 1-6.

notationnels, c'est-à-dire des représentations formelles dénotant le monde, leur caractère symbolique et effectif leur permet d'être interprétables linguistiquement et de renvoyer aux connaissances dont ils sont la transcription formelle, et donc d'être utilisés comme des modèles.

C'est donc ce double constat qui constitue le point de départ de notre travail : d'une part les systèmes formels automatiques ne sont pas des modèles formels ou dénotationnels du domaine, mais d'autre part ils fonctionnent néanmoins comme des modèles. Un tel constat se démarque du formalisme comme principe de modélisation pour insister sur le caractère effectif et symbolique des représentations formelles qu'une démarche formaliste permet d'élaborer.

Interprétabilité, dénotation et modélisation

Comme nous l'avons souligné, le système formel n'est pas un modèle du fait de son formalisme, mais du fait de l'interprétation linguistique qui l'investit et de la systématisation de cette interprétation par l'effectivité du système. Dans le cas du formalisme comme principe de modélisation, les connaissances opérationnalisées dans le système formel automatique font sens dans la mesure où elles correspondent au savoir que l'on a sur le domaine (voir la figure 1.2). Dans le cas où ce système formel automatique ne peut rendre compte par lui-même des connaissances du domaine, mais en propose seulement une image possible de celles-ci du fait de son interprétabilité linguistique, le problème est de savoir comment constituer à l'aide du système formel un sens correct dans le domaine permettant de remplir la tâche pour laquelle il a été conçu. A partir du moment où ce n'est pas un modèle dénotationnel, la manière dont il fait sens est livrée au hasard du choix des unités linguistiques pour exprimer la signature du système, et à l'arbitraire de la compréhension que peut en avoir l'utilisateur.

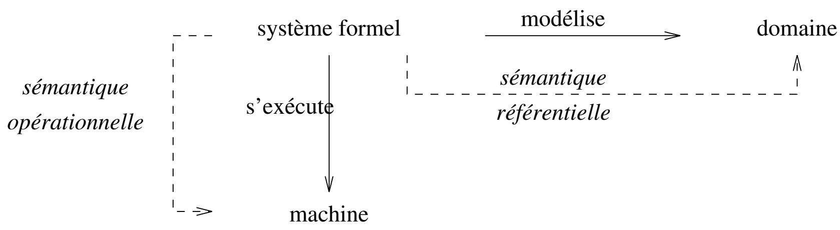


FIG. 1.2 –*Le système formel est à la fois un modèle du domaine dans la mesure où il le dénote, et un système effectif.*

L'enjeu est donc de contrôler la manière dont ce système fait sens dans le domaine, puisque, de toute évidence, il fait sens. Il faut donc faire en sorte que l'interprétation spontanée du système formel en fasse un modèle, c'est-à-dire qu'il soit investi d'un sens correct dans le domaine. Mais, puisque le formalisme n'est pas un principe de modélisation, le contrôle devra s'effectuer avec des moyens autres que le formel. Le système formel n'est pas un modèle du domaine par lui-même, comme dans le cas où il dénoterait la réalité du domaine, mais seulement du fait de son investissement linguistique de sens, c'est-à-dire en considérant le statut linguistique des symboles qu'il mobilise. C'est l'interprétation en langue qui lui permet de véhiculer un savoir sur le domaine, et donc de devenir à ce titre un modèle du domaine. Puisque les connaissances sont exprimées dans la langue et n'existent, ne sont élaborées et comprises, que dans la langue, il est nécessaire que le système formel soit appréhendé au niveau du sens linguistique qu'il peut recevoir pour que l'on puisse reconnaître quelles connaissances il exprime et contient. Mais, puisque le formalisme ne rend pas compte du linguistique, c'est depuis le linguistique qu'il faut contrôler ce que le système formel veut dire, et non contrôler depuis le formel ce que les connaissances en langue doivent vouloir dire (voir la figure 1.3).

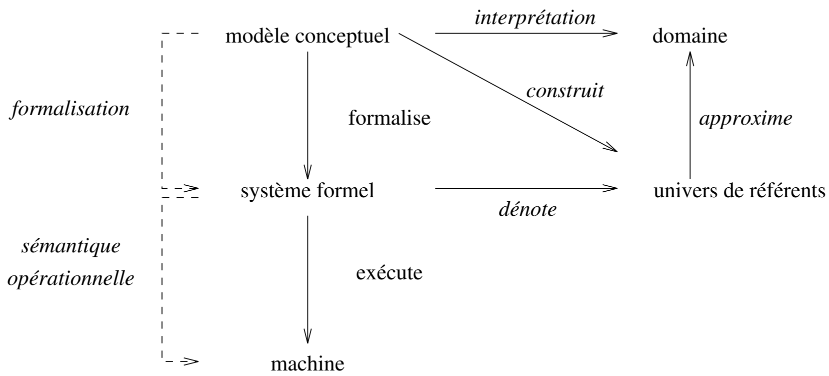


FIG. 1.3 –*Le système formel ne modélise pas le domaine mais formalise le modèle du domaine. Pour le considérer comme un modèle, il faut l'interpréter en lui réinjectant la sémantique du modèle qu'il formalise. Il est néanmoins nécessaire de disposer d'un tel système dans la mesure où le formalisme comme principe d'effectivité assure son opérationnalisation.*

Le problème de l'approche formaliste est donc que, malgré l'effectivité des modèles de connaissances et leur interprétabilité en langue, l'exploitation computationnelle ou calculatoire des modèles ne rend pas compte du contenu des expressions linguistiques des connaissances. En d'autres termes, les réécritures en langue permettant d'expliciter le contenu de ces connaissances ne sont pas reconstructibles dans le cadre de leur représentation syntaxico-formelle. Cependant, le système fait sens dans le domaine et assiste l'utilisateur dans son action dans le domaine et l'élaboration de ses connaissances. En effet, l'outil, par son effectivité informatique, permet d'exprimer en langue, c'est-à-dire au niveau où elles se donnent, des connaissances qui resteraient implicites voire inexistantes sinon. Le support informatique montre des rapports symboliques dont l'interprétabilité en langue correspond à la formulation de nouvelles connaissances dans le domaine, bien que celles-ci ne soient pas formalisables.

Nos propositions

Notre travail est construit autour de ce double constat : d'une part la non formalité des connaissances, c'est-à-dire la non validité du formalisme comme principe de modélisation, et d'autre part la constitution de connaissance dans le domaine à partir de l'effectivité du formel et de son caractère symbolique. Le point de départ de notre réflexion est que les systèmes formels automatiques sont des modèles non pas parce qu'ils axiomatisent le domaine et le dénotent, mais parce qu'ils sont des systèmes symboliques dynamiques de signifiants linguistiques commandant l'investissement sémantique de manière à en faire

TAB. 1.1 –La notion de signature dans les systèmes formels.

Un langage (formel) se définit à partir d'un alphabet A de symboles ; une suite finie de symboles constitue un *mot*, les mots forment l'ensemble A^* sur lequel est définie une opération binaire de concaténation. Cette opération, associative et admettant comme élément neutre le mot vide, définit sur A^* une structure de monoïde libre engendré par A. Tout partie de A^* est un *langage formel* défini sur A.
On appelle « signature » d'un langage formel un ensemble Σ muni d'une application $ar : \Sigma \mapsto N$.
La signature d'un langage du premier ordre, tel que ceux qui sont utilisés en IA pour constituer les systèmes formels représentant les connaissances, est l'ensemble Σ des symboles de fonction et de relation qui figurent dans ses formules. On distingue dans une signature du calcul des prédicats deux signatures disjointes, la signature fonctionnelle et la signature relationnelle.
Les symboles contenus dans les signatures fonctionnelles et relationnelles doivent, pour que le système formel possède une sémantique, être interprétés dans un domaine. Par exemple, la signature du langage de l'arithmétique est $\{0, s, +, x, <, =\}$ où $0, s, +, x$ sont les symboles fonctionnels et $=, <$ sont les symboles relationnels. Ces symboles sont ininterprétés dans le calcul symbolique mais peuvent être interprétés dans le cadre usuel de l'arithmétique et ainsi 0 renvoie à l'élément nul et + à l'addition entre des entiers naturels. Nous nous réfererons par la suite aux signatures fonctionnelles et relationnelles par le terme de « primitive non logique » dans la mesure où elles ne possèdent pas de signification logique comme un connecteur, mais doivent être interprétées dans un domaine pour qu'il y ait un sens résultant de la formule. Sur tous ces points, voir (Lalement, 1990).

un système de connaissances, c'est-à-dire un modèle. Nous proposons dans une première partie, une méthodologie qui exploite ce fait, l'*artéfacture*, et dans une seconde partie une philosophie qui essaie de comprendre ce fait, l'*herméneutique matérielle*.

Notre propos est en effet de proposer une autre méthodologie pour l'IA qui contourne l'obstacle que constitue la non formalité des connaissances à une approche formaliste tout en conservant les points forts de cette approche : son caractère symbolique et son caractère effectif. C'est à la mise en évidence de moyens linguistiques d'élaborer des connaissances à l'aide de représentations formelles que se consacre l'*artéfacture*, nom que nous donnons à l'intelligence artificielle conçue et réalisée dans ce cadre puisque l'*artefact* informatique construit doit produire un sens dans le domaine qui soit reconnu par les acteurs du domaine comme un modèle et donc assiste à la résolution du ou des problèmes posés. Notre première partie, qui précise les principes de cette méthodologie, l'illustre sur un problème important en IA aujourd'hui, les « ontologies du domaine »[8].

Mais, il ne suffit pas de constater le fait que l'artefact symbolique qu'est un système formel automatique fasse sens et produise du sens pour l'exploiter dans un programme méthodologique, il faut en outre s'efforcer de comprendre ce fait. C'est à cela que se consacre la deuxième partie de notre travail. Puisque le formalisme ne peut être une théorie de la connaissance, puisqu'il ne rend pas compte des connaissances exprimées dans la langue, il faut recourir à d'autres principes théoriques pour étudier ce que sont les connaissances et comprendre pourquoi leur formalisation peut produire de nouvelles connaissances quand cette formalisation est rendue effective sur un artefact. Il faut par conséquent comprendre

8. Une définition précise de ce terme, dans son acception IA est donnée dans la section 1-7.2.

le statut de ce qui est effectif et développer une philosophie de l'informatique. Puisqu'il ne peut être question d'adopter le formalisme comme philosophie pour comprendre le formalisme comme science de l'effectivité ou informatique, nous recourons à la philosophie kantienne pour mener cette réflexion qui nous amène à poser que l'informatique est une science de la nature, en un sens élargi. En d'autres termes, le calcul doit être considéré comme un objet « naturel », c'est-à-dire comme un objet « matériel » au sens où il appartient à ce que produit la nature.

D'autre part, il nous faut comprendre ce que signifie le sens pour un esprit humain qui en investit les représentations formelles qu'il considère. Nous adoptons dans ce but une approche phénoménologique. Tant l'approche kantienne de l'informatique que l'approche phénoménologique de la connaissance débouchent sur une conception de la connaissance comme constituée par la matière qui la véhicule : toute connaissance n'est connaissance que si elle est inscrite (en un sens à préciser) sur un support matériel qui la constitue. Une telle conception aboutit au projet d'une *herméneutique matérielle* dont nous ne faisons qu'esquisser quelques traits généraux. Ces deux thèses, (*i*) l'informatique est une science de la nature, (*ii*) toute connaissance repose sur la médiation d'une inscription matérielle, donnent ensemble la problématique de l'artéfacture : si le calcul est un objet matériel, on peut le considérer comme un support matériel dynamique d'inscription et donc on peut considérer les artefacts informatiques comme des supports dynamiques matériels d'inscription et étudier la constitution des connaissances qu'ils autorisent.

Le propos de ce chapitre d'introduction, avant d'entrer dans le détail des deux parties composant notre mémoire, est de fixer le contenu des notions que nous utilisons ici et de présenter quelles sont nos positions et propositions. Le plan adopté est le suivant : nous précisons tout d'abord ce que nous entendons par IA dans la section 1-2, puis ce que nous entendons par *formalisme* dans la section 1-3 pour préciser ensuite en quoi formalisme et IA peuvent s'unir en un programme de recherche cohérent (section 1-4). Nous revenons alors sur les deux constats autour desquels tournent notre méthodologie et notre réflexion : nous précisons tout d'abord quels problèmes rencontre ce programme (section 1-5) pour insister sur l'interprétabilité des systèmes qu'il réalise. Cette interprétabilité se théorise grâce à la notion d'« orthothéticité » que nous empruntons à B. Stiegler (Stiegler, 1996) qui nous permet d'expliciter l'importance du caractère *symbolique* des représentations formelles traitées en IA et de la nécessité de conserver ce caractère (section 1-6). Nos propositions consistent alors à montrer comment concevoir un programme de recherche, l'artéfacture (section 1-7), construit sur ce constat et à esquisser quelle philosophie du sens pourrait rendre compte de la pertinence d'un tel programme (section 1-8). Nous terminons cette introduction sur le plan du mémoire (section 1-9).

1 Un exemple

Considérons les phrases suivantes formulées dans la langue naturelle :

> Le logroton contient du chlortalidone ;
> la chlortalidone possède une fonction diurétique ;
> le logroton possède une fonction diurétique
> le logroton possède une fonction antihypertenseur

Une formalisation, c'est-à-dire une transcription dans un langage formel, donnerait à peu près ceci, en adoptant le formalisme des graphes conceptuels de Sowa (Sowa, 1984) (voir le chapitre 2 pour une définition complète) : :

```
[logroton]--(contient)-->[chlortalidone]
[chlortalidone]--(a-pour-fonction)-->[fonction-diuretique]
[logroton]--(a-pour-fonction)-->[fonction-diuretique]
```

L'avantage de ces graphes est d'une part d'être lisible spontanément du fait que les types des concepts correspondent à des signifiants linguistiques, par exemple `logrotron` dans le concept `[logrotron]` ou `contient` dans la relation `(contient)` . D'autre part, il est possible de calculer sur ces représentations puisqu'elles sont formelles : on peut ainsi déduire que

```
[fonction-diuretique]-
  (a-pour-fonction)<--[chlortalidone]
  (a-pour-fonction)<--[logroton]
```

en faisant la jointure entre les deux derniers graphes.

Mais il faut s'assurer que la sémantique opérationnelle induite par les possibilités calculatoires du formalisme soit conforme au sens que les notions possèdent dans le domaine. Or cette conformité n'est pas assurée par principe, dans la mesure où la sémantique de la langue n'est pas compositionnelle alors que l'est la sémantique des termes formels. Par conséquent, ce qui compris et attendu au niveau de la langue et de sa signification n'est pas forcément ce qui est calculé par le système. Ainsi, suivant le principe selon lequel le concept `[fonction-diuretique]` possède une signification non contextuelle, on a pu considérer d'un point de vue calculatoire qu'il était possible de fusionner les deux graphes où ce concept apparaît. Or, du point de vue du domaine, le résultat obtenu est faux. En effet, la fonction diurétique ne qualifie pas le même type de fonction selon qu'il s'agit d'une molécule biochimique comme la chlortalidone ou d'un médicament comme le logroton.

Dans le cas d'une molécule biochimique, la fonction correspond à l'effet biochimique produit par l'administration de la molécule. Dans le cas d'un médicament, la fonction correspond à l'indication thérapeutique, c'est-à-dire le problème que doit traiter l'administration du médicament. Ainsi, alors que le logroton a bien pour effet physiologique un effet diurétique, sa fonction en tant que médicament est d'être un hypertenseur.

Autrement dit, le concept de « fonction diurétique » possède une signification dépendante du contexte si bien que des calculs effectués sur sa transcription formelle, supposant sa non-contextualité, infèrent des résultats contredisant l'interprétation attendue au niveau du domaine. Il faut, pour éviter ce problème, distinguer la fonction biochimique des molécules de la fonction thérapeutique des médicaments. On dira ainsi que la notion de « fonction » est ambiguë et que l'on retient « fonction biochimique » pour ne qualifier que les molécules et « fonction thérapeutique » pour ne qualifier que les médicaments.

On écrirait alors les graphes suivants :

```
[logroton]-
  (contient)-->[chlortalidone]-
    (a_pour_fonction_biochimique)-->[fonction_diurétique]%
  (a_pour_fonction_therapeutique)-->[hypertenseur]
  (a_pour_effet_physiologique)-->[fonction_diurétique]%
```

A ce niveau, il devient possible de composer les concepts tout en produisant une signification pertinente dans le domaine. Par exemple, le logroton possède un autre principe actif, le métoprolol, qui est un bêta-bloquant. Par conséquent, si l'on nous dit que « le métoprolol a pour fonction d'être un bêta-bloquant » et que « le logroton a une fonction bêta-bloquante », la clarification sémantique effectuée ci-dessus permet de poser que

```
[metoprolol]--(a_pour_fonction_biochimique)-->[beta-bloquant]
```

et d'inférer que

```
[logroton]-
  (a_pour_effet_physiologique)-->[beta-bloquant]
  (a_pour_effet_physiologique)-->[fonction_diuretique].
```

On peut remarquer que distinguer fonction biochimique et fonction thérapeutique revient à interdire l'usage d'unités linguistiques dont le sens serait modifié par le contexte. En d'autres termes, le travail de clarification sémantique consisterait à se ramener aux hypothèses nécessaires à une formalisation des connaissances. Et c'est effectivement ainsi qu'il faut procéder : du fait qu'il est nécessaire d'avoir des représentations symboliques effectives dont le comportement computationnel reflète la réalité du domaine, il est nécessaire de se plier aux conditions de possibilité d'une opérationnalisation computationnelle. Seulement, il faut pour cela effectuer un travail de modélisation explicite où l'on dégage de l'expression linguistique des connaissances des notions dont le contenu sémantique est déterminé de telle manière qu'elles puissent être univoques et non contextuelles, c'est-à-dire exploitables informatiquement.

En d'autres termes, il est nécessaire de faire reposer la construction formelle sur un travail de désambiguisation linguistique. Le problème est de savoir comment l'effectuer. Il n'est pas en effet possible de considérer par exemple les objets du domaine indépendamment de leur description linguistique pour constituer des représentations formelles qui les dénotent. Pour être capable de systématiser les expressions linguistiques, il est nécessaire de s'en remettre au sens qu'elles véhiculent et de construire un système de significations dans lequel les notions du domaine sont déterminées de manière univoque et non contextuelle. Le chapitre 4 montre en détail comment effectuer une telle construction et montre à quel prix cela est possible : la décontextualisation ne peut se faire qu'en fixant un contexte particulier si bien que désambiguer n'est pas représenter tous les sens possibles, mais n'en retenir qu'un seul.

L'enjeu d'une méthodologie comme l'artéfacture, qui constate que les modèles conçus en IA doivent leur capacité de modélisation à leur emprunt aux expressions des connaissances du domaine de signifiants linguistiques, est de donner un sens non contextuel et univoque à ces signifiants pour qu'il soit possible que leur manipulation formelle et computationnelle produise du sens dans le domaine. Notre travail sur les « ontologies » consiste précisément à préciser comment donner de telles significations aux signifiants retenus dans la signature du formalisme.

Notre propos n'est donc pas d'interdire la formalisation, mais au contraire de la rendre possible alors que le domaine que l'on traite n'est pas formel. En d'autres termes, la formalisation n'est qu'une étape intermédiaire du processus consistant à construire un modèle opérationnel qui consiste à « mathématiser » (au sens du calcul formel) une description linguistique du domaine.

Alors que cette étape est laissée implicite habituellement (mais elle est obligatoirement faite), puisque rationaliser l'expression en langue naturelle des connaissances pour en faire un modèle consiste justement à construire un modèle formel dénotant un univers de référents, notre propos est d'expliciter cette étape et d'y voir l'*essentiel* de l'effort de modélisation. Dans la mesure où le monde n'est pas un univers de référents, il est nécessaire de le modéliser comme tel au préalable si l'on veut faire fonctionner la représentation formelle comme modèle.

Ce petit exemple permet d'illustrer les points suivants :

- la représentation formelle des connaissances effectuée ici modélise la réalité du domaine non pas parce que l'on a considéré les objets du domaine en eux-mêmes pour y voir la dénotation des expressions symboliques utilisées dans la représentation, mais parce que l'on a utilisé des symboles correspondant à des signifiants et que l'on s'en est remis à la langue pour assurer la médiation entre la représentation et le monde ;
- la manipulation formelle des représentations, compositionnelle, ne respecte pas obligatoirement la manière dont les signifiants linguistiques s'unissent pour composer de nouvelles significations ;
- le calcul symbolique, en introduisant une systématicité computationnelle entre les représentations formelles permet de mettre en évidence les phénomènes sémantiques du domaine qui, contredisant les présupposés du formalisme, doivent être éclaircis pour que les représentations formelles puissent être considérées comme un modèle du domaine ; c'est en effet le calcul sur les graphes qui rend nécessaire de distinguer entre fonction biochimique et fonction thérapeuthique ; à ce niveau, le formalisme est un outil puissant d'aide à la modélisation du domaine ; cette modélisation s'effectue au niveau linguistique de l'expression des connaissances du domaine et consiste à donner une signification univoque et non contextuelle aux signifiants linguistiques retenus pour constituer la signature du formalisme de représentation des connaissances. Ainsi, il est possible d'obtenir une composition sémantique des représentations formelles qui fasse sens dans le domaine.

On retrouve ainsi les trois termes de nos constats négatifs et positifs : des modèles symboliques non référentiels interprétables en langue et effectifs. C'est la signifiance linguistique conjuguée à l'effectivité du symbolisme qui permet de faire de l'outil informatique une aide à la modélisation, c'est-à-dire à la constitution de connaissances.

2 L'INTELLIGENCE ARTIFICIELLE

Dans le cadre de notre travail, nous retenons comme caractérisation de l'intelligence artificielle[9] le projet technologique visant à la résolution informatique de problèmes ou l'automatisation informatique de tâches pour lesquels on dispose de connaissances exprimées en langue naturelle. Selon cette définition, relèvent d'une approche IA les problèmes de diagnostic, de raisonnement naturel, de traitement automatique de la langue, etc., c'est-à-dire tous les problèmes pour lesquels on ne dispose pas de théories dites scientifiques. Une telle caractérisation de l'IA n'est pas neutre ni évidente et doit être justifiée.

En effet, elle tranche sur les définitions que l'on peut parfois rencontrer qui voient dans l'IA le projet de doter une machine des facultés cognitives qui lui permettent d'adopter un comportement correspondant à ce qui est habituellement nommé intelligence. John Haugeland, philosophe ayant travaillé sur les aspects philosophiques des fondements de l'IA affirme ainsi que l'objectif poursuivi est de développer des machines « intelligentes », au sens fort du mot, c'est-à-dire « dotées d'esprit » :

> « L'« IA » veut de l'authentique, des machines dotées d'esprits, au sens le plus littéral de l'expression. » (Haugeland, 1985, p. 8)
>
> il poursuit en note :
>
> « Peut-être devrait-on appeler l'intelligence artificielle « intelligence synthétique » pour être plus près du langage industriel. En effet, les diamants artificiels ne sont que des pâles imitations, alors que les diamants de synthèse sont de véritables diamants, mais fabriqués et non extraits du sol (ou bien prenez un parfum artificiel à la vanille par opposition, par exemple, à de l'insuline synthétique). En dépit de son nom, l'IA veut atteindre la véritable intelligence et non une pâle imitation de celle-ci. » (Haugeland, 1985, p. 8 note)

Selon Haugeland, l'intelligence naturelle est « naturelle » dans la mesure où des systèmes concrets, les êtres humains, manifestent la propriété naturelle d'être intelligent, c'est-à-dire d'avoir un esprit. L'intelligence artificielle a pour tâche de trouver quels principes sont à la base de ce phénomène et d'être capable de construire des systèmes qui manifesteraient, non pas une simulation de l'intelligence, mais une véritable intelligence. En d'autres termes, l'intelligence humaine n'est qu'un cas particulier de sytèmes intelligents, et l'IA est la science de l'intelligence générale ou la science générale des systèmes intelligents, de la même manière que la physique de l'électricité est la science des phénomènes électriques en général, qu'ils soient produits naturellement dans la nature ou artificiellement dans les centrales. L'intelligence est le phénomène qui se produit quand on réalise des systèmes formels automatiques (Haugeland, 1981).

Si l'on se tourne vers les praticiens de l'IA, on retrouve une conception analogue selon laquelle l'objectif poursuivi est la construction de système « intelligents ». Une figure importante de la communauté IA, Roger Shank, affirme ainsi qu'« il y a deux objectifs principaux en IA sur lesquels la plupart des praticiens s'accorderaient. Premièrement, et principalement, le but de l'IA est de construire une machine intelligente. Et, deuxièmement, son but est de découvrir la nature de l'intelligence. » (Shank, 1990, p.4). Mais comme le remarque Shank, le projet de l'IA ne se comprend que si la notion d'« intelligence » est

9. On peut consulter (Ganascia, 1990; Ganascia, 1993) pour des introductions claires et bien documentées à cette discipline et aux enjeux et problèmes qu'elle suscite.

précisée. Or, et c'est là que le bât blesse, il n'y que peu de consensus sur la nature de l'intelligence. Mais, plutôt que rechercher une définition unique de l'intelligence pour asseoir le projet de l'IA, il vaut mieux, remarque Shank, énumérer quelles capacités ou fonctions semblent indispensables à un système quelconque pour qu'il puisse être dit intelligent.

Ainsi, le projet de construire un système intelligent consiste dans le projet de réaliser un système capable de faire preuve d'intelligence, c'est-à-dire d'être capable de réaliser des tâches pour lesquels, quand elles sont réalisées par un humain, elles sont dites requérir de l'intelligence. Ces tâches sont des tâches dites « cognitives ». De telles tâches sont les tâches pour lesquelles il faut raisonner, mémoriser, apprendre, se construire une représentation, etc.

Newell, l'un des pères fondateurs de l'IA, ne dit pas autre chose. Il considère en effet que « the ultimate goal of work in cognitive architecture is to provide the foundation for a system capable of general intelligent behaviour » . Le but est en effet de mettre au point une architecture permettant à un système d'accomplir l'ensemble des tâches cognitives (« the full range of cognitive tasks ») en utilisant pour cela les méthodes de résolution de problèmes et de représentation des connaissances. Si bien que le système SOAR développé par Newell et son équipe est présenté comme « une architecture pour un système devant être capable d'une intelligence générale » (*an architecture for a system that is to be capable of a general intelligence* (Laird *et al.*, 1987, p.2)).

Notre caractérisation tranche donc sur les définitions rappelées ci-dessus dans la mesure où elle ne fait pas appel à la notion d'intelligence ou de cognition, encore moins de faculté cognitive. En effet, notre définition de l'IA ne relève pas d'un projet scientifique visant à construire des machines exhibant un comportement intelligent, en cherchant par là-même à comprendre la nature de l'intelligence, mais d'un projet technologique et pratique visant à traiter une classe de problèmes restant rebelles aux techniques classiques de résolution. Sans y être identique, une telle caractérisation est par conséquent proche de celle de Jean-Louis Laurière (Laurière, 1987, p.1) qui considère que l'objet même de l'IA, consiste dans le traitement de « tout problème pour lequel aucune solution algorithmique n'est connue ».[10].

Cependant, l'écart peut ne pas être si grand: notre définition de l'IA est neutre par rapport aux recherches sur l'esprit ; si elle ne détermine pas leur nature, elle ne les interdit pas. On peut s'en rendre compte si l'on se souvient que la langue naturelle peut se considérer selon un double point de vue. Les expressions de langue sont en effet l'expression première et immédiate de l'esprit : l'étude de la langue renvoie à l'étude de la pensée et de l'esprit (Auroux, 1996). Mais, d'un autre côté, la langue permet de parler de quelque chose et d'exprimer notre rapport au monde. En d'autres termes, les connaissances que la langue exprime peuvent être soit des connaissances sur le sujet exprimant soit des connaissances sur l'objet dont on parle.

On retrouve une telle distinction dans l'évolution des paradigmes dans l'acquisition des connaissances, étape méthodologique nécessaire à la réalisation des SBC consistant à définir quelles connaissances un système doit avoir pour résoudre un problème, et à définir

10. On pourrait croire que ces caractérisations se distinguent dans la mesure où des problèmes pour lesquels il existe des théories formelles non algorithmiques appartiennent au domaine de l'IA selon J.-L. Laurière alors qu'ils en seraient exclus selon nous. Mais, pour résoudre informatiquement ces problèmes formels, on use d'heuristiques qui sont exprimées en langue naturelle. Par conséquent, ces deux caractérisations se rejoignent en définissant l'IA par la classe des problèmes qu'elle aborde plutôt que par la recherche de la nature de l'intelligence.

comment les lui donner. Dans une première époque, que l'on peut appeler époque « psychologique », l'acquisition des connaissances se comprend comme un « transfert d'expertise » . Les spécialistes ou experts possèdent un modèle mental du domaine leur permettant de résoudre les problèmes rencontrés dont il faut doter le système. Le SBC réplique alors la démarche psychologique de l'expert en opérationnalisant son modèle mental du monde.

Le principe de base de ce courant repose sur l'idée intuitive qu'il est possible d'écrire directement dans un programme les connaissances du domaine exprimées en langue naturelle. Hérité des premiers systèmes experts[11], ce principe stipule que les règles de production fournissent un moyen naturel d'expression des connaissances de l'expert. L'objectif d'un processus d'acquisition des connaissances consiste alors dans ce contexte à permettre à un expert d'écrire ses connaissances et à maintenir lui-même la base ainsi construite.

En effet, dès que l'on entreprend de construire un système qui résolve des problèmes pratiques, concrets et réels, c'est-à-dire des problèmes qui ne relèvent pas de cas d'école ou d'exemples jouets tirés de la littérature, comme par exemple le diagnostic médical, la conduite de haut fourneau, etc., plutôt que les tours de Hanoï, le jeu d'échecs, le tic-tac-toe, etc., il est nécessaire, pour déterminer la nature du problème, de collecter les connaissances dont on dispose. Or, toutes les approches adoptées pour mener à bien cette acquisition des connaissances sont confrontées au problème de l'expression des connaissances en langue naturelle. Cela ne signifie pas que ces approches mobilisent des théories linguistiques spécifiques à l'étude de la langue, mais simplement que le matériau à partir duquel mener l'acquisition se présente à travers la langue. Comme le notent Wielinga et Breuker (Wielinga & Breuker, 1984) dans un travail qui allait devenir la méthodologie KADS (que nous présentons avec quelques détails dans le chapitre 2),

> « le problème majeur dans le développement des systèmes ("experts") à base de connaissances est l'acquisition des connaissances dans les domaines d'expertise. La méthode la plus usitée pour une telle acquisition des connaissances est l'interprétation des données verbales, généralement obtenue pendant des entretiens avec l'expert, dans les termes d'un formalisme adapté à l'implémentation (*e.g.* les règles de production). »

L'effort est par conséquent d'être capable, à partir des connaissances du domaine que l'on trouve exprimées dans la langue naturelle, d'obtenir une contrepartie formelle et opérationnelle de ces connaissances pour les exploiter informatiquement.

Le système exemplaire de cette approche est le système TEIRESIAS, développé par Davis (Davis, 1979). Le principe de ce programme est de proposer une *traduction simplifiée des règles en langue naturelle* pour faciliter l'édition, la modification et l'ajout de nouvelles règles dans la base. Ce système est construit autour du principe que l'expert analyse les conclusions erronées du programme à l'aide de ce dernier et modifie ainsi sa base de règles.

Un expert formule ses connaissances sous la forme de phrases ou énoncés de la langue naturelle. Les mots de la phrase sont tout d'abord réduits à une forme canonique (*e.g.* son lemme) pour être ensuite associés à des « connotations » possibles. Une connotation

11. Rappelons qu'un système expert comprend une base de règles de production contenant les connaissances du domaine, une base de faits contenant les données d'un problème à traiter, un moteur d'inférence contenant un principe de raisonnement, le *modus ponens* dans le cas du « chaînage avant » ou le *modus tollens* dans le cas du « chaînage arrière », permettant de déduire de nouveaux faits à partir des données du problème. Une règle de production se définit selon un schéma SI CECI, ALORS CELA. Mycin (Shortliffe & Buchanan, 1975) constitue le système emblématique de ce courant.

The	client is following	up	on	market	trends	carefully
	↓	↓		↓	↓	
	predicate function	value		object	attribute	
	↓					
	SAME					

FIG. 1.4 –*L'association des mots de la langue à des primitives d'un langage de représentation des connaissances dans* TEIRESIAS

correspond à une primitive conceptuelle du langage formel permettant de construire les règles. Ces primitives peuvent donc être des prédicats, attributs, objets ou valeurs. Une forme est associée à chaque prédicat qui spécifie l'ordre et le type des arguments du prédicat. Une procédure stipulant comment parcourir la liste des connotations pour trouver un item du type approprié est associée à chaque type possible d'argument. La figure 1.4 montre comment le groupe verbal de la phrase citée connote un prédicat, le prédicat SAME qui a pour forme OBJECT ATTRIBUTE VALUE. En parcourant les connotations de la phrase, ce prédicat se détermine dans la fonction LISP (SAME OBJECT TREND UPWARD), en retenant le terme TREND pour l'attribut, le terme UP pour la valeur, et le terme MARKET pour l'objet, mais la procédure d'association renvoie alors la valeur OBJECT pour que le prédicat produit reste générique.

Il s'agit ainsi, dans TEIRESIAS, d'appliquer à l'expression linguistique des connaissances des formes (*patterns*) logiques pour en dégager le contenu. Il est notable que ce passage s'effectue très simplement et ne relève pas d'un traitement sophistiqué de la langue :

> « Natural language has not been a major focus of this work, and for the most part the simplest techniques have been used that would support the level of performance required. For instance, all questions and responses from TEIRESIAS are either preformed, or manufactured by filling in templates with appropriate words [...]. Where answers are free text (rather than multiple choice), the system relies primarily on the keyword- and template-oriented techniques. » (Davis, 1979, p.128)

C'est qu'en effet la signification des unités linguistiques consiste dans les prédicats logiques qu'elles « connotent » si bien que des techniques aussi sommaires que la recherche de mots clefs ou d'instanciation de schémas formels suffisent à passer de la langue au formalisme du programme.

Ce travail tout à fait remarquable montre cependant rapidement les limites d'une telle approche dans la mesure où les règles de production constituent une technique de programmation et non un langage de représentation. En effet, si les règles de production sont bien un langage de programmation dans la mesure où elles autorisent une calculabilité identique à une machine de Turing universelle, elles ne possèdent pas l'expressivité nécessaire à la transcription de connaissances structurées, comme on peut les rencontrer dans les domaines d'expertise. Par exemple, les règles ne permettent pas de décrire les connaissances dites « profondes », c'est-à-dire les connaissances correspondant aux théories du domaine et à leur compréhension systématique (comme par exemple une théorie causale du domaine). Les règles de production ne permettent pas de représenter un *système* de connaissances dans la mesure où elles sont par définition des *atomes* de connaissance manipulés indépendamment les uns des autres. Pour conserver ce que le formalisme des règles de production peut avoir d'efficace et d'agréable à utiliser, on a conçu des systèmes de représentations de connaissances profondes (la plupart du temps, des modèles causaux) qui étaient par la suite « traduites » en des connaissances de surface (Charlet, 1992) formulées en règles de production.

Cela signifie que, si l'on peut potentiellement tout faire d'un point de vue calculatoire, l'utilisateur expert n'en retrouve pas pour autant les structures correspondant à ses connaissances si bien qu'il se retrouve incapable de les exprimer. Enfin, cette approche montra qu'il était impossible de séparer une connaissance de son utilisation, *i.e.* de maintenir une stricte séparation entre la base de connaissance et le moteur d'inférence, des règles de la première pouvant contenir des instructions de contrôle pour le second.

TEIRESIAS avait dégagé l'utilité de disposer de modèles de règles, comme on l'a vu plus haut, qui permettent d'interpréter ce que l'expert exprime. Mais la nécessité de fonder les inférences du système expert sur des connaissances profondes (Keravnou *et al.*, 1992) du domaine comme les modèles causaux (Charlet *et al.*, 1996) a suscité un nouveau courant pour la conception des SBC, la modélisation constructive.

En effet, dans une seconde époque, que l'on peut appeler époque « constructiviste », l'acquisition se comprend comme une « construction de modèle ». Les experts et les spécialistes fournissent un matériau linguistique reflétant leur introspection, les données du problème, l'objectif poursuivi, etc., et la tâche de l'acquisition des connaissances consiste alors à *construire* un modèle à partir de ce matériau. Le résultat obtenu n'est donc pas un modèle qui rendrait compte du savoir des experts consultés, mais un modèle qui expose une théorie du domaine et du problème du point de vue du système devant traiter ce problème. Il ne s'agit plus de retrouver le modèle de l'expert à partir de la verbalisation qu'il donne de ses connaissances et de sa démarche de résolution de problème, pour ensuite le répliquer en l'opérationnalisant sur une machine, mais de construire un modèle original pour les besoins de la réalisation d'un SBC. Ainsi, on peut lire :

> « Knowledge acquisition is increasingly considered to be a *model-based* activity. This is in contradistinction to the traditional "extract-transfer-and-simulate" view, in which knowledge acquisition is seen as extracting knowledge from a human expert, transferring or translating this into a knowledge representation, and mimicking or simulating the human problem solving in a knowledge-based system. In the model-based approach to knowledge acquisition, engineer-type model are constructed that relate to various (restricted) aspects of interest of problem-solving and that are able to perform certain problem-solving functions. » (Akkermans *et al.*, 1993)

propos que confirme Linster (Linster, 1993) :

> « Early work in knowledge acquisition for knowledge-based systems emphasized a transfer view of knowledge acquisition. Knowledge was considered to be in an expert's head, and the knowledge engineer, using a tool, elicits the knowledge and represents it in a suitable operational formalism. Today, we see knowledge acquisition as a constructive model-building process. The interaction between domain expert, knowledge engineer, and tool creates knowledge. »

On peut analyser ce changement de perspective à l'aide de la distinction qu'établit Norman entre les modèles mentaux que des spécialistes se forgent à propos du sytème qu'ils utilisent ou doivent concevoir, et les modèles conceptuels qui constituent une représentation abstraite du système que l'on peut communiquer. Alors que le modèle mental se caractérise par le fait qu'il permet à l'expert de comprendre et d'anticiper le comportement du système, le modèle conceptuel est inventé pour fournir une représentation précise, consistante et complète du système que l'on peut enseigner, transmettre, appliquer :

> « Conceptual models are devised as tools for the understanding or teaching of physical systems. Mental models are what people really have in their heads and what guides their use of things. » (Norman, 1983, p.12)

Mais, en pratique, la notion de modèle mental est problématique. En effet, Norman insiste sur le fait que si, « Ideally, there ought to be a direct and simple relationship between the conceptual and the mental model », il s'avère qu'en réalité, les modèles mentaux sont incomplets, au pouvoir prédictif limité, instables, non spécifiques (différents systèmes peuvent donner lieu à des modèles mentaux possédant d'importants recouvrements), etc. En d'autres termes, ils sont de mauvais candidats à l'élaboration de modèle conceptuel. De plus, les modèles mentaux, privés par définition, sont difficile à mettre en évidence :

> « discovering what a person's mental model is like is not easily accomplished. For example, you cannot simply go up to the person and ask. Verbal protocols taken while the person does a task will be informative, but incomplete. Moreover, they may yield erroneous information, for people may state (and actually believe) that they believe one thing, but act in quite a different manner. All of a person's belief structures are not available to inspection, especially when some of those beliefs may be of a procedural nature. Finally, there are problems with what is called the "demand structure" of the situation. If you ask people why and how they have done something, they are apt to feel compelled to give a reason, even if they did not have one prior to your question. They are apt to tell you what they believe you want to hear (using their mental models of your expectations). Having then generated a reason for you, they may then believe it themselves, even though it was generated on the spot to answer your question. On-line protocols generated while in the act of problem solving and that give descriptions of activities rather than explanations are much more reliable » (Norman, 1983, p.11, note)

Comme le remarque (Gaines *et al.*, 1993), il vaut mieux considérer que le transfert d'expertise correspond, en fait, à l'élaboration d'un modèle conceptuel que le SBC opérationnaliserait, plutôt qu'à la conceptualisation du modèle mental de l'expert. Il s'agit, dans ce changement de perspective, de passer d'une modélisation de l'expert et de son fonctionnement cognitif à une modélisation du système et de son fonctionnement effectif.

Par conséquent, caractériser le fait d'acquérir des connaissances à partir de leur expression en langue naturelle permet de rendre compte de la pratique effective, passée et actuelle, mise en œuvre dans la réalisation de SBC. Mais, on voit que les enjeux ne sont pas les mêmes selon le point de vue adoptée sur ces connaissances : expressions d'un modèle mental de l'expert, ou expressions à partir desquelles élaborer un modèle conceptuel du système à construire. Dans le premier cas, on s'intéresse à une science de l'intelligence et de l'esprit dont la détermination des principes permet d'en construire d'artificiels. Dans le second cas, l'IA modélise le monde tel qu'il est visé par les connaissances exprimées en langue pour être capable d'y agir. Alors qu'une IA psychologique est davantage concernée par une psychologie scientifique, l'IA constructiviste est davantage concernée par l'étude scientifique du monde où prennent leur valeur et leur signification les connaissances exprimées en langue.

La majorité des travaux actuels concernant l'élaboration des SBC s'inscrit dans le cadre d'une IA constructiviste. C'est donc cette dernière qui sera l'objet essentiel de notre travail et de notre critique, même si parfois, des excursions historiques dans l'IA psychologique seront nécessaires pour éclairer les notions utilisées. Notre critique consistera à adopter, tout comme l'IA constructiviste, le point de vue selon lequel l'acquisition des connaissances et la conception des SBC est une activité de construction de modèle. Mais, contrairement à l'IA constructiviste, notre position consistera à abandonner le formalisme de ces modèles, c'est-à-dire à abandonner la thèse selon laquelle ces modèles sont des modèles parce qu'ils formalisent les connaissances et dénotent le monde. En résumé, il existe trois conceptions possibles qui émergent de la caractérisation de l'IA comme la résolution de problème par l'utilisation de connaissances exprimées en langue naturelle ; elles se distinguent d'une part par l'objet qu'elles modélisent et d'autre part par le statut qu'elles confèrent au système opérationnel obtenu :

le système comme théorie cognitive des agents : opérationnaliser les connaissances linguistiques revient à développer une théorie scientifique de la manière dont un agent peut utiliser et exploiter ces connaissances et disposer ainsi du savoir qu'elles véhiculent ; c'est une théorie sur le fait que les agents humains savent traiter un problème à l'aide de ces connaissances et qu'elles expriment par conséquent le modèle mental qui leur permet de comprendre et d'anticiper le comportement du système ; en répliquant ce modèle en le formalisant et l'opérationnalisant sur un ordinateur, on obtient un système capable de traiter le problème ; cette approche consiste à voir dans le système informatique une théorie psychologique des agents où c'est l'expert, et non le problème qui est modélisé ;

le système comme théorie formelle du domaine : les connaissances linguistiques sont des connaissances qui n'expriment pas les processus mentaux de l'expert mais un savoir sur le monde ; elles permettent de construire un modèle du domaine et du problème que le système réalisé opérationnalise ; ce système possède le même statut que l'opérationnalisation de théories scientifiques comme les résolutions numériques des théories mathématiques de systèmes physiques à ceci près qu'il n'est pas considéré comme une simple opérationnalisation mais comme un véritable modèle du domaine dans la mesure où, formel, il le dénote ;

le système comme théorie sémiotique du domaine : le système informatique, formel puisque computationnel, ne dénote pas le domaine ; mobilisant des primitives non logiques exprimées dans un symbolisme emprunté à la langue, le système reçoit une interprétation linguistique ; le but de la modélisation sémiotique du domaine est de contrôler l'investissement linguistique du système formel de manière à ce que le sens produit corresponde à la réalité du domaine ; il s'agit de construire un système formel dont l'interprétation linguistique permette d'y voir un modèle conceptuel du domaine et du problème ;

Les deux premières approches sont formelles, les deux dernières modélisent le domaine. Notre proposition, qui coïncide avec la troisième approche, consiste donc à refuser, à l'instar de la seconde approche, le psychologisme de la première, et à nous démarquer de son formalisme comme principe de modélisation.

Toutes ces approches ont en commun de vouloir opérer au « niveau des connaissances ». Ce concept a été proposé par Newell (Newell, 1982) pour dépasser la cacophonie qui régnait à l'époque en matière de formalismes de représentation des connaissances, où aucun consensus ne semblait émerger. Il s'agissait de distinguer les notions de « représentations » et de « connaissances » en opposant le niveau des programmes ou des symboles où des représentations formelles sont effectives et le niveau des connaissances, correspondant au savoir véhiculé et implémenté par les programmes. Le fait que le niveau des connaissances soit un « niveau » signifie que le savoir véhiculé par le programme est en fait indépendant du programme dans la mesure où la connaissance peut être indépendante de sa représentation :

> « There exists a distinct computer systems level, lying immediately above the symbol level, which is characterized by knowledge as the medium and the principle of rationality as the law of behavior. » (Newell, 1982, p.25)

Un système décrit au niveau des connaissances est compris comme étant un agent possédant des buts à atteindre, des croyances sur le mondes, et des actions qu'il peut entreprendre. Son comportement peut être prédit grâce à au « principe de rationalité » qui stipule qu'un agent entreprend l'action qui, au vu de ses croyances sur le monde, lui permet d'atteindre ses objectifs.

Le problème d'une telle notion est que l'on se demande ce peut être une connaissance si elle n'est pas une représentation, même s'il est évident qu'il est impossible de l'y réduire. La question est : où est le niveau des connaissances ? Pour Newell, le niveau des connaissances correspond aux connaissances que l'on attribue à un système pour expliquer et anticiper son comportement. Mais cela déplace le problème sans le résoudre. Que le niveau des connaissances correspondent à un niveau d'observation et d'explicitation du comportement ne dit pas en quoi il consiste, ni par quoi il se concrétise.

Notre thèse (Bachimont & Collinot, 1993) est que le niveau des connaissances correspond à la description en langue naturelle du comportement des systèmes. Cette thèse est compatible avec la définition que donne Newell du niveau des connaissances sans s'y réduire. En effet, décrire en langue ce que fait un système revient à le catégoriser selon le sens commun où l'on attribue des qualités comme les croyances et les buts pour anticiper le comportement des objets et systèmes autour de soi ; le principe de rationalité est un principe de « bon sens » que le sens commun met en œuvre pour associer buts et croyances. Il s'agit d'une version psychologique.

Mais la description en langue ne permet pas seulement de définir des états de croyances et de buts que l'on attribue à un système pour en faire un agent, mais aussi de définir des connaissances sur le monde que le système opérationnalise. C'est comme on l'a vu, l'attitude de l'IA constructiviste.

Notre thèse sur le niveau des connaissances est par conséquent neutre sur l'interprétation qu'il faut en donner, psychologique ou constructiviste, selon que l'on a pour projet de construire des systèmes cognitifs ou experts ou des systèmes à base de connaissances. De toute manière, le problème de fond reste toujours le même : comment se confronter au donné linguistique par lequel les connaissances, expressions d'états mentaux ou d'états du monde, se révèlent.

Notre point de vue dans ce travail est celui de la résolution de problème non d'un point de vue psychologique mais d'un point de vue d'ingénierie. Notre interprétation du niveau des connaissances n'est donc pas qu'il s'agit d'états « mentaux » attribués à un agent, mais de connaissances sur le monde.

Par conséquent, notre caractérisation de l'IA depuis le point de vue de l'expression linguistique des connaissances permet de rendre compte d'une part des travaux effectivement réalisés en IA, et d'autre part de rester neutre quant aux statut cognitif et philosophique qui leur est accordé. Enfin, cette caractérisation est compatible, comme nous venons de le voir, avec les affirmations méthodologiques courantes en IA selon lesquelles l'IA constructiviste construit des modèles au niveau des connaissances (voir le chapitre 2 pour des détails). Il s'agit à présent de considérer le type de savoir que véhiculent des connaissances exprimées en langue et le type de problème que cela pose.

2.1 Les connaissances exprimées en langue naturelle

Le projet même de vouloir opérationnaliser les connaissances exprimées en langue naturelle leur reconnaît une dignité épistémologique que parfois l'on réserve aux seules connaissances dites scientifiques parce qu'établies selon les méthodes expérimentales de la science moderne et dans le langage formel qu'elle adopte. Il faut donc accorder que ces connaissances décrivent un monde dont il faut reconnaître, d'une manière ou d'une autre, la réalité. Il est nécessaire dans ces conditions de caractériser le type de savoir véhiculé par ces connaissances ainsi que le type de la réalité visée.

De manière générale, on peut dire que la langue naturelle nous permet de rapporter le savoir que nous avons de notre rapport au monde, c'est-à-dire de nos actions dans le monde. Grossièrement, pour fixer les idées, on peut dire que la langue naturelle véhicule un savoir sur nos pratiques, qu'elles expriment un savoir pratique, et on peut les distinguer alors des langages artificiels qui sont davantage conçus pour véhiculer un savoir sur les objets du monde et qui expriment plutôt un savoir théorique.

La langue naturelle nous permet de rendre compte de la manière dont le monde se présente spontanément à nous, dans notre vie quotidienne, dans nos pratiques. Ce rapport au monde englobe le sens commun, thème activement étudié en IA (voir notamment (Hobbs & Moore, 1985; Hobbs, 1995), (Lenat & Guha, 1990), (Kathleen, 1995)). Mais il peut être explicitement thématisé et rationalisé pour s'élaborer en art technique, au sens où par exemple la médecine est souvent présentée davantage comme un art que comme un science. Ces pratiques ont alors bien peu de rapport avec le sens commun au sens habituel du terme, et sont incompréhensibles à qui n'est pas de la partie. Mais, la langue naturelle permet dans ce contexte d'exprimer ce qui est le « sens commun du domaine », c'est-à-dire ce qui est compris et su par tout acteur du domaine dès lors qu'il maîtrise la langue du domaine et de la discipline.

Si on considère par exemple les dictionnaires encyclopédiques médicaux comme le Dorland (Dorland, 1994), ou le Harrison (Harrison, 1992) pour la médecine interne, on y trouve un système de connaissances usant de la langue naturelle comme seul moyen possible d'expression et exprimant un savoir sur les objets du monde et sur les actions possibles dont ils peuvent faire l'objet. En effet, ces ouvrages, destinés à transmettre des connaissances, ne font appel, comme seul médium d'expression, qu'à la langue naturelle.

La connaissance formulée en langue exprime dans ce cadre un savoir sur des objets du domaine en tant qu'ils constituent un objet pour des actions possibles. La connaissance, dans ce contexte, correspond à une action possible, et l'objet du domaine n'existe que comme corrélat de l'action possible. C'est d'ailleurs le fait que ces connaissances renvoient à des pratiques qui les rend incontournables et qui rend crucial l'enjeu de les opérationnaliser dès lors que l'on veut automatiser ces pratiques. C'est ainsi que des formalismes de représentation des connaissances se donnent pour objectif explicite de représenter le contenu d'énoncés linguistiques. C'est notamment le projet des graphes conceptuels de John Sowa (Sowa, 1984) :

> « The structure of a knowledge representation language depends critically on its ultimate goal. For conceptual graphs, the goal is a system of logic that can express the propositional content of sentences in natural language in as simple and direct a manner as possible. » (Sowa, 1991, p. 157)

Une autre manière de caractériser ces connaissances exprimées en langue naturelle est de remarquer qu'elles décrivent l'interaction entre des systèmes. Cette caractérisation est la même que la précédente dans la mesure où une pratique est l'interaction entre un agent et son environnement. Clancey (Clancey, 1993) insiste particulièrement sur ce point et note que les connaissances comme les classifications ou les taxinomies ne sont pas réductibles à des connaissances scientifiques consistant dans la théorie causale d'un système. En effet, les classifications ou taxinomies, c'est-à-dire des connaissances reposant sur le pouvoir de catégorisation véhiculé par la langue naturelle, décrivent non pas la nature physique d'un système mais ses interactions avec son environnement.

> « Many papers in the literature suggest that classification models are inferior to simulation models and can be entirely reduced to or compile from them. According to this point of view, physicians talk in terms of syndromes and disease classifications because they do not understand the causal mechanisms causing these processes. A real model would reduce disease descriptions to descriptions of physical structure and function. For the most part these assumptions are false and belie a fundamental misunderstanding about the nature of system modeling and, more generally, how systems interact.
>
> « Disease descriptions characterize *the result of recurrent interaction* between an individuel person and his or her environment. [...] The systems we are now describing is the environment and the embedded device interacting over time, not the device in isolation. Thus classification models constitute a level of system description, but they cannot be reduced to or mapped onto preexisting physical structures in individual devices. » (Clancey, 1993, p.36)

Cette longue citation permet de bien situer l'enjeu d'une description formulée en langue naturelle. Il ne s'agit pas en effet de décrire comment un système isolé se comporte, mais de décrire comment ce système interagit avec son environnement dans des situations concrètes. Or, c'est précisément cet aspect qui concerne le projet de l'IA: traiter des problèmes concrets en utilisant les connaissances portant sur l'interaction entre des systèmes (en général, les spécialistes du domaine) et l'environnement où se pose le problème à traiter. Il ressort que ces connaissances sont indispensables dans tout SBC, et que la langue naturelle véhicule le sens commun dont a besoin tout système raisonnablement « intelligent » pour adopter un comportement de résolution de problème adapté à son contexte d'utilisation :

> « In most artificial intelligent systems, some knowledge of particular problems domains has been encoded, but most of this work has been too specific to be of much use to the field in general. There is a certain minimum of "core knowledge", however, that *any* reasonably sophisticated intelligent agent must have to make its way in the world. This is not knowledge

> of the world as science perceives it, but the "commonsense theories" of the world that ordinary people have – their everyday beliefs about what the world is like. » (Hobbs & Moore, 1985, p.xi)

Il s'agit pour l'IA d'utiliser ces connaissances, de les opérationnaliser en un système informatique. Ce système sera mis en œuvre pour résoudre le problème et entrera en interaction soit avec le monde concret directement ou avec ses utilisateurs. Il s'agira alors d'un nouveau système établissant un nouveau type d'interaction qui donnera lieu, selon les principes établis par Clancey, à de nouvelles connaissances descriptives, taxinomiques ou classificatoires, dégageant ce que l'utilisation du système permet de faire émerger comme comportement de résolution de problème dans le domaine. Le fait même d'utiliser le système engendre inévitablement de nouvelles connaissances sur le domaine : l'existence concrète du système devient un outil pour la constitution de nouvelles connaissances. On retrouve ainsi notre constat selon lequel l'effectivité du système le détermine comme un outil de modélisation. Mais, ici on s'aperçoit que ce caractère d'effectivité est plus générique que l'effectivité computationnelle ou calculatoire : c'est l'effectivité comme existence concrète et interaction réelle avec un environnement.

Cela suggère que la constitution de connaissances provient, de manière générale, de la mobilisation d'outil matériel dans un contexte concret et que le comportement dynamique de l'ordinateur n'est qu'un cas particulier d'outil matériel dont l'utilisation permet la constitution de nouvelles connaissances. En d'autres termes, l'informatique ne serait pas un modèle de l'esprit, mais l'étude d'objets matériels particuliers, les algorithmes, dont la mise en œuvre dans des contextes concrets permet la constitution de nouvelles connaissances.

Cela suggère par conséquent deux choses : d'une part, toute connaissance procède de la médiation d'un objet concret dont elle décrit l'interaction avec le monde réel dans le cadre d'un contexte concret. Ce sera le thème de l'herméneutique matérielle de préciser les modalités de cette médiation. D'autre part, l'informatique propose la médiation d'objets matériels particuliers, les calculs symboliques. La section suivante et le chapitre 5 s'attachent à montrer que les calculs sont bien des objets matériels effectifs et qu'il est légitime de considérer que l'IA est un cas particulier de l'herméneutique matérielle appliquée à la médiation du calcul symbolique.

2.2 Réalisme pratique et réalisme scientifique

En voulant caractériser le contenu des connaissances véhiculées par la langue, nous sommes tombés avec Clancey sur une opposition entre les connaissances descriptives des pratiques ou interactions et les connaissances scientifiques décrivant un système isolé. Cette opposition se traduit par une irréductibilité : les connaissances en langue ne sont pas des connaissances scientifiques mal formulées, mais des connaissances visant un aspect de la réalité que les connaissances scientifiques ne prennent pas en compte.

Or, cela peut entraîner de nombreuses difficultés pour un projet comme celui de l'IA. Nous avons en effet montré plus haut que l'IA avait besoin d'un principe de modélisation et d'un principe d'effectivité. Ce que nous venons de voir montre qu'il peut exister un conflit entre ces deux principes :

le principe de modélisation explicite comment dégager le savoir contenu dans les connaissances exprimées en langue et déterminer de quelle manière de savoir se rapporte au monde dont il parle ; le savoir obtenu ainsi n'est pas obligatoirement compatible avec une expression scientifique, c'est-à-dire dans les langages de la science ;

le principe d'effectivité explicite comment opérationnaliser le modèle ou savoir obtenu grâce au principe de modélisation ; or, ce principe ressortit par définition aux connaissances scientifiques, puisqu'il correspond à l'informatique comprise comme science des calculs effectifs ;

Par conséquent, affirmer l'irréductibilité des connaissances exprimées en langue à des connaissances scientifiques, du fait qu'elles ne visent pas la même chose, entraîne une tension méthodologique entre les principes de modélisation et d'effectivité.

Il est en effet bien connu que la réalité exprimée dans la langue naturelle n'est pas celle construite par la science et il peut avoir un conflit entre d'une part le monde de la science et d'autre part le monde que nous appelerons « monde de la pratique » étant donné ce que nous avons souligné sur les relations entre langue et pratique. En d'autre terme, le réalisme « pratique » lié aux connaissances exprimées en langue naturelle constituant le monde de la pratique se heurte au réalisme « scientifique » lié aux connaissances scientifiques. Putnam exprime cette opposition de la manière suivante :

> « Ainsi nous voyons que le nom « Réalisme » peut être réclamé par, ou peut donner lieu à, au moins deux attitudes philosophiques différentes (et en fait à plusieurs). Le philosophie qui prétend que seuls les objets de la science *existent réellement* et que la plus grande part, sinon tout le monde, du sens commun est une simple « projection », se veut réaliste : mais le philosophie qui insiste sur la *réalité* des choses et des cubes de glace fait de même, et ces deux attitudes, ces deux représentations du monde, peuvent conduire à plusieurs programmes différents pour la philosophie. » (Putnam, 1991, p. 84)

Pour faire bref dans le cadre de cette introduction (le chapitre 5 reprend en détail l'argumentation que nous allons esquisser à présent), l'opposition du réalisme pratique au réalisme scientifique est une opposition entre le réalisme de la *chose* telle qu'elle s'offre à nous « naturellement » et « spontanément », c'est-à-dire sans la médiation des constructions théoriques des sciences expérimentales de la nature, et le réalisme de l'*objet*, tel que la science expérimentale permet d'en construire.

Le problème est qu'un objet n'est pas une chose, et réciproquement. Les objets construits (couleur comme longueur d'onde) par la science ne sont pas ces objets spontanés ou choses (couleur dans son sens commun) au moyen desquels nous catégorisons notre monde quotidien. Les choses, ressortissant à notre rencontre immédiate avec le monde, sont d'un point de vue scientifique naïves et pré-théoriques : elles ne sont pas vraiment des objets, mais plutôt des « pseudo-objets ». En d'autres termes, elles ne peuvent revendiquer la même objectivité que les objets scientifiques dans la mesure où elles ne peuvent posséder l'exactitude de la détermination scientifique.

On retrouve ainsi l'opposition que faisait Husserl dans les *Ideen* (Husserl, 1950) entre les objets susceptibles d'une approche géométrique, c'est-à-dire mathématique, et les objets qui ne pouvaient relever que d'une approche descriptive. Il existe donc à ce niveau une fracture dans le savoir où les connaissances que l'on possède sur les choses ne se laisse pas transcrire dans le cadre des sciences mathématiques, et où les connaissances que l'on possède sur les objets ne laissent pas transcrire dans les sciences descriptives. En effet, la notion d'« électron » ne se laisse adéquatement décrire que depuis le formalisme mathématique de la physique et les verbalisations auxquelles ce formalisme se prête explicitent en langue le contenu de cette notion sans pour autant y être adéquates[12].

Par conséquent, la compréhension du monde que prodiguent les connaissances exprimées en langue naturelle, et les actions qu'elles permettent de mener à partir de cette compréhension, ne possèdent pas d'emblée des corrélats dans le cadre de l'explication scientifique des phénomènes. Par conséquent, il peut y avoir un conflit entre d'une part l'explicitation du savoir contenu dans ces connaissances et d'autre part leur opérationnalisation. Il peut exister un clivage entre les connaissances dites « scientifiques » ou formelles et les connaissances non formelles. Nous avions déjà noté ce clivage dans (Bachimont, 1994) et nous l'avions alors appelé *différence phénoménologique*.

Tout programme méthodologique pour l'IA doit donc trancher la question de la nature des connaissances exprimées en langue et leur irréductibilité éventuelle à la science. Plus précisément, deux questions permettent de déterminer l'approche méthodologique adoptée :

- quelle différence peut introduire le fait de s'exprimer en langue naturelle plutôt que dans un langage formel logico-mathématique ;
- de manière liée au point précédent, quelle différence de nature peut-il exister entre le rapport au monde exprimé par la langue naturelle, appelé généralement le sens commun, et le savoir théorique sur le monde exprimé par les langages formels de la science.

L'une des forces du formalisme comme principe de modélisation est de répondre précisément à ces deux questions en affirmant l'identité de nature entre sens commun et science, entre savoir exprimé en langues naturelles et savoir exprimé en langages scientifiques, résolvant ainsi la tension méthodologique inhérente au projet de l'IA, comme nous venons de le voir.

3 Le formalisme

Le formalisme est une notion qui recouvre de nombeux aspects distincts et nous avons vu plus haut qu'il fallait distinguer le formalisme comme modélisation des connaissances d'une part et le formalisme comme effectivité computationnelle d'autre part. Le formalisme comme modélisation des connaissances consiste dans l'affirmation que formaliser des connaissances exprimées en langue, c'est dégager en quoi elles sont des connaissances et en quoi consiste leur contenu. Le formalisme est alors une théorie épistémologique sur

12. Cette opposition est également proche de celle souligner par Kant dans la *Critique de la faculté de Juger* (Kant, 1984a) entre l'explication finaliste et l'explication mécanique, la première permettant de comprendre et de rationaliser les phénomènes mais échouant à en donner une explication scientifique, et la seconde, expliquant les scientifiquement les phénomènes mais échouant de ce fait à expliquer tous les phénomènes.

la nature de la connaissance. Le formalisme comme effectivité computationnelle consiste dans le fait que le syntaxico-formel correspond à la description de l'effectivité de processus naturels en termes d'exécutions d'algorithmes ou calculs. Le formalisme est alors une théorie particulière de la nature physique, bien qu'entendue en un sens élargi.

3.1 Précisions terminologiques

Mais avant de préciser ces deux aspects du formalisme, il convient de fixer notre usage des termes comme « modélisation », « formalisation », « formel », « modèle », etc., dans la mesure où nous en ferons abondamment usage et également dans la mesure où il n'existe pas de réel consensus quant à l'usage correct de ces termes. Nous appelons donc :

modèle tout système de connaissances permettant de raisonner à propos d'une réalité et d'en anticiper l'évolution sous un certain point de vue ; un modèle se caractérise par

- **la systématicité** qu'il introduit entre les notions, c'est-à-dire que les notions s'interdéfinissent par l'intermédiaire des relations qui les relient ; la systématicité implique qu'il faut être « systématique » : par exemple, si « droit » est défini, « gauche » doit l'être ; le système doit définir les notions qui, appartenant à un même niveau de généralité, s'interdéfinissent et s'entre impliquent ; par exemple, si un modéle médical pour la coronarographie contient la notion de sténose, il doit également contenir la notion de diamètre, d'artère, et débit sanguin ; en revanche, il n'est pas nécessaire qu'il possède les notions ressortissant au métabolisme des graisses, dans la mesure où ce n'est pas le point de vue de la coronarographie ni le niveau de détail qui la concerne ;
- **la cohérence** qu'il impose aux relations défissant les notions ; formellement il ne doit pas être possible d'inférer P et $\neg P$;
- **la correction** qu'il recherche en autorisant que des raisonnements vrais dans le domaine ;
- **la complétude** qu'il recherche en déterminant les notions ; le modèle doit contenir tous les éléments permettant de raisonner sur la réalité visée en fonction de l'objectif poursuivi ;

formel tout ce qui est consigné dans les termes d'un langage artificiel comme la logique ou les mathématiques, où la compréhension des énoncés est commandée par la forme plutôt que par le contenu des structures symboliques employées ;

formalisation le fait d'exprimer des connaissances dans un langage formel ;

modélisation formelle la conception d'un modèle formel ; parmi les modèles formels, il convient de distinguer une

- **axiomatisation** où le modèle prescrit à la réalité décrite une structure syntaxico-formelle ; par exemple, c'est la description de la réalité linguistique à partir de grammaires formelles, ou bien la description d'un système physique comme la régulation cardiaque à partir d'un système de règles formelles ; en d'autres termes, une axiomatisation élabore une grammaire des phénomènes à expliquer ;

mathématisation où le modèle ne prescrit pas à la réalité décrite une structure syntaxico-formelle mais une réalité décrite par des équations ; c'est par exemple la description de la régulation cardiaque à l'aide d'un système dynamique ; en d'autres termes, une mathématisation met en équation les phénomènes à expliquer.

Dans la mesure où tout système de connaissance est un modèle, une formalisation peut être comprise comme une modélisation formelle ; « formalisation » n'est donc pas synonyme d'« axiomatisation », que l'on appelera parfois « modélisation syntaxico-formelle » ou encore « formalisation syntaxique » ; cependant, comme il ne s'agit dans notre travail que de formalisation syntaxique, nous employerons le terme de « formalisation » dans ce sens, précisant quand le contexte l'exige s'il faut le considérer de manière générique ;

modélisation informelle tout système de connaissances exprimé dans un langage non formel; par exemple, les connaissances exprimées dans la langue naturelle à propos d'un problème donné et rationalisées en une théorie du domaine constituent un modèle informel du problème et de la réalité du monde dans laquelle il s'insère.

La principale distinction opposant « axiomatisation » de « mathématisation » repose sur le fait que dans une axiomatisation, la syntaxe du modèle renvoie à la réalité du monde alors que même si un modèle mathématique possède sa propre axiomatisation, ce n'est pas cette structure syntaxico-formelle qui est attribuée à la réalité mais les équations déduites de cette structure. Après ces clarifications terminologiques, revenons au formalisme et montrons comment il fournit à l'IA le principe de modélisation et le principe d'effectivité dont elle a besoin pour mener à bien son projet.

3.2 Formalisme et modélisation des connaissances

Le formalisme comme modélisation des connaissances affirme que la formalisation des connaissances permet de dégager en quoi des connaissances sont des connaissances. On rencontre une telle conception dans les thèses des philosophes formant le Cercle de Vienne [13], notamment sous la plume de Schlick qui écrit :

> « Toute connaissance est donc par essence une connaissance des formes, des relations, et rien d'autre. Seules les relations formelles au sens défini sont accessibles à la connaissance, au jugement au sens purement logique du terme. » (Schlick, 1985, p.188-189)

Confronté à une connaissance exprimée en langue naturelle, il est nécessaire, pour dégager le savoir qu'elle recèle, de faire abstraction de tout contenu, au sens de vécu « évoqué » ou « éveillé » par un terme de l'énoncé ; en effet, le vécu est privé, individuel et incommunicable. Le vécu ne peut donc être une connaissance puisqu'une connaissance est

13. La bibliographie est abondante sur un mouvement intellectuel aussi important, par son contenu et par sa postérité, que le Cercle de Vienne. Ce mouvement s'inscrit dans les différents courants philosophiques nés de la révolution Frégéenne de la logique. On peut consulter à ce sujet (Jacob, 1980b), (Jacob, 1980a), (Malherbe, 1981), (Engel, 1989).

par définition une vérité sur le monde qui comme telle est partageable. Schlick note ainsi :

> « Le vécu est contenu alors que le connaître est par nature dirigé vers la pure forme. [...] On exprimera peut-être mieux la tâche et la fonction purement formelles de la connaissance en disant que tout connaître est, sans cesse, un ordonner et un calculer, jamais un regarder et un vivre les choses. » (Schlick, 1985, p.188)

En affirmant que toute connaissance est formelle et que tout contenu est incommunicable car ressortissant au vécu privé qui s'évoque et s'éveille chez autrui sans se communiquer, Schlick affirme par là même que tout ce qu'un mot veut dire, s'il possède du sens, c'est-à-dire s'il véhicule une connaissance, ne peut consister que dans les relations formelles qu'il décrit entre des objets. Ainsi, qu'un mot soit utilisé par un poète ou un scientifique ne changera rien à son contenu ; seulement dans un cas, on l'utilise non pour son sens mais pour son pouvoir d'évocation – c'est le cas du poète, et dans l'autre, on l'utilise pour son contenu – c'est le cas du scientifique. En d'autres termes, il ne peut exister de différence de nature entre les langues naturelles et les langages artificiels dans la mesure où ces derniers sont spécialement conçus pour dégager le « contenu » formel des expressions linguistiques en tant qu'elles véhiculent ou expriment des connaissances. Le formalisme du Cercle de Vienne abolit donc la différence entre langue naturelle et langage scientifique dans l'expression des connaissances :

> « Car le sens de ces mots, utilisés par le poète ou par le psychologue, ne peut en toutes circonstances être donné et expliqué qu'en le réduisant aux relations formelles entre les objets. Le mot « vert » n'est en rien plus riche (au contraire, il est même plus pauvre) que le concept de fréquence des oscillations lumineuses par lequel le physicien l'a remplacé. Le mot « vert » n'exprime pas réellement ce qui est vécu lorsqu'on regarde une prairie verte, le mot n'a aucune affinité de contenu avec le vécu du vert ; il n'exprime qu'une relation formelle selon laquelle touts les objets que nous nommons verts sont reliés les uns aux autres. » (Schlick, 1985, p.187)

Cela implique que, selon le Cercle de Vienne, il n'y a pas lieu d'opposer le sens commun et la vérité scientifique. Si, en effet, le sens commun possède une vérité sur le monde, c'est qu'il possède un contenu formel qui le rend analogue en cela avec les connaissances scientifiques :

> « Aussi paradoxal que cela puisse paraître, il est littéralement vrai que tous nos énoncés, des plus triviaux dans la vie de tous les jours aux plus complexes dans les sciences, ne reflètent que des relations formelles au sein du monde, et que n'entre en eux tout simplement rien de la qualité du vécu. [...] Nous voyons maintenant qu'à cet égard les énoncés de la physique théorique ne diffèrent en rien de tous les autres énoncés de la vie quotidienne, pas plus que des énoncés des sciences de l'esprit. » (Schlick, 1985, p.186-187)

Toute connaissance est donc formelle. Mais il s'agit de savoir s'il s'agit d'un formel syntaxique ou d'un formel mathématique, si les modèles que l'on conçoit doivent être compris comme des axiomatisations ou des mathématisations. Or, là dessus, les auteurs du Cercle de Vienne sont également très clairs : la formalisation est une analyse logique du langage si bien que les relations formelles dégagées par la formalisation des connaissances sont des relations logiques syntaxiques.

Si le projet de l'IA est résoudre des problèmes informatiques en opérationnalisant les connaissances exprimées en langue naturelle dont on dispose sur ces problèmes, le formalisme du Cercle de Vienne fournit le principe modélisateur de la démarche de l'IA: l'IA modélise la réalité du problème à résoudre en formalisant les connaissances puisque la formalisation ne fait pas autre chose que de dégager en quoi ces connaissances sont des connaissances du problème à résoudre.

Mais il manque encore un principe opérationnel qui rend le modèle ainsi obtenu effectif, pour qu'il puisse être opérationnalisé en un système matériel concret qui assure la résolution du problème posé. Il s'agit d'obtenir un principe général garantissant que tout modèle obtenu par la formalisation des connaissances soit effectif et opérationnalisable. Il ne suffit en effet pas, pour un projet comme celui de l'IA, qu'il y ait des constructions effectives de systèmes concrets de résolution selon des méthodes établies au cas par cas ou de manière *ad hoc*. Il faut avoir un principe opérationnel qui garantisse *par principe* que toute formalisation est effective. Un tel principe est fourni par la notion de machine de Turing universelle.

3.3 Formalisme et effectivité computationnelle

Le formalisme n'est pas seulement une conception épistémologique de la connaissance affirmant la formalité de la science. Le formalisme précise également en quel sens la formalité renvoie à l'effectivité. Mais ce n'est pas le formalisme du cercle de Vienne qu'il nous faut considérer pour cela, mais celui de Hilbert qui, lors de la crise des fondements des mathématiques, énonça la conception que le raisonnement mathématique et logique devait être compris comme une manipulation syntaxique de symboles considérés uniquement en fonction de leur type symbolique (par exemple en étant tel ou tel symbole de constante ou de variable) indépendamment de toute signification éventuellement attachée à ce symbole.

3.3.1 Le programme de Hilbert

On sait en effet qu'au tournant du siècle se posèrent de graves questions quant au statut des objets mathématiques. Les avancées extraordinaires qu'avaient connues les mathématiques au cours du XIX^{e} siècle avaient coïncidé avec une abstraction croissante des objets manipulés qui se retrouvaient ainsi échapper à toute tentative d'y avoir accès par une intuition intellectuelle ou représentation mentale. On en vint dès lors à distinguer dans les mathématiques deux champs principaux :

- des mathématiques dites *abstraites* ou *idéales* qui portent sur des entités qui n'ont rien de concret ou de physique, par exemple les nombres transfinis de Cantor ou l'infini actuel, qui ne se rencontrent ni dans la réalité *physique*, dans la réalité *psychologique* : concevables, ces entités ne sont pas représentables dans la pensée ni effectuables par la pensée ;
- des mathématiques dites *concrètes* ou *finitistes* car elles portent sur des entités obtenues et contrôlables par des moyens combinatoires *finis* ; à l'époque de Hilbert, il n'était pas évident de définir avec précision ces mathématiques finitistes même si la notion est intuitivement facile et claire ; cela correspond à toutes les mathématiques et entités que l'on peut obtenir en effectuant des *calculs finis*, c'est-à-dire en termes

modernes, par des moyens algorithmiques tel qu'un ordinateur pourrait les mettre en œuvre ; les mathématiques finitistes correspondent aux mathématiques que l'on pourrait confier à un ordinateur.

Mais, pour Hilbert, les mathématiques finitistes correspondent à des mathématiques s'effectuant selon un mécanisme reposant sur « l'attention au contenu concret et finitiste de la pensée », à la manière dont les vérités arithmétiques s'édifient « uniquement par des constructions numériques au moyen de considérations intuitives à contenu » (Hilbert, 1972).

Autant les mathématiques concrètes semblent hors de doute car elles reposent sur des entités comme les nombres entiers à l'objectivité incontestables, autant les mathématiques abstraites renvoyent à une fondation incertaine. Si comme l'avait souligné Kronecker, « Dieu a créé les nombres entiers laissant le reste aux mathématiciens », il est bon de s'interroger sur la réalité des objets mathématiques autres que les nombres entiers pour les y confronter [14].

Des objets mathématiques idéaux comme les transfinis rendent crucial et urgent ce type d'interrogation dans la mesure où de tels objets soulevaient de graves questions quant à leur éventuelle réalité: penser et réfléchir sur ces objets, est-ce penser et réfléchir à quelque chose d'objectif? Peut-on conférer aux mathématiques abstraites la même indubitabilité que celle des mathématiques concrètes? Faut-il au contraire les éliminer au profit de ces dernière? Mais plus grave encore, des contradictions et des paradoxes furent formulés qui montraient comment ces notions, au contenu difficile à maîtriser, pouvaient conduire à des contradictions insolubles :

> « Or l'application immédiate des conceptualisations et des argumentations s'étaient progressivement imposés à l'usage courant, produisait des contradictions : contradictions d'abord isolées, puis qui peu à peu apparurent de plus en plus profondes et sérieuses. C'étaient les paradoxes de la théorie des ensembles. » (Hilbert, 1972)

La situation devenait dès lors intolérable: « Qu'on y réfléchisse: en mathématiques, cette science réputée pour un modèle de sûreté et de vérité, les notions et les inférences, celles que chacun a apprises, que chacun enseigne ou emploie, conduisent à des insanités! Si la pensée mathématique elle-même échoue, où trouvera-t-on sûreté et vérité? » (Hilbert, 1972).

3.3.1.a Le but

Le programme de Hilbert consiste à atteindre les deux buts suivants :

> 1. Nous voulons examiner soigneusement les conceptualisations et les types d'inférence féconds de la théorie des ensembles, et, partout où c'est possible avec une chance de

14. Notre évocation du problème des fondements et de l'objectivité en mathématiques est fort superficielle et ne saurait tenir lieu d'une discussion de ces problèmes. N'étant évoquées ici que pour situer la nature du formalisme syntaxique qui nous importe ici, ces questions trouvent un traitement approfondi, tant sur le plan historique qu'épistémologique dans (Salanskis & Panza, 1995), (Largeault, 1992), (Salanskis & Sinaceur, 1992), (de Rouilhan & Rivenc, 1992).

réussite, les étayer ou les rendre utilisables. Car il ne faut pas qu'on nous chasse du paradis que Cantor a créé pour nous.

2. Il est indispensable de donner partout aux inférences la sûreté qu'elles ont en arithmétique élémentaire, théorie à l'abri du doute et dans laquelle contradictions et paradoxes, s'il en survient, ne sont imputables qu'à notre manque de soin.

(Hilbert, 1972, p.228)

Le but de Hilbert est donc sauver les mathématiques abstraites. Pour cela, il va élaborer une procédure d'élimination ou de réduction qui va les rapporter aux mathématiques concrètes : toute opération mathématique portant sur des entités abstraites doit se convertir en une opération ne mobilisant que des entités concrètes ; les entités abstraites sont donc *éliminées* lors de cette réduction.

Il ne s'agit pas d'une réduction dans le sens où il faudrait *supprimer* les mathématiques abstraites, mais d'une réduction où l'on veut les *garantir* et lever tout doute éventuel quant à leur objectivité et existence[15]. Le principe est donc d'autoriser les mathématiques abstraites dans la mesure où elles se ramènent aux mathématiques concrètes, et non de les interdire puisqu'elles se ramènent aux mathématiques concrètes. En d'autres termes, si l'élimination du transfini pouvait toujours en principe réussir, on serait alors assuré que toutes les notions mathématiques mobilisant l'infini, comme le calcul infinitésimal, ou encore toute formule possédant des quantificateurs comme « il existe » ou « pour tout » portant sur les individus d'un ensemble infini, ne sont que des manières de parler, mais des manières de parler autorisées et non à proscrire comme le voulait un Kronecker par exemple.

3.3.1.b Le principe

Le principe de la réduction est fourni par la logique mathématique. L'invention de l'axiomatique formelle et de la notion de système formel montre que des propriétés peuvent être vraies de systèmes de choses dont la nature est indifférente : les relations se transportent d'un domaine d'entités à un autre. Ainsi, de manière triviale, les lois de l'arithmétique sont valables que l'on compte des choux ou des carottes. La structure de corps est valable pour les nombres réels et les nombres complexes ; l'algèbre de Boole-Schröder permet de prendre en compte tant de la logique des classes que la logique des propositions ; etc.

Or, dans les systèmes formels, les symboles sont utilisés *indépendamment de tout contenu abstrait*, mais seulement en fonction des *relations formelles* qu'ils entretiennent avec les autres symboles. Par exemple, l'élément neutre d'un corps pour une loi se définit fonctionnellement par la loi $x * e = x$ pour tout x, indépendamment du fait de savoir quels objets on veut associer à x et e : si la loi est la multiplication dans N, e vaut 1, s'il s'agit de l'addition, e vaut 0. Par conséquent, les opérations mathématiques effectuées dans un système formel sont indépendantes du contenu mathématique, abstrait ou concret, associé aux symboles du système.

L'autre point fondamental qu'il faut relever pour comprendre la démarche hilbertienne, est que les opérations effectuées dans le cadre des manipulations formelles prescrites par

15. Comme le souligne fort bien Jean Largeault, « on élimine les concepts et les propositions transfinies *en théorie* pour avoir le droit de les conserver *en pratique* » (Largeault, 1972, p.215).

le système formel relèvent des mathématiques finitistes concrètes. Cela signifie que ce qui est entrepris dans le cadre des systèmes formels possède une objectivité hors de doute. Pourquoi ces opérations sont-elles concrètes ? Tout simplement parce qu'elles sont des manipulations de type mécaniques, que l'on applique de manière machinale, sans réfléchir, sur des symboles concrets, des signes sur le papier qui relèvent de l'intuition empirique et de sa catégorisation.

A présent, il est facile de décrire la démarche de Hilbert : rendre compte des mathématiques abstraites comme l'interprétation de la manipulation formelle de symboles concrets définis par un système formel. Ce qui est décrit en termes abstraits est pris en charge par des mathématiques finitistes portant sur les symboles concrets dont ils sont l'interprétation. Il suffit dès lors de prouver de manière finitiste, dans la traduction concrète des mathématiques abstraites, que ce qui est entrepris dans le cadre de ces dernières n'introduit pas de contradiction dans les mathématiques concrètes. L'adjonction de nouveaux énoncés, dépourvus de sens puisque l'on ne peut recourir à leur contenu dans la mesure où il est abstrait, revient à adjoindre de nouveaux énoncés *idéaux* : l'adjonction n'est donc légitime que si elle n'introduit pas de contradiction dans le domaine auquel ils s'ajoutent : « une extension par adjonction d'éléments idéaux n'est en effet recevable que s'il n'en résulte pas de contradiction à l'intérieur de l'ancien domaine, c'est-à-dire par conséquent si les relations apparaissant entre les anciens objets en termes desquels on élimine les objets idéaux sont toujours valables dans l'ancien domaine. »

Le programme de Hilbert repose donc sur le fait que toutes les mathématiques, abstraites ou concrètes, s'écrivent, et que pour comprendre l'écriture des premières, il faut faire appel à des entités abstraites à l'objectivité douteuse alors que l'écriture des secondes se définit et se manipule à partir de règles formelles complètement explicites.

3.3.1.c Son échec

Le programme de Hilbert ne put et ne peut être mené à bien : des théorèmes prouvent l'impossibilité de son succès ; c'est notamment le théorème d'incomplétude de Gödel (1931) (Gödel, 1989).

Le théorème d'incomplétude de Gödel affirme que tout système formel suffisamment complexe, c'est-à-dire permettant de formaliser l'arithmétique, contient des formules indécidables : il est possible de trouver des formules exprimées dans le langage du système formel qui ne sont pas des théorèmes et dont les négations ne sont pas non plus des théorèmes. Mais, le plus important concernant ces formules indécidables est qu'elles *sont néanmoins vraies*. Le théorème de Gödel prouve que dans tout système formel suffisamment complexe, *il existe des formules vraies non démontrables*, c'est-à-dire des formules qui *abstraitement sont vraies et qui, concrètement ne sont ni vraies ni fausses,* i.e. *indécidables* [16].

Autrement dit, il existe des formules finitistes, c'est-à-dire exprimées dans un système formel, exprimant (interprétée par) un contenu abstrait, pour lequel il n'existe pas de moyen finitiste de savoir si la formule est vraie ou non : être un théorème pour une for-

16. « Ainsi, la proposition qui, *dans le système PM* [système formel des *Principia Mathematica* que Gödel a considéré pour mener sa démonstration], est indécidable peut pourtant être décidée par des considérations métamathématiques. » (Gödel, 1989, p.112).

mule dans un système formel signifie qu'il existe une dérivation depuis les axiomes du système reposant sur l'utilisation mécanique et formelle des règles de déduction. Etre un théorème pour une formule, c'est avoir sa vérité prouvée de manière finitiste. S'il existe une formule vraie en mathématique abstraite qui d'un point de vue finitiste, ne peut être ni vraie ni fausse, cela prouve que les mathématiques abstraites échappent par définition aux mathématiques concrètes. Or, l'existence de telle formule et le moyen d'en construire est précisément ce qui est donné par le théorème de Gödel. On est alors incapable de prouver de manière finitiste que cette formule est compatible, cohérente avec les autres formules puisque sa vérité est inaccessible aux procédés finitistes. On est ainsi incapable de prouver par des moyens finitistes qu'un système concret supplémenté des formules idéales transcrivant les mathématiques abstraites est consistant.

3.3.2 La mécanisation du calcul

Le programme de Hilbert, malgré son échec[17], a cependant permis d'expliciter les notions et principes indispensables à la mécanisation du calcul, voire de la pensée. En effet, pour sauver les mathématiques abstraites, il a fallu expliciter et mathématiser ce qu'il fallait comprendre par « mathématiques concrètes », c'est-à-dire les mathématiques correspondant à des manipulations finies de symboles concrets ou de nombres entiers. C'est ainsi que Gödel a posé la notion de « fonction récursive » pour mathématiser les calculs et expliciter la traduction des formules abstraites en calculs concrets.

Mais Turing (Turing & Girard, 1995) déplaça sensiblement la question en élaborant le concept d'une machine abstraite permettant d'effectuer les manipulations formelles de signes dépourvus de signification. Alors que le formalisme logique de Hilbert posa la notion de raisonnement mathématique comme manipulation de symboles ininterprétés, Turing introduisit la notion de manipulation *automatique* de symboles ininterprétés.

On connaît le principe d'une machine de Turing : une tête de lecture se déplace sur une bande mémoire infinie composée de case contenant un blanc ou un caractère pris dans un alphabet fini. La tête de lecture ne lit le contenu que d'une seule case à la fois : elle n'est jamais à cheval sur plusieurs cases. Seuls un nombre fini de cases contiennent un caractère non blanc. La tête de lecture peut se trouver dans un certain nombre d'états internes : ces états sont en nombre fini et mutuellement exclusifs. La tête de lecture peut lire ou écrire un symbole sur une case, se déplacer sur la case de droite ou sur la case de gauche. Le comportement de la tête de lecture est déterminé par un ensemble de quintuplets, constituant le programme qu'exécute la machine, dictant la conduite à tenir en fonction de deux critères : l'état interne de la machine, le symbole lu dans la case où pointe la tête de lecture à un instant donné. Selon l'état interne q_i courant et le contenu p de la case lu par la tête de lecture, un quintuplet (p, q_i, L, q_j, r) prescrit d'écrire le symbole r à la place de p sur la case courante, de se déplacer d'une case vers la gauche (L) et de se mettre dans

17. Il faut cependant relativiser la portée de cet échec dans la mesure où un programme de recherche en mathématiques, le courant des *Reverse Mathematics*, étudie comment fonder les théories mathématiques habituelles sur des axiomatiques ne mobilisant que des nombres entiers. Plus précisément, on recherche sur une base fixée, quels sont les axiomes d'existence qui permettent d'obtenir les résultats centraux de l'analyse et de l'algèbre. Il s'est avéré que des systèmes « faibles », c'est-à-dire (entre autres) dont la consistance est équivalente à celle de l'arithmétique de Peano, permettent d'aller très loin et en particulier suffisent pour mener à bien toute l'analyse et toute l'algèbre d'application courante. (Feferman, 1988) et (Simpson, 1985) sont les références classiques. Voir (Delahaye, 1995) pour une introduction à ces questions.

l'état interne q_j. On appelle « machine de Turing » un tel dispositif muni d'un ensemble de quintuplets ou programme [18].

On constate que la manipulation des symboles s'effectue bien indépendamment de tout contenu attaché à ses symboles et que seul importe leur type pris dans un alphabet fini (savoir qu'un « a » est un *a* plutôt qu'un *b*). Une telle machine possède des propriétés remarquables : elle peut calculer toute les fonctions récursives. Réciproquement, elle permet de préciser la notion de calcul : une fonction est calculable s'il existe une machine de Turing pouvant calculer son graphe. De plus, il existe une machine de Turing universelle qui peut simuler l'exécution de n'importe quelle machine de Turing simple. Une machine de Turing universelle est universelle dans la mesure où toute fonction calculable est calculable par une telle machine.

Le concept de machine de Turing universelle permet de donner un contenu mathématique précis à la notion de calcul. Le formalisme hilbertien avait, nous l'avons vu, caractérisé un calcul comme une manipulation de signes ininterprétés. Le formalisme turingien rapporte cette manipulation à un mécanisme en décrivant la structure d'une machine à laquelle confier la manipulation. En rendant la manipulation automatique, le formalisme turingien permet de considérer un calcul comme un processus naturel, c'est-à-dire un processus produit par les seules forces de la nature.

En effet, un calcul peut être défini comme étant un processus matériel permettant, avec des ressources finies et en un temps fini, de décider si un objet donné appartient à un ensemble donné. En d'autres termes, on définit par la notion de calcul par la notion d'« effectivement décidable ». Selon les ouvrages de logique mathématique, par exemple (Mendelson, 1987), un ensemble est effectivement décidable s'il existe « une procédure mécanique qui détermine, pour tout objet, s'il appartient ou non à l'ensemble. Par une *procédure mécanique*, nous voulons dire une procédure qui s'exécute automatiquement sans faire aucun recours à une intuition ou une originalité dans son application. ».

En définissant la notion d'algorithme comme étant ce que calcule une machine de Turing, Turing précise formellement ce qui restait désigné de manière flou et vague comme étant une « procédure mécanique ne faisant pas appel à l'intuition (*ingenuity*) ». Un calcul, c'est désormais une procédure mécanique exécutée par une machine de Turing.

Mais, il faut remarquer que la machine de Turing ne s'encombre pas de description sur la manière de construire réellement, physiquement une telle machine. La machine est spécifiée formellement et abstraitement. Cela signifie que ce que fait une telle machine est indifférent à la manière dont physiquement elle le réalise. En d'autres termes, des machines physiquement différentes, construites selon des principes physiques différents (par exemple l'une reposant sur une physique des semi-conducteurs, une autre reposant sur une physique de relais mécaniques) peuvent être identiques du point de vue des calculs qu'elles exécutent. La machine de Turing permet donc de caractériser des processus matériels d'un point de vue abstrait, du point de vue du calcul qu'ils exécutent. L'informatique peut donc être définie comme la science des algorithmes ou du calcul, dans la mesure où elle décrit un niveau autonome du réel indépendant des niveaux physiques qui le réalisent.

La science des algorithmes s'intéresse en effet à un niveau abstrait du réel qui possèdent ses propres lois et donc son autonomie. Cependant, ce niveau n'est qu'une abstraction dans

18. Voir par exemple (Mendelson, 1987) ou (Lalement, 1990) pour des détails sur la machine de Turing.

la mesure où il ne peut exister d'algorithme que physiquement réalisé, de machine que réellement construite. Mais de la même manière qu'il est possible de décrire les phénomènes de vorticité et d'en faire la modélisation mathématique sans faire appel au niveau corpusculaire à la base des tourbillons observés, il est possible de décrire le comportement d'un programme, d'analyser ses propriétés, par exemple sa complexité, indépendamment du substrat physique qui permet de le réaliser.

Il est donc tout à fait cohérent d'affirmer avec Newell et Simon que l'informatique est une science de la nature, même s'il faut entendre par nature une nature élargie au niveau computationnel, élargie à un niveau abstrait et autonome mais qui n'en décrit pas moins que des processus « naturels ».

> « Computer science is an empirical discipline. We would have called it an experimental science, but like astronomy, economics and geology, some of its unique forms of observation and experience do not fit a narrow stereotype of the experimental method. Nonetheless, they are experiments. Each new machine that is built is an experiment. Actually constructing the machine poses a question to nature ; and we listen for the answer by observing the machine in operation and analyzing it by all analytical and measurement means available. [...] Inspection of the program in the light of a few runs reveals the flaw and lets us proceed to the next attempt. » (Newell & Simon, 1981, 36)

Le questionnement empirique de la nature ne porte pas sur la réalisation physique des programmes, mais sur l'examen du programme lui-même et de ses exécutions. Le niveau computationnel ou algorithmique est autonome dans la mesure où les phénomènes étudiés ainsi que leur modélisation sont formulés au même niveau : c'est par exemple la maîtrise de l'explosion combinatoire qui est observée lors de l'exécution et qui permet, selon qu'elle est effective ou non, de conjecturer quelles modifications du programme permettront de la juguler (Newell & Simon, 1981, 36)[19].

Il est à présent clair que les formalismes hilbertien et turingien fournissent un principe d'effectivité qui, couplé au principe de modélisation fourni par le formalisme du Cercle de Vienne, donne à l'IA les moyens théoriques de réaliser son projet. Selon ce dernier, les connaissances linguistiques s'analysent en représentations logico-formelles et tout raisonnement mené sur ces connaissances revient à un calcul sur les représentations correspondantes. Or, selon les formalisme hilbertien et turingien, tout calcul se ramène à une manipulation automatique, donc effective, de représentations syntaxiques ininterprétées. Par conséquent, on obtient bien un système qui peut utiliser effectivement (principe d'effectivité : le formel est opérationnel) le savoir exprimé par les connaissances linguistiques (principe de modélisation : le formel dégage le savoir du linguistique).

4 Intelligence artificielle et formalisme

L'IA trouve dans le formalisme une théorie épistémologique de la connaissance et un principe technologique pour réaliser son projet. L'IA, dans une telle approche formaliste, est un projet consistant à vouloir *axiomatiser* les phénomènes étudiés, que ces derniers soient

19. Nous partageons totalement cette vue selon laquelle l'informatique est une science expérimentale de la nature. Nous donnons une analyse détaillée de cette conception dans le chapitre 5 en nous appuyant sur la philosophie kantienne des sciences.

mentaux et relèvent d'une psychologie scientifique ou qu'ils soient externes et relèvent d'une science de la nature physique. Le but est par conséquent de concevoir un système formel automatique dont le caractère modélisateur repose sur sa capacité à axiomatiser les phénomènes étudiés.

Selon les strates historiques que nous avons déjà dégagées, il faut distinguer dans l'IA un projet d'axiomatisation de l'esprit, correspondant à ce que nous avons appelé « IA psychologique », et un projet d'axiomatisation du monde, correspondant à ce que nous avons appelé « IA constructiviste ».

4.1 L'IA comme axiomatisation de l'esprit

Un des principaux apports du formalisme hilbertien fut de pouvoir considérer l'activité mathématique comme une manipulation de signes ininterprétés, compris comme des signes concrets : cette manipulation se fait avec autant d'évidence que l'arithmétique des nombres entiers. A partir du moment où l'activité mathématique n'est plus considérée comme une activité intellectuelle, dont la vérité repose sur l'évidence intellectuelle, mais où elle se ramène à une activité de manipulation de signes dont la vérité repose sur l'évidence sensible, on rapporte l'activité de raisonnement mathématique et calculatoire à des comportements externes, à des manipulations externes de symboles.

Le formalisme hilbertien renvoie en cela au formalisme leibnizien qui instaura cette conception du calcul comme manipulation de signes ininterprétés en opposition à l'intuitionisme cartésien. En effet, Leibniz, tout comme Descartes, voulait donner à la raison un guide sûr et infaillible pour mener ses raisonnements. Tout comme Descartes, il se tourne vers les connaissances dont la certitude paraissait la plus inébranlable, les mathématiques, pour élargir à l'ensemble des connaissances le principe de leur certitude. Mais, alors que Descartes pensait que la certitude mathématique reposait sur l'évidence des raisonnements qui s'adresse à l'intuition de l'esprit, Leibniz considérait au contraire que leur certitude procédait du formalisme avec lequel les démonstrations sont menées. Pour Descartes, les démonstrations mathématiques sont certaines parce que évidentes, claires et distinctes ; pour Leibniz, les démonstrations sont certaines car effectuées en suivant des règles de calcul formel dont l'application ne fait pas appel au génie de celui qui les manipule, mais au simple respect mécanique et machinal de règles comme celles que l'on applique pour effectuer les opérations élémentaires comme l'addition ou la multiplication.

> « Ce que je demande, ce sont des critères de la vérité qui soient palpables, et qui ne laissent pas plus de place au doute que des calculs sur des nombres. Ce qu'il faut, ce sont des marques palpables de ce qui est clair et distinct, puisque souvent les hommes sont en désaccord là-dessus. » Leibniz ; cité dans (Couturat, 1901, p.100 note 2)

En voulant rapporter les mathématiques abstraites aux mathématiques concrètes, Hilbert ne fait pas autre chose que substituer à une intuition intellectuelle chancelante l'intuition concrète des signes composant les systèmes formels et aux raisonnements incertains car reposant sur le génie du mathématicien la règle de calcul à l'application certaine. Le formalisme n'est pas la suppression de toute intuition, mais la substitution de l'intuition sensible, à la certitude indiscutable. Comme le rappelle Yvon Belaval,

> « Si l'on refuse à l'intuition rationnelle la vision immédiate de l'absolu, il ne reste pour « arrêter » notre pensée que l'intuition empirique. Le symbole concret, écrit, est perçu sans

> contestation. Il n'a de sens, bien entendu, que par l'ensemble des opérations dans lesquelles on le fait entrer : mais ces opérations suivent les lois nécessaires de la raison. » (Belaval, 1960)

Leibniz ne dit lui-même pas autre chose :

> « La véritable méthode nous doit fournir un *filum Ariadnes*, c'est-à-dire un certain moyen sensible et grossier qui conduise l'esprit comme font les lignes tracées en géométrie et les formes des opérations qu'on prescrit aux apprentis en Arithmétique. » Leibniz

En d'autres termes, le formalisme leibnizien *externalise* la source de certitude de l'intuition intellectuelle comme évidence à l'intuition sensible comme perception. La mécanisation effectuée par le formalisme turingien apporte un point final à cette externalisation de la pensée calculante en montrant comment une machine peut assurer l'exécution des prescriptions d'un algorithme. Autrement dit, si les règles de calcul sont machinales et imposent une évidence empirique que tout un chacun peut voir et doit prendre en compte, c'est que ces règles correspondent à la description d'un processus naturel, mécanique et physique. Il est dès lors possible de voir dans la notion de système formel automatique un principe permettant d'expliquer la nature de la pensée, si l'on admet avec Hobbes (Hobbes, 1651) que « penser, c'est calculer »[20].

En affirmant que tout raisonnement sur des connaissances consiste dans l'application mécanique de règles calculatoires sur des symboles concrets, les principes formalistes peuvent être compris comme l'affirmation que la pensée, raisonnant sur des symboles pour atteindre la vérité et la connaissance, est un système *physique* appliquant à ce titre *machinalement* des règles *mécaniques* sur des symboles *concrets* comme des signes matériels consignés sur le papier ou sur un support neuronal. En d'autres termes, la pensée est un *système symbolique physique* (Newell & Simon, 1981; Newell, 1980), un système qui, en tant qu'il manipule des symboles est un système manipulant des signes ou symboles ininterprétés et qui, en tant que ces symboles sont physiquement réalisés, est une machine. Une machine de Turing, réalisable matériellement mais agissant sur des symboles, est en effet un système symbolique physique :

> « A physical symbol system consists of a set of entities, called symbols, which are physical patterns that can occur as components of another type of entity called an expression (or symbol structure). Thus a symbol structure is composed of a number of instances (or tokens) of symbols related in some physical way (such as one token being next to another). At any instant of time the system will contain a collection of these symbol structures. Besides these structures, the system also contains a collection of processes that operate on expressions to produce other expressions : processes of creation, modification, reproduction, and destruction. A physical symbol system is a machine that produces through time an evolving collection of symbol structures. Such a system exists in a world of objects wider than just these symbols expressions themselves. » (Newell & Simon, 1981, p. 40)

Mais un symbole est ici un signe ininerprété, un signe sans signification. Il correspond à une classe de token ou objets physiques : il se définit comme la configuration physique (*physical pattern*) que réalisent ces objets physiques. Par conséquent, le symbole au sens des systèmes symboliques physiques ne possède pas la principale caractéristique que l'on

20. D'ailleurs, Haugeland (Haugeland, 1981; Haugeland, 1985) voit en Hobbes le véritable précurseur de l'IA au sens où lui-même la comprend (voir *supra*).

attribue au symbole quand on affirme par exemple que la pensée humaine est symbolique, ou que les connaissances sont symboliques ; on veut en effet dire par là que les symboles mis en œuvre dans la pensée ne sont pas dépourvus de sens, mais que bien au contraire ils en sont chargés.

Or, affirment Newell et Simon, l'un des résultats de la recherche *empirique* sur les systèmes symboliques physiques est de montrer comment les symboles qu'ils mobilisent peuvent *désigner* un référent, par exemple un processus, et dans ce dernier cas, ils peuvent être interprétés, l'interprétation produisant l'exécution du processus désigné (si le système possède la capacité d'exécuter ce processus). Et en effet, tout langage de programmation fondé sur la manipulation de listes (*list processing*) comme LISP vérifie ces propriétés de désignation et d'interprétation [21].

En donnant un définition purement computationnelle de la notion de symbole, l'informatique émet une hypothèse empirique, physique, sur la nature du symbole dans la pensée humaine, donnant une réponse à la question, plagiée sur un article célèbre de McCullogh [22], de savoir ce que doit être un symbole pour que l'intelligence puisse l'utiliser, et l'intelligence pour qu'elle puisse utiliser un symbole.

Puisque l'intelligence repose en son principe sur le symbole, et que les langages de manipulation de listes rendent compte du caractère « symbolique » des symboles, à savoir de désigner et de s'interpréter, les systèmes symboliques physiques constituent l'hypothèse qu'il faut formuler quant à la nature de l'intelligence : « A physical symbol system has the necessary and sufficient means for general intelligent action » (p. 41). Les auteurs de cette hypothèse précisent clairement son statut :

> « One of the fundamental contributions to knowledge of computer science has been to explain, at a rather basic level, what symbols are. This explanation is a scientific proposition about Nature. It is empirically derived, with a long and gradual development. » (Newell & Simon, 1981, p. 37)

Il s'agit en effet d'une hypothèse sur la nature de l'intelligence, c'est-à-dire sur l'intelligence comme phénomène *naturel* appartenant à la « Nature ». Il s'agit par conséquent d'une hypothèse mettant en place un programme de recherche dans le cadre des sciences expérimentales de la nature portant sur l'explication scientifique de l'intelligence, comprise ici comme le comportement intelligent, ou l'action intelligente (*intelligent action*).

Les principes formalistes, conduisant naturellement à la notion de système symbolique physique, conduisent par conséquent, au terme du raisonnement que nous avons mené en

21. (Newell, 1980) précise cette notion de désignation en la rapprochant de l'universalité de la machine de Turing universelle : en effet, cette dernière est universelle parce qu'elle peut simuler toute machine de Turing particulière ; elle possède donc dans ses données une description symbolique qui désigne la machine particulière qu'elle simule. Par conséquent, la notion d'universalité implique celle de désignation. Réciproquement, la désignation implique l'universalité puisque si une machine peut désigner toute autre machine, elle peut la simuler et donc être universelle. En d'autres termes, si LISP exemplifie de manière paradigmatique que les symboles informatiques désignent, cette propriété est inhérente à tout calculateur pour peu qu'il soit universel. Il s'agit donc d'une argumentation qui vaut pour toute l'informatique et non simplement pour LISP.

22. « What Is a Number, That a Man May Know it, and a Man, That he May Know a Number? » *General semantics Bulletin*, Nos 26-27, 7-18, 1961.

suivant Newell et Simon, à une hypothèse scientifique sur la nature de l'intelligence. Par conséquent, un projet technologique comme celui d'exploiter des connaissances exprimées en langue naturelle conduit au projet scientifique de naturaliser l'esprit, c'est-à-dire de l'expliquer depuis des hypothèses formulées dans le cadre des sciences de la nature, en comprenant cette dernière de manière suffisamment large pour inclure le niveau computationnel.

Ainsi, si le projet de l'IA tel que nous l'avons formulé renvoie en quelque sorte à un projet « faible », à une IA « faible » ou « prudente », en revanche, la formulation formaliste du projet de IA renvoie à une science de l'esprit et donc à un projet « fort » pour l'IA, à une IA « forte ». John Searle proposa d'ailleurs de distinguer entre une IA forte et une IA faible :

> « Il me paraît utile de faire la distinction entre ce que j'appelerai l'IA « forte » et l'IA « faible », ou « prudente ». D'après l'IA faible, la principale valeur de l'ordinateur pour l'étude de l'esprit, c'est qu'il est pour nous un outil très puissant. Ainsi, il nous permet de formuler et de tester des hypothèses de façon plus rigoureuse et plus précise. D'après l'IA forte, en revanche, l'ordinateur n'est pas simplement un outil de l'étude de l'esprit ; l'ordinateur convenablement programmé est véritablement un esprit, en ce sens que des ordinateurs munis des bons programmes comprennent et ont d'autres états cognitifs. En IA forte, l'ordinateur programmé ayant des états cognitifs, les programmes ne sont pas simplement des outils nous permettant de tester des explications ; ils sont, eux-mêmes, des explications» (Searle, 1981)

Mais cette distinction n'est pas très satisfaisante dans la mesure où on ne voit pas pourquoi l'ordinateur serait un outil très puissant pour l'étude de l'esprit en particulier plutôt que pour un autre domaine d'investigation. En d'autres termes, si l'ordinateur est un outil privilégié pour l'étude de l'esprit, c'est que le formalisme computationnel possède une valeur particulière dans l'étude de l'esprit. L'ordinateur n'est donc alors pas qu'un outil de test rigoureux et précis, c'est aussi un outil de modélisation. Dans ce cas, on passerait alors à une IA forte, qui utilise l'ordinateur pas seulement comme un outil pour l'étude de l'esprit mais comme un modèle de l'esprit. Mais, il ne va pas de soi qu'utiliser l'ordinateur comme étude de l'esprit implique qu'un ordinateur soit lui-même un esprit quand il est correctement programmé. C'est cependant ce que soutiennent les zélateurs du cognitivisme computationnel comme Pylyshyn (Pylyshyn, 1984) (voir *infra*).

La méthode de l'IA est en effet de décrire la résolution en termes de connaissances, telle qu'elle se réalise chez les êtres humains. Ces descriptions sont symboliques et doivent être reproduites telles quelles sur la machine. L'IA cherche donc à simuler des comportements humains dits intelligents, et doit utiliser des modèles de ces comportements pour les reproduire, ou tout au moins, en reproduire les effets. Il s'agit dans cette perspective de résoudre certains problèmes en simulant le raisonnement d'un expert humain de référence. La méthode de l'IA est donc de *simuler* le comportement cognitif humain. Mais, et c'est là toute la question, la simulation est à comprendre ici non dans le sens de la modélisation scientifique, qui reconstruit (mathématiquement ou axiomatiquement) les phénomènes sans prétendre en dévoiler la nature ultime, mais selon l'acception logique que Turing (1990) a fixée :

> « Le terme de simulation est employé ici dans le sens fort, mythique depuis Turing, qui fait généralement référence en IA: il suggère une production indiscernable, ou à tout le moins fonctionnellement substituable à un original sans provoquer de déficit significatif en situation d'interaction réelle. » *in* (Visetti, 1991, p. 266)

C'est ainsi en effet que Pylyshyn écrit :

> « Je veux soutenir que le calcul computationnel est un modèle littéral de l'activité mentale, et non une *simulation* du comportement, comme cela a été quelque fois affirmé dans les premières années de la modélisation cognitive. Contrairement à la simulation, disons, d'un processus chimique ou du flot d'un trafic, je n'affirme pas simplement que le modèle engendre une séquence de prédictions du comportement, mais bien plutôt qu'il le fait essentiellement de la même manière ou en vertu des mêmes mécanismes fonctionnels (et non évidemment des mêmes mécanismes biologiques) et *en vertu du fait* de posséder quelque chose qui correspond aux mêmes pensées ou états cognitifs que ceux qui commandent le comportement de l'organisme modélisé. Etre la même pensée implique posséder le même contenu sémantique (*i.e.* des pensées identiques possèdent des contenus sémantiques identiques). » *in* (Pylyshyn, 1984, p.43)

Au lieu d'une simulation approximative, on a bien une simulation reproductrice, et même *productrice* du comportement cognitif. Et c'est donc avec raison que l'on peut affirmer qu'avec une telle simulation, simuler la pensée, c'est penser, simuler la parole, c'est parler etc. Le modèle n'est donc pas une approximation du phénomène, mais son essence.

Or, une telle attitude renvoie le projet d'une IA forte non pas à un projet scientifique, mais à un projet métaphysique dans la mesure où la simulation à laquelle se livre l'IA forte n'est pas simplement une explication scientifique puisque les phénomènes observés sont reproduits par le modèle construit, mais une thèse sur la la nature même des choses qui produisent le phénomène.

Il faut en effet distinguer dans l'étude de l'esprit l'attitude selon laquelle le modèle computationnel permet de formuler des hypothèses scientifiques quant à la nature de l'esprit de l'attitude selon laquelle le modèle de l'ordinateur révèle la nature essentielle de l'esprit. Dans le premier cas en effet, un modèle computationnel permet de décrire et de rendre compte des phénomènes mentaux et de l'utilisation des connaissances. Dans le second cas, le modèle computationnel ne reconstruit pas seulement les phénomènes mentaux, mais en révèlent l'essence. Ainsi, alors que la première attitude fait une science computationnelle de l'esprit sans trancher sur la nature essentielle de l'esprit, la seconde considère que l'esprit est un logiciel dont le cerveau est le matériel et que par conséquent tout ordinateur correctement programmé pense.

En tant que telle, une science ne peut répondre à ce type de question concernant l'essence des phénomènes qu'elle reconstruit dans ses modèles. Si l'IA forte, à travers les thèses de Pylyshyn, y répond, c'est qu'il ne s'agit plus de sciences cognitives mais de philosophie cognitive. On ne peut en dire autant de la thèse des systèmes symboliques physiques qui, en tant que thèse, formule simplement le cadre dans lequel il faut entreprendre la description et explication scientifiques des phénomènes de l'esprit. En revanche, il est tout-à -fait concevable d'associer à cette thèse un postulat ontologique et philosophique sur la nature essentielle de l'esprit affirmant que l'esprit est *réellement* un système symbolique physique. Mais les articles cités ici de Newell et Simon ne l'affirment pas explicitement.

4.2 L'IA COMME AXIOMATISATION DU MONDE

Comme nous l'avons déjà souligné, l'approche formaliste peut être consacrée non pas à une axiomatisation du sujet ou de l'esprit, mais à une axiomatisation du monde. Il ne s'agit plus de voir dans les bases de connaissances des modèles psychologiques d'un expert, mais des modèles du monde. En effet, le point de vue psychologique fait oublier que le principal intérêt des bases de connaissances, ce n'est pas qu'elles représentent le raisonnement de l'expert, mais qu'elles modélisent le monde :

> « When we refer to knowledge bases as expert or consultation systems, we may forget that they describe some system of the world – not just expert reasoning or the consultative process. » (Clancey, 1989, p.10)

En considérant les bases de connaissances comme des modèles qualitatifs du monde physique, social, cognitif ou perceptif, on restitue la véritable nature de ces bases ainsi que leur enjeu. Ainsi, au lieu de reconstruire la vie mentale à partir de l'hypothèse des systèmes symboliques physiques, il sera désormais question d'axiomatiser les phénomènes du monde en un système formel. Cela est particulièrement patent dans le domaine de la « physique qualitative » :

> « L'approche de base du physicien naïf qui travaille en intelligence artificielle consiste en ceci qu'il prend un échantillon d'inférences inductives faisant partie d'un domaine donné et qu'il se soucie de développer des langages formels capables d'exprimer axiomatiquement la connaissance en question et de saisir formellement les procédures inférentielles pertinentes. » (Smith & Casati, 1993)

On peut vérifier la justesse de cette remarque en consultant par exemple (Hobbs & Moore, 1985) où l'on peut lire en effet que « we should aim for an axiomatization of spatial relationships, shape, structure, and motion, that, as a test of adequacy, satisfies three different criteria. It should mesh well with the view of space that seems to be implicit in language [...] with vision research [...] and should be adequate for detailed discussion of geography. » (p. xii). Bref, il s'agit d'une « axiomatization of the common sense domain » (p. xvii).

Le problème est de savoir si une reconstruction syntaxico-formelle des phénomènes tels qu'ils sont donnés à travers une description linguistique est possible. Or, et c'est là le constat négatif motivant notre réflexion, la réponse est négative.

5 LES PROBLÈMES DE L'APPROCHE FORMALISTE EN IA

Le principal obstacle à une reconstruction informatique syntaxico-formelle est la non formalité de la langue naturelle et des connaissances qui y sont exprimées. En d'autres termes, le projet de réaliser le projet de l'IA en s'appuyant sur une approche formaliste de la connaissance est impossible.

La non formalité des connaissances linguistiques a été souvent relevée et certains auteurs y ont vu sa caractéristique essentielle. Notamment, Terry Winograd a clairement soulevé ce point et y a reconnu une pierre d'achoppement fondamentale pour toute approche formaliste de la connaissance telle qu'une IA classique pouvait adopter. Dans l'exemple suivant, Winograd montre comment le contexte est indispensable pour comprendre et analyser le contenu d'une assertion en langue naturelle :

> « Imagine that a doctor asks a nurse "Is the patient eating?" If they are deciding whether to perform an examination, the request might be paraphrased "Is she eating at this moment?" If the patient is in the hospital for anorexia and the doctor is checking the efficacy of the treatment, it might be more like "Has the patient eaten some minimal amount in the past day?" If the patient has recently undergone surgery, it might mean "Has the patient taken any nutrition by mouth," and so on. In responding, a person interprets the sentence as having relevance in the current situation, and will typically respond appropriately without conscious choosing among meanings. » (Winograd, 1990, p. 180)

Le sens d'un énoncé dépend de son insertion dans le contexte, de la même manière que le sens des mots qu'il contient dépend de l'énoncé. En d'autres termes, le sens véhiculé par la langue n'est pas compositionnel : le sens du tout n'est pas composé par le sens des parties. Or, une approche formaliste repose au contraire sur l'hypothèse de la compositionnalité du sens dans la mesure où la notion même de syntaxe générative permet, par des règles de construction, l'assemblage de symboles pour en faire des formules, et où la signification associée à de telles formules est déterminée par le sens associé aux symboles primitifs et par les règles de composition sémantique permettant de composer le sens des symboles pour déduire le sens de la formule qu'ils constituent.

On retrouve le problème de la non compositionnalité des termes également dans les taxinomies construites dans les domaines non formels. L'idéalisation correspondant à la systématisation des connaissances et des dénominations n'entraîne pas la fixation de la signification des termes. C'est ainsi qu'en zoologie, on peut rencontrer semblables phénomènes :

> « Même dans un domaine bien défini, les normes qui fixent la signification du mot, et par suite sa référence, n'ont pas de régularité prévisible. Prenons un exemple dans le domaine de la zoologie, bien structuré par ses traditions académiques et scientifiques. Dans *buse variable*, *variable* signifie « dont la couleur varie selon les individus », alors que dans *lièvre variable*, *variable* signifie « qui change de couleur selon les saisons ». » (Rastier *et al.*, 1994, p. 60)

Par ailleurs, nous avons vu plus haut (section 1-1) que le terme de « fonction diurétique » possédait une signification contextuelle. Il s'agit d'un phénomène très général qui concerne tous les domaines reposant sur la langue naturelle pour exprimer leur connaissance.

Comme le note Winograd, ce problème « lies in the symbol system hypothesis itself. In order to characterize a situation in symbolic form, one uses a system of basic distinctions, or terms. Rules deal with the interrelations among the terms, not with their interpretation in the world ». Or, ce qui nous intéresse dans les connaissances exprimées en langue naturelle, c'est précisément leur interprétation dans le monde, ce qu'elles commandent de comprendre et de faire dans le monde.

Ce fait est également fortement souligné en linguistique par F. Rastier qui y voit une manière de congédier toute approche formelle de la langue :

> « Si le principe de compositionnalité est invalide en sémantique linguistique et si l'interprétation échappe ainsi au paradigme du calcul, la détermination du global sur le local s'exerce

> par les contraintes initiales imposées ou proposées au parcours interprétatif. » (Rastier *et al.*, 1994, p. 37)

C'est qu'en effet, le sens d'énoncés linguistiques ne leur est pas immanent, il ne leur est pas propre ou local. Il est ainsi impossible de décontextualiser un énoncé formulé en langue exprimant un savoir pour obtenir une représentation de ce savoir. Si l'exemple pris par Winograd se situe dans un contexte médical, et si les connaissances médicales ne s'expriment qu'en langue naturelle, il ne s'agit pas d'un hasard. En effet, le savoir médical est pratique et contextuel :

> « the concept of context is vital in medical knowledge and appears at all levels of expertise in different forms : (*i*) context for the *interpretation of data*. Context reflects the history of the patient, the diagnosis, or the response to the treatment. (*ii*) context to account for various *temporal phases* in the course of a patient's disease. [...] Although clinical practice is largely based on consensual knowledge, each clinical team, specialized in a particular medical field, accumulates specific experience. This difference with commonly accepted knowledge reflects habits, cultural factors, and specific skills. [...] Medical knowledge is inherently unstable ans varies according to theoretical and experimental research as well as advances in technology. » (Dojat & Pachet, 1995, p. 405)

De la même manière, si la langue naturelle nous permet d'exprimer notre rapport au monde tel que le caractérise la notion de sens commun, c'est que le monde qui se présente ainsi n'est pas « atomiste » et compositionnel, mais holiste. Relevé dans (Hayes, 1985), le holisme des faits relevant du sens commun, comme ceux de la physique naïve, interdit leur description par des moyens purement formels, ou du moins ceux relevant d'une théorie atomiste telle que la théorie des ensembles :

> « les représentations contemporaines de l'expérience de sens commun sont extrêmement étroites, au point d'être suspectes en tant que représentations qui se veulent fidèles. Nous avons suggéré que cette étroitesse découle d'une adhésion trop précipitée aux traits attirant du point de vue de l'implantation, traits qui renvoient à leur tour à l'emploi des moyens essentiellement atomistes (non holistiques) de la théorie des ensembles. » (Smith & Casati, 1993)

Par ailleurs, comme nous l'avons souligné plus haut, la langue naturelle permet d'exprimer un rapport au monde et une attitude à prendre dans un contexte donné, une conduite à tenir. Elle seule permet d'exprimer le savoir constituant une pratique comme la médecine. Seule la langue permet d'exprimer avec des termes généraux (par exemple « le patient a-t-il mangé ? », « la sténose est serrée », etc.) un rapport au contexte qui prescrit une conduite à tenir. Dès lors, le projet de décontextualiser une telle connaissance en l'extrayant de sa gangue linguistique n'a plus du sens : en passant d'une sémantique contextuelle de la langue naturelle à une sémantique compositionnelle, localiste et non contextuelle d'un langage formel, on perd le contenu ou le savoir véhiculé par l'énoncé, c'est-à-dire le rapport au monde ou au contexte (cf. ce qu'affirme Winograd ci-dessus).

> « Le sens d'un texte ne lui est pas immanent, il est immanent à la situation de communication à laquelle appartient le texte. Il varie donc avec les paramètres de cette situation. Comme tous les objets culturels, il est soumis à une évolution historique. » (Rastier *et al.*, 1994, p. 16 (note))

Les principes formalistes mettent en place une conception objectiviste de la connaissance dans la mesure où une connaissance est rapportée à une représentation syntaxique qui possède des réalisations physiques. La connaissance est dès lors un objet que l'on range en mémoire et que l'on manipule dans des processus symboliques :

> « la conception objectiviste de la connaissance inspire la définition des connaissances comme des chaînes de symboles, que l'on peut « manipuler » et « stocker ». Cette réification est inacceptable, car ces connaissances supposées sont produites et traitées indépendamment de tout contexte et de toute situation. Et une connaissance n'existe pas en soi, même une vérité factuelle : elle est toujours située dans une culture. » (Rastier *et al.*, 1994, p. 14)

Comme le souligne F. Rastier, cela revient à vouloir reconnaître une objectivité aux connaissances indépendamment du contexte de leur utilisation: s'analysant comme une représentation syntaxico-formelle, une connaissance existe en tant qu'entité computationnelle, au même titre qu'un programme. Or, ce que signifie la non compositionnalité de la langue, c'est qu'une connaissance n'existe qu'en contexte. En d'autres termes, il est nécessaire, pour exploiter des connaissances exprimées en langue de renoncer au principe formaliste de modélisation qui voit en une connaissance une représentation syntaxique possédant une autonomie objective. Il faut adopter une théorie de la connaissance adaptée au holisme de la langue, adaptée à la détermination du local par le global. Une telle théorie de la connaissance doit s'adosser sur une théorie de la sémantique de la langue : ce n'est que l'étude des expressions linguistiques des connaissances que l'on peut atteindre le contenu de ces connaissances. Une telle théorie sémantique nous est founie par la sémantique différentielle. La théorie de la connaissance associée sera celle des essences différentielles. La sémantique différentielle est le paradigme adopté par F. Rastier pour élaborer une théorie de la sémantique des textes. Prenant congé du paradigme formaliste, elle s'oriente vers une dimension herméneutique :

> « Elle [la sémantique différentielle] peut être dite herméneutique, car elle reconnaît comme constitutive la détermination du local par le global, à l'inverse de la conception combinatoire de la systématicité qui a justifié le paradigme formel en linguistique. » (Rastier *et al.*, 1994, p.34)

Mais une théorie sémantique ne dit pas comment modéliser le savoir exprimé dans les énoncés puisque, par principe, elle ne s'occupe pas du rapport au monde[23]. Une théorie permettant de décrire le sens n'est pas une théorie permettant de modéliser les connaissances qu'il exprime. Par conséquent, congédier le principe formaliste de modélisation du fait qu'il est incompatible avec le sens des énoncés linguistiques laisse un vide qu'il faudra combler.

6 Symbole et orthothéticité

Il s'agit à présent, après avoir précisé le constat négatif selon lequel les modèles développés en IA ne peuvent être des modèles dénotationnels, de reprendre le constat positif selon lequel les modèles de l'IA modélisent dans la mesure où d'une part ils reposent sur des

23. Nous traitons ce point en détail dans le chapitre 4.

signifiants linguistiques empruntés à l'expression en langue des connaissances et d'autre part ils opérationnalisent les relations entre signifiants explicitant ainsi le contenu des connaissances qui les interprètent.

6.1 Interprétabilité du symbolique

Le fait que les symboles utilisés dans les systèmes formels automatiques que sont les artefacts IA soient empruntés à la langue naturelle, *i.e.* soient des unités linguistiques prélevées sur l'expression linguistique des connaissances du domaine, est essentiel. Il est en effet à l'origine d'une caractéristique tout à fait paticulière de ces systèmes, à savoir qu'ils sont « investis d'esprit ». Précisons ce que nous entendons par cette notion.

6.1.1 Des objets investis d'esprit

Il est important de remarquer que les systèmes conçus en IA sont, en pratique, utilisés comme de véritables *auxiliaires cognitifs*, c'est-à-dire qu'on les utilise comme produisant du sens, indépendamment du problème de savoir s'ils simulent cette production de sens qu'on leur prête ou s'ils la produisent effectivement.

Le fait est que bien que l'on sache qu'ils ne pensent pas, bien qu'ils ne soient pas le siège de phénomène cognitif, on interagit avec ces systèmes comme si c'était le cas. Même si le système expert médical ne pense pas, et qu'il ne viendrait pas à l'idée d'un utilisateur médecin de le croire, on considère que son résultat est un diagnostic, et pas seulement une inférence formelle. On suit ses prescriptions en tant que diagnostic, et pas en tant qu'inférence formelle. Quand Mycin infère sur les infections causées par la bactérie E. Coli, l'utilisateur ne s'intéresse au résultat du programme que d'un point de vue médical. On ne veut pas que le système *simule* l'inférence d'un diagnostic, on veut qu'il le produise, exactement comme on ne veut pas qu'une voiture simule le déplacement dans l'espace, mais qu'elle le produise effectivement, même si on ne lui prête pas la locomotion propre aux êtres vivants.

L'expérience concrète des artefacts conçus en IA est qu'ils sont effectivement *anthropomorphisés*, et qu'on leur prête raisonnements et pensées. C'est d'ailleurs la seule stratégie rationnelle que peut posséder un utilisateur pour construire son interaction avec le système. Et cet utilisateur entend bien que cette stratégie soit payante: qu'il soit allé plus loin en termes de connaissances en utilisant un système informatique que d'avoir fait du sur-place avec un simulateur de raisonnement.

La situation des artefacts IA est en effet différente des simulations informatiques habituelles, où l'on explore les possibilités d'un modèle. Dans une simulation informatique, le résultat produit correspond à la forme prédite du phénomène mais non à sa mesure effective : c'est une prédiction théorique et non la donnée d'une mesure expérimentale par un capteur. Dans un système IA, le résultat n'est pas considéré comme une prédiction d'un raisonnement, mais comme étant le raisonnement lui-même. Ainsi, en télémédecine, l'utilisateur se conformera aux prescriptions du système comme s'il était en face d'un praticien ; dans un serveur de connaissances encyclopédiques, les sorties sont utilisées comme des connaissances du domaine et non comme des simulations de connaissances.

Par ailleurs, cela implique que dans le cas d'une simulation informatique, le système *produit* de véritables connaissances sur le modèle lui-même, alors qu'il ne fait que simuler, par l'intermédiaire du modèle, la réalité modélisée. C'est d'ailleurs ce qui se produit dans le cas des modèles complexes dont on ne connaît pas les propriétés, et que l'on opérationnalise informatiquement pour les étudier : la réalité explorée n'est pas le monde mais le modèle lui-même, et le travail n'est pas un travail de prédiction théorique mais un travail d'investigation expérimentale sur la nature du modèle.

On constate que les modèles, entités théoriques, motivent des explorations expérimentales. Cela signifie que des entités théorique comme les modèles doivent investir des supports matériels pour acquérir davantage de détermination, c'est-à-dire de précision idéale et théorique. Il faut en passer par l'observation empirique des supports matériels dans lesquels s'investissent les modèles pour accroître la connaissance que l'on en possède. Cela n'est pas propre à l'opérationnalisation informatique. L'écriture est la première forme, et la forme exemplaire, de l'investissement de la matière par l'idéalité et c'est ainsi qu'elle en permet la contemplation et la connaissance. Toutes les idéalités sont écrites, sont inscrites sur des supports, il n'y pas d'idéalité non matérielle. L'ordinateur n'est donc à cet égard qu'un support de plus, à la suite des parchemins, livres, figures, etc., mais avec cette caractéristique essentielle que désormais le support est dynamique.

Le caractère dynamique du support est ce qui permet d'utiliser l'ordinateur comme un auxiliaire cognitif dans la mesure où un livre propose des liens statiques que seul le lecteur peut parcourir dynamiquement. En parcourant dynamiquement les liens existant entre les différentes inscriptions, l'ordinateur s'investit d'esprit. Cette dynamicité est à la source de la production de nouvelles connaissances, nouvelles pour l'utilisateur qui les reçoit comme telles. C'est la raison pour laquelle l'ordinateur paraît pouvoir penser alors qu'il ne le fait pas. En donnant au livre la dynamicité qui lui manque, l'ordinateur transforme cet objet investi d'esprit en un processus investi d'esprit, c'est-à-dire ce qui ressemble à s'y méprendre à un agent.

C'est qu'en effet, l'ordinateur est une machine physique particulière, une machine « informatique » qu'il faut distinguer d'une machine physique au sens habituel. Les machines physiques permettent d'opérationnaliser une action physique dans le monde réel : elles automatisent un savoir-faire. Ces machines sont construites à partir des connaissances pratiques, comme le moulin à vent, ou bien à partir de connaissances théoriques appliquées, comme les machines mécaniques ou thermiques. Dans ces derniers cas, la machine opérationnalise non pas la connaissance théorique en tant que telle, mais son schème, c'est-à-dire son rapport à la réalité physique et concrète. Les machines physiques sont construites à partir de connaissances portant sur la forme des phénomènes physiques dont on veut obtenir la reproduction. Par conséquent, la machine physique met en œuvre le pouvoir causal de la matière physique de manière à produire dans le monde les conséquences physiques prédites par les connaissances théoriques. La machine physique opérationnalise donc ces connaissances dans la mesure où elle produit les phénomènes dont elles prédisent la forme théorique.

La situation est sensiblement différente pour les machines informatiques, dans la mesure où elles n'opérationnalisent pas un savoir-faire, mais un savoir théorique : ce qui est automatisé n'est pas une action dans le monde physique, mais une action dans le monde théorique, conceptuel. Les machines informatiques automatisent la forme des phénomènes physiques indépendamment de leur matière. Les machines informatiques permettent donc d'opérationnaliser les connaissances théoriques dans leur fonctionnement théorique et conceptuel, et non dans leur rapport au monde réel. Ce qui est reproduit, c'est le fonctionnement théorique – plus exactement son fonctionnement symbolique – du modèle et des connaissances, et non les processus naturels décrits par les connaissances. Les machines physiques utilisent des connaissances théoriques pour mettre en jeu des processus naturels dont elles décrivent la légalité, alors que les machines informatiques mettent en jeu les connaissances elles-mêmes et non la réalité qu'elles décrivent.

C'est pourquoi l'informatique a pu devenir une science cognitive et proposer un temps les modèles les plus convaincants pour naturaliser l'esprit. En reproduisant le fonctionnement des connaissances et non ce qu'elles décrivent, les machines informatiques semblent reproduire l'activité même de la pensée et de ce fait s'investir d'esprit.

On peut se demander ce qu'il faut exactement comprendre par objet ou processus « investis d'esprit ». Nous empruntons la notion d'« objet investis d'esprit » à Husserl qui, dans les *Ideen II* (Husserl, 1982), montre comment un objet comme le livre ne peut être considéré seulement comme un objet physique et matériel auquel se surajouterait un sens, le sens du contenu qu'il véhicule. Le livre, comme livre, est un objet qui traversé de part en part par un sens qui l'« anime » :

> « Le livre avec ses feuilles de papier, sa reliure, etc., est une chose. A cette chose, il n'y en a pas une autre, le sens, qui se raccrocherait, mais au contraire celui-ci, d'une certaine manière, *pénètre de fond en comble* le tout physique *en « l'animant »* ; à savoir dans la mesure où il anime chaque mot, mais, encore une fois, non pas chaque mot pour soi mais des chaînes de mots qui sont liées de façon à former des configurations douées de sens, celles-ci à leur tour des configurations plus élevées, etc. Le sens spirituel, en animant les apparences sensibles *fusionne* d'une certaine manière avec elles, au lieu de leur être lié dans un simple juxtaposition. » (Husserl, 1982, p. 326)

L'objet s'investit d'esprit dans la mesure où il « fusionne » avec le sens qu'il véhicule : il ne fait qu'un avec lui. Autrement dit, un objet comme un livre ne peut recevoir le qualificatif de livre qu'à partir du moment où il n'existe que par et pour le sens qui l'anime. Etant d'emblée animé par un sens « spirituel », le livre s'adresse à l'esprit qui y reconnaît une unité de sens, une intentionnalité qui s'adresse à lui. Par conséquent, l'esprit reconnaît dans le livre la trace de l'esprit.

En apportant une dynamicité au support physique d'inscriptions symboliques, les représentations syntaxico-formelles élaborées en IA par exemple, l'ordinateur apporte au sens fusionné avec le support symbolique physique une mobilité qui lui est propre : le sens spirituel n'anime plus simplement des apparences sensibles statiques, mais des processus sensibles dynamiques qui font un avec le sens qui les anime. L'esprit du lecteur ne reconnaît plus alors simplement la trace d'un esprit, mais la présence d'un esprit lui-même dans la mesure où un processus sensible animé d'un sens spirituel renvoie au sens vivant, au sens présent à lui-même, c'est-à-dire au sens de la conscience.

Que le processus investi d'esprit soit une notion essentielle est un fait bien connu, même s'il n'est pas toujours formulé en ces termes. En effet, ce fait a été clairement mis en évidence par l'expérience du programme ELIZA de J. Weizenbaum (Weizenbaum, 1966). Un système, laissé en libre accès simulait le comportement d'un psychiatre. Pour cela, il reprenait les phrases de l'utilisateur pour les transformer en une question qu'il lui retournait. Si l'utilisateur disait par exemple *Quelle belle journée aujourd'hui*, ELIZA répliquait par *qu'est-ce qui vous fait dire cela ?*. Egalement, si on lui disait *je me sens déprimé*, ELIZA répondait *je suis désolé de savoir que vous êtes déprimé*. ELIZA fonctionnait en repérant des mots-clefs dans les phrases de l'utilisateur et en choisissant une transformation pour produire une réplique. Le plus surprenant est que, malgré les mécanismes rudimentaires à la base d'ELIZA, un tel système a grugé des utilisateurs au point qu'ils refusaient de croire Weizenbaum quand il leur affirmait que ce n'était pas un véritable psychiatre mais un programme rudimentaire qui leur avait donné la réplique pendant qu'ils se confiaient aveuglément.

Ce que prouve un système comme ELIZA, c'est que ce que nous avons appelé « la plausibilité interprétative » dans (Bachimont, 1994) ne repose pas en son principe sur une « plausibilité cognitive ». Cette dernière notion renvoie à la situation dans laquelle le système informatique produit du sens en opérationnalisant un modèle de l'esprit. La plausibilité interprétative renvoie au fait que le système fasse sens, c'est-à-dire qu'il est possible d'attribuer une signification aux processus symboliques (physiques) exécutés par le système. Or, selon les critères des partisans d'une IA forte, un système comme ELIZA ne possède aucune plausibilité cognitive. En particulier, les structures computationnelles mises en œuvre dans ce programme sont incapables de produire des actions intelligentes ou des performances cognitives (cf. section 1-2). Pourtant, la plausibilité interprétative fut manifeste pour ses utilisateurs.

En d'autres termes, la notion d'« investissement d'un sens spirituel » que nous avons empruntée à Husserl ne qualifie pas l'essence réelle de l'objet mais le statut de son apparaître pour l'esprit. Pour qu'il y ait investissement d'un sens spirituel, la question n'est pas de savoir si le système informatique est réellement un esprit ou non, en donnant à ce terme le sens que l'on voudra, mais de savoir si les processus symboliques apparaissent comme étant animés d'un sens spirituel.

6.1.2 DES REPRÉSENTATIONS ORTHOTHÉTIQUES

Comment se fait-il que les livres et les processus symboliques puissent s'investir ainsi de sens ? Cela provient du fait que les symboles consignés dans les livres, les symboles exhibés dans les processus computationels ne sont pas n'importe quels symboles. Ce sont des symboles qui s'annoncent comme possédant d'emblée du sens.

> « Il s'agit partout d'un *mode d'aperception fondamental*, d'une attitude expérientielle tout à fait particulière, dans laquelle un apparaissant sensible (un prédonné) ne devient pas un donné sensible, un perçu, un objet d'expérience, mais dans son « fluidum psychique », dans l'unité même d'une appréhension d'un autre type, contribue à constituer une objectité d'un type propre. Il est manifeste que ce mode d'aperception a, dans ce cas, une « fonction subjective » autre, un mode subjectif autre que celui qu'il a dans le cas d'une expérience thématique externe. » (Husserl, 1982, p. 327)

Husserl convoque, comme symbole possédant d'emblée du sens, tous les objets d'usage et du monde de la culture : il cite pêle-mêle « un verre, une maison, une cuillère, un théatre, un temple, etc. ». Cependant, il est difficile de mettre sur un même plan un livre et une cuillère, même si tous deux possèdent une signification qui excède leur sens en tant qu'objet matériel sensible que peut considérer une attitude naturelle. En effet, l'un contient des représentations structurées symboliquement, alphabétiquement, l'autre non.

La fusion du sens avec son support sensible n'est pas identique dans les deux cas. Dans le premier, celui des caractères orthographiques consignés dans un livre, l'objet se constitue comme étant l'enregistrement exact d'un sens : par essence, l'objet qu'est l'écrit est de renvoyer à un sens dont il donne le contenu exact, tandis que dans le second cas, celui d'un objet d'usage par exemple, l'objet se constitue comme ayant d'emblée un sens, mais sans en révéler la teneur exact, ou le contenu. Il faut donc distinguer parmi les objets investis d'esprit ceux qui s'animent d'un sens spirituel dont ils révèlent le contenu exact, et ceux qui s'animent d'un sens sans dire précisément lequel.

En effet, nous avons déjà noté que la langue naturelle, la parole par conséquent, permettait d'exprimer le rapport au monde, le rapport au contexte, c'est-à-dire la manière dont les choses du monde font sens pour nous dans un contexte particulier. L'écriture, en consignant la parole, fixe dans la lettre ce rapport au monde, et renvoie de manière exacte à un sens. Cela ne contredit pas le fait que nous avons relevé dans notre constat négatif soulignant la contextualité du sens de la langue naturelle. En effet, le fait que ces représentations *prétendent* renvoyer exactement à un sens n'implique pas qu'elles renvoient exactement au même sens quel que soit le contexte. Le contexte de lecture de l'utilisateur détermine l'exactitude des représentations à renvoyer exactement au sens pertinent du contexte, exactement comme la lecture d'un énoncé écrit formulé dans la langue renvoie exactement au sens que le lecteur lui attribue dans son contexte de lecture. Dire que le sens est contextuel, c'est affirmer que le sens n'est pas assignable hors contexte. Mais, en pratique, on est toujours situé dans un contexte. C'est pourquoi l'écriture renvoie de manière exacte à *un* sens, mais non *au* sens : le fait de consigner la parole énoncée en contexte dans une écriture qui devient un objet non contextuel implique que l'écriture prétend renvoyer à un sens unique sans pouvoir dire précisément lequel du fait de la perte du contexte d'énonciation. En d'autres termes, l'écriture se caractérise d'une part par l'exactitude de son enregistrement de la parole et donc par sa capacité à prétendre renvoyer à un sens unique, et d'autre part par son incapacité à dire quel sens exactement.

On comprend ainsi pourquoi les connaissances dans les domaines pratiques ne peuvent s'élaborer et se théoriser que dans la langue naturelle. En effet, seule la langue naturelle permet d'exprimer le rapport au monde dans lequel consiste la pratique. L'écriture alphabétique, en transcrivant *exactement* cette expression, fournit le matériau sur lequel édifier un système de connaissance qui sera permanent et partageable. Mais il restera un fossé irréductible entre la théorie et la pratique. En effet, la transcription exacte du savoir décontextualise la connaissance si bien qu'un travail de recontextualisation est nécessaire ; et il est bien connu que dans les domaines de la pratique, comme la médecine pour reprendre cet exemple, l'enseignement repose pour une large part sur un apprentissage de la contextualisation du savoir tel que l'écriture l'a transcrit. Si les médecins se forment davantage à l'hôpital qu'en faculté, et que les ingénieurs se forment davantage dans les entreprises que dans les écoles, c'est qu'il est nécessaire de savoir constituer un contexte de lecture qui soit propre à assigner le sens correct auquel l'écrit prétend renvoyer de manière exacte.

Comme nous le verrons au chapitre 5, les langages formels de la science ne possèdent ce problème dans la mesure où ils prescrivent d'eux-mêmes comment effectuer la recontextualisation. En effet, dans la mesure où les langages de la science prescrivent comment construire l'objet dont on parle, le rapport au contexte est internalisé dans la théorie. En revanche, la langue naturelle exprime un rapport un contexte que l'écriture consigne sans l'internaliser : il est donc nécessaire, lors de la lecture, de lui redonner un contexte que l'écrit ne détermine pas même s'il contient des contraintes interprétatives ; il est en effet possible de discriminer les contextes pertinents des contextes impertinents de lecture. Mais ces contraintes ne suffisent pas à déterminer quel rapport au monde, quelles choses se donnant dans ce rapport, est exprimé. L'écrit consigne la rencontre passée et vécue avec des choses et un contexte : le lire et le comprendre, c'est produire une nouvelle rencontre qui possède une communauté de sens avec la précédente, une analogie, mais non une identité. L'écrit scientifique consigne la manière de susciter une rencontre, l'expérimentation, avec tel ou tel objet, dont les critères d'identité sont déterminés par la définition même de l'expérimentation. Il ne s'agit pas alors d'analogie entre deux expérimentations, mais d'incertitudes sur les résultats qui permettent de parler d'identités ou de différences.

Si les représentations symboliques de l'IA sont des écritures dynamiques s'investissant spontanément, d'emblée d'esprit, c'est qu'elles partagent avec l'écriture alphabétique des livres d'être exactes dans leur renvoi à un sens. L'écriture alphabétique constitue un cas exemplaire pas tellement parce qu'elle consigne la parole, parce qu'elle est phonétique, mais parce qu'elle la consigne exactement et donc parce qu'elle autorise de retrouver le sens que possède le message oral dont elle est l'enregistrement. Bref, ce qui compte dans l'écriture alphabétique, c'est son caractère ortho-graphique, c'est-à-dire d'être une écriture exacte. Les représentations symboliques de l'IA sont, pour emprunter un concept à B. Stiegler, *orthothétiques* dans la mesure où elles posent exactement, dans une structure matérielle constituant son enregistrement, une signification. Comme le souligne B. Stiegler,

> « *Il faut abandonner ici la compréhension primordialement phonologique de l'écriture alphabétique pour privilégier son caractère ortho-graphique.* [...]
>
> « Le phénomène essentiel de l'écriture orthographique dite phonologique est l'*exactitude* de l'enregistrement de la voix plus tôt (sic) que l'exactitude de l'enregistrement *de la voix* : la question est l'enregistrement plus tôt que la voix. De même, la photographie est un enregistrement exact. » (Stiegler, 1996, p.23-24)

L'orthothéticité est un concept qui souligne l'authenticité, l'exactitude, du support sensible vis-à-vis du sens qui l'anime, et qui font de lui un *enregistrement* fiable de ce sens. Tel objet renvoie exactement à tel sens : la confusion n'est pas permise. L'écriture phonétique n'est qu'une manière parmi d'autres de réaliser des représentations orthothétiques. Comme le souligne B. Stiegler, une photographie renvoie par essence à un « ça a été ». Alors qu'une peinture laisse toute licence quant à l'interprétation de ce qui est représenté, une photographie renvoie à des événements réels dont, par son existence même, elle certifie la réalité. C'est même pour cela qu'une peinture ne peut être truquée, alors qu'une photographie le peut : en jouant sur son orthothéticité et donc sa prétention à renvoyer exactement à un contenu, on peut travestir cette exactitude. La peinture ne prétendant ni l'exactitude ni la fidélité à une réalité originale, il est sans objet de vouloir travestir la signification qui l'anime.

L'écriture dynamique que propose l'IA en manipulant automatiquement des représentations symboliques dont la structure et la terminologie reflètent l'expression linguistique des connaissances qu'elles représentent impose donc l'authenticité du processus quant aux contenus qui lui sont associés. Si l'IA reste un projet technologiquement important malgré l'incompatibilité entre la non formalité de la langue et des connaissances linguistiques et la formalité de son principe de modélisation, c'est que la méthode formaliste permet de manipuler des représentations orthothétiques de connaissances. Orthothétiques voulant dire ici non pas que les représentations sont des connaissances mais qu'elles prétendent, dans leur essence même de représentations symboliques, en être, et qu'on leur accorde toujours cette prétention, indépendamment du fait de savoir si elles en sont vraiment.

Insister sur l'orthothéticité des représentations syntaxico-formelles vis-à-vis des connaissances dont elles sont la retranscription symbolique permet de comprendre pourquoi l'IA peut s'emparer de problématiques comme la « capitalisation des connaissances » ou les « très grandes bases de connaissances » (Lenat *et al.*, 1986; Lenat & Guha, 1990; Lenat & Feigenbaum, 1991) et donc de partager des préoccupations avec l'informatique documentaire : en formalisant l'expression linguistique des connaissances par une structure syntaxico-formelle conservant la forme symbolique des expressions, l'IA est en effet bien placée pour rendre compte des connaissances décrites orthothétiquement par ces supports documentaires que sont les documents écrits. L'IA, sous ce rapport, devient une informatique documentaire qui, prenant explicitement en compte le caractère orthothétique des représentations qu'elles manipulent, définit les processus dynamiques pour qu'ils se conforment au sens qui les investit et permettent ainsi de l'expliciter.

6.2 Effectivité du symbolique

Les systèmes IA sont des auxiliaires cognitifs dans la mesure où, manipulant des représentations orthothétiques, c'est-à-dire prétendant renvoyer à un sens exact dans le domaine, ils font sens.

Mais le système formel automatique ne fait pas seulement sens, il *produit* du sens, et ce, parce qu'il n'est pas seulement un système formel dont on interprète les représentations formelles qu'il permet de dériver, mais parce qu'il s'agit d'un système formel *automatique* dont on interprète le comportement dynamique. En effet, le fait que ce soit le modèle lui même, dont l'interprétation a un sens dans le domaine, qui tourne, permet de mieux saisir le contenu du modèle en tant que modèle, c'est-à-dire ce qu'il affirme à propos de la réalité modélisé. Dès lors, le système formel automatique est un outil de modélisation puisqu'il permet, du fait qu'il actualise le potentiel inférentiel du système formel, de confronter les propriétés du système au monde visé et d'en faire alors un modèle.

On retrouve cette propriété essentielle des systèmes formels automatiques de *coupler l'effectivité du symbolique ininterprété du formel computationnel à l'interprétabilité du symbolique adoptant la signifiance de la langue*, dans les travaux de Marc Linster (Linster, 1993) qui constate que l'entreprise IA de construire des SBC consiste à « on the one hand, *modeling to make sense* of observed behavior and, on the other hand, *modeling to implement systems* to reconstruct comparable problem-solving actions. ». Il est essentiel, à son point de vue, que le système formel qui tourne, c'est-à-dire, le système formel computationnel ininterprété, soit *le même* que le système formel constituant un modèle du domaine : le programme doit être le modèle, et pas seulement son opérationnalisation. En effet, il est difficile, selon lui, de comprendre un modèle si l'on ne dispose pas d'outil permettant de le rendre effectif, et l'opérationnalisation permet par elle-même de donner des informations en retour pour la poursuite de l'élaboration du modèle :

> « The implementation of the behavior with operational languages cannot be separated from the model-construction process, as on the one hand, models tend to become so complex that their behavior can hardly be predicted on the basis of a document and on the other hand, the feedback from a runnning system contains important cues for the ongoing model elaboration. »

Le propos de Linster est donc d'affirmer que l'effectivité du computationnel est nécessaire à l'élaboration du modèle :

> « Furthermore, an operational model provides important feedback for the model-construction process, as complex interactions become transparent and the computer can be used to find ambiguous parts or loop holes in the model. » (Linster, 1993, p.212)

L'intervention de l'artefact informatique comme système formel automatique dans la constitution ou l'élaboration de nouvelles connaissances est encore plus évidente si l'on remarque que les connaissances exprimées en langue ne forment pas nécessairement une théorie du monde ou un modèle du monde. Ce que le système opérationnalise n'est pas un modèle qui préexisterait dans le domaine ou l'esprit des spécialistes. Le système permet d'élaborer un modèle original; Linster souligne ainsi :

> « We consider knowledge acquisition to be a constructive modeling activity that make sense of previously ill-structured or largely not understood situations and behaviors. » (Linster, 1993, p.211)

Mais, M. Linster n'affirme pas, comme nous le faisons ici, que les systèmes formels produisent sens parce qu'ils sont interprétables en langue. En effet, il ne s'interroge pas explicitement sur le lien unissant le modèle à son modélisé, et reprend la conception standard selon laquelle la notion de dénotation permet d'expliciter le lien entre le contenu du modèle et la réalité qu'il modélise. Mais il insiste, comme nous l'avons fait, sur le fait que le système formel automatique qu'est le SBC est à la fois un modèle et un programme dans la mesure où les primitives du langage de modélisation sont opérationnelles, c'est-à-dire sont des primitives d'un langage de programmation. En d'autres termes, les primitives non logiques sont à la fois les briques de la modélisation et le matériau de l'opérationnalisation. Ces primitives possèdent une sémantique dans le domaine – c'est une sémantique dénotationnelle ou référentielle, et une sémantique dans le programme – c'est la sémantique opérationnelle :

> « the knowledge structuring primitives that are used for the model construction are operational constructs of a high-level knowledge representation formalism. They can be interpreted directly, or compiled automatically into lower level languages. »
>
> et plus loin :
>
> supporting « model construction to make sens and to build systems ... can best be realized if the terms for the operational model are used in the making-sense phase too. Additionally, this gives the terms operational semantics. » (Linster, 1993, p.211-212)

On saurait trop insister sur l'importance de cette remarque. Ce que permet de faire l'informatique dans le cadre de l'IA, c'est de permettre d'exprimer dans un format proche du linguistique des connaissances pour ensuite calculer sur les expressions obtenues. En d'autres termes, l'informatique permet de passer d'une représentation exprimant des connaissances à une représentation exploitable par un calcul. Il s'agit en d'autres termes d'une compilation des connaissances où l'empilement des niveaux de langages permet par un jeu de compilations successives de rendre effectif un symbolisme linguistique. C'est cette propriété qui permet d'avoir, mises en œuvre sur un processeur symbolique, des représentations orthothéthiques. Sans cette capacité à traduire les symbolismes entre eux, il est impossible de comprendre pourquoi le symbolisme de la langue peut avoir un rapport quelconque avec les signes ininterpétés qu'une machine de Turing met en œuvre. Evidemment, cela ne signifie pas que la compilation conserve le sens de l'interprétation linguistique. Mais, cela signifie que la compilation assure un changement de format de représentation aux connaissances qui héritent de ce fait d'une sémantique opérationnelle. En d'autres termes, elles possèdent un nouveau contenu du fait de cette compilation. C'est ce nouveau contenu qu'il faut contraindre pour qu'il soit possible d'y voir une opérationnalisation des connaissances et un moyen de traiter ou d'assister automatiquement la résolution de problèmes.

Nous avons terminé à présent notre parcours critique pour cette introduction et il est temps de passer à notre propositions.

7 Artéfacture

Nous avons posé plus haut que pour mener à bien le projet de traiter informatiquement des problèmes sur lesquels seules des connaissances en langue sont disponibles, il fallait mobiliser d'une part un principe de modélisation et d'autre part un principe d'effectivité. En adoptant une perspective formaliste héritée du Cercle de Vienne pour la modélisation et du formalisme hilbertien et turingien pour l'effectivité, l'IA a construit un programme de recherche cohérent et une méthodologie conséquente mais qui achoppent en pratique sur la non formalité de la langue et des connaissances qu'elle permet d'exprimer.

Par conséquent, le formalisme comme principe de modélisation n'est plus tenable. En revanche, le formalisme comme principe d'effectivité reste indispensable. En effet, ce formalisme est en fait la seule théorie de l'effectivité dont on dispose. Cela implique que si l'on veut avoir des modèles des savoirs exprimés en langue qui soient *par principe* effectifs ou opérationnels, il faudra en passer par ce que la théorie de l'effectivité prescrit.

Par conséquent, il s'agit de toujours construire des programmes, puisque ces derniers sont de droit effectifs. Mais, en délaissant le formalisme comme modélisation, on n'affirme plus qu'ils constituent des théories ou explications de l'esprit, comme le faisait par exemple la thèse des systèmes symboliques physiques. Ces programmes opérationnalisent un modèle du savoir exprimé par les connaissances linguistiques sans en être la théorie, de la même manière qu'un programme qui résout numériquement les équations de la chaleur n'est pas une théorie du phénomène physique de la diffusion de la chaleur. Le problème est donc de savoir quel type de modèle il s'agit d'opérationnaliser pour réaliser le programme de l'IA.

Abandonner le formalisme comme modélisation, c'est abandonner en fait l'axiomatisation de l'esprit ou du monde. On pourrait alors lui substituer une mathématisation de l'esprit dans la mesure où une telle approche ne rencontre pas les mêmes difficultés à traiter le holisme sémantique de la langue. C'est d'ailleurs le point fort mis en avant par les partisans du connexionisme comme Smolensky (Smolensky, 1988) : le modèle connexionisme propose de traiter l'esprit depuis une approche mathématique continuiste, en particulier depuis une approche par les systèmes dynamiques. Dans ce cas, le formalisme comme effectivité opérationnalise le modèle mathématique obtenu de l'esprit et on peut considérer que c'est là le moyen de mener à bien le projet de l'IA.

Même si un tel programme de recherche est tout à fait pertinent, ce n'est pas celui qui est adopté ici. Nous avons vu en effet comment le caractère symbolique des systèmes IA était crucial et leur permettait de remplir leur tâche de modélisation. C'est donc dans cette voie que nous voulons persister.

7.1 Artéfacture comme programme de recherches

Comme nous l'avons dit plus haut, l'artéfacture est le programme de recherches que nous proposons pour tenir compte d'une part de l'inadéquation du formalisme comme modélisation pour extraire le savoir des expressions linguistiques de connaissances et d'autre part de la capacité du formalisme comme modélisation à produire des représentations qui, étant toujours symboliques, sont orthothétiques, et donc s'investissent d'un sens pour l'utilisateur. Par ailleurs, le projet de l'IA étant de produire des systèmes effectifs, il importe de conserver un principe d'effectivité et nous avons déjà évoqué les raisons pour lesquelles le formalisme comme effectivité est incontournable.

L'artéfacture repose ainsi sur le fait *empirique* que l'effectivité des pratiques d'un domaine de spécialité ne peut être traduite dans l'effectivité computationnelle en mobilisant comme principe de traduction l'axiomatisation des pratiques. L'artéfacture repose sur la distinction de deux notions fondamentales :

le matériel-régional qui renvoie à l'effectivité d'un domaine pratique et aux savoirs qu'il mobilise ; le constat négatif montre que le matériel-régional ne s'axiomatise pas ; en d'autres termes, l'empiricité d'un domaine ne se laisse pas décrire par des moyens purement formels ou analytiques : par conséquent, pour rationaliser ce qu'il faut appeler la synthéticité du domaine, il faudra faire appel à des principes *a priori* non analytiques, ou encore des principes synthétiques *a priori* ; c'est pourquoi un tel constat négatif renvoie sur des approches transcendantales mobilisant le synthétique *a priori* pour d'une part comprendre l'informatique et d'autre part caractériser les connaissances du domaine [24] ;

le matériel-effectif qui renvoie à l'effectivité computationnelle ; le constat positif montre que le matériel-effectif confère aux symboles qu'il manipule une signifiance les mettant au principe d'une constitution de nouvelles connaissances.

Le point clef est que le matériel-effectif s'investit d'un sens matériel-régional alors que le syntaxico-formel que le matériel-effectif opérationnalise n'est pas un modèle dénotationnel du domaine. Le syntaxico-formel permet en effet de donner une forme symbolique *orthothétique* au matériel-effectif qui ainsi s'investit d'un sens dans le domaine.

Nous avons déjà dit que l'enjeu de l'artéfacture est de pouvoir contrôler la manière dont le matériel-effectif permet au syntaxico-formel symbolique de faire sens. De manière à présent plus méthodologique, l'artéfacture se constitue autour des tâches suivantes :

la spécification d'un modèle régional dont la nature est linguistique ; il s'agit dans ce modèle d'une part de déterminer le sens ou contenu de savoir que les expressions linguistiques des connaissances véhiculent et d'autre part de le déterminer d'une telle manière qu'une représentation syntaxico-formelle pourra en rendre compte ; le

24. Notre seconde partie entreprend cette investigation philosophique.

modèle régional est donc un appauvrissement du sens du domaine effectué en vue de sa formalisation ;

la formalisation du modèle régional doit fournir en un étape suivante une représentation syntaxico-formelle du domaine ; cette représentation est un modèle du domaine dans la mesure où elle peut s'interpréter dans les termes des connaissances du domaine grâce au modèle régional ; le principe est que la représentation formelle mobilise des primitives non logiques dont l'interprétation linguistique est contrôlée par le modèle régional de telle manière que la composition de leur signification dans les représentations formelles selon les règles de composition sémantique du formalisme soit conforme au sens du domaine ; la formalisation du modèle régional fournit donc un modèle formel du domaine, qui modèlise le domaine parce qu'il s'interprète dans les termes du domaine par l'intermédiaire du modèle régional et non parce qu'il dénote le domaine ; le modèle formel, du fait de son formalisme, peut être associé à une caractérisation du domaine sous forme d'objets qu'il dénote ; mais cette caractérisation n'est pas le domaine ; elle n'en est qu'une « caractérisation » dans la mesure où seul le modèle régional permet de la définir ; s'il est possible d'adjoindre au modèle formel une sémantique référentielle dénotationnelle, c'est que le modèle régional permet de déterminer des objets qu'un modèle formel peut dénoter (voir la figure 1.3) ;

l'opérationnalisation du modèle formel consiste enfin à doter le modèle formel d'un contenu ou sémantique opérationnels ; il s'agit que les termes du domaine retenus par le modèle régional pour dégager les connaissances, qui ont reçu une sémantique référentielle dans le modèle formel, reçoivent à présent une sémantique opérationnelle pour qu'ils puissent assumer le rôle de primitives de programmation après avoir été successivement des composantes de modélisation (régional) et des primitives de formalisation (formel).

En résumé, la question que se pose l'artéfacture, pour plagier encore une fois un titre célèbre dans les sciences cognitives, est la suivante : *que doit-être une modélisation des connaissances du domaine pour qu'elle soit compatible avec une opérationnalisation computationnelle, et que doit être une opérationnalisation computationnelle pour prendre en charge le sens d'un modèle du domaine, étant donné que le domaine n'est pas axiomatisable ?*

Un tel programme est en principe *possible* puisque de fait les systèmes IA sont des constructions syntaxico-formels effectives faisant du sens dans le domaine. Le problème est donc de passer du constat de ce fait à une mise en œuvre effective des principes méthodologiques présentés ci-dessus. C'est ce que propose notre traitement de la question des « ontologies » en IA.

7.2 Artéfacture et ontologie en IA

Le but de notre première partie est d'expliciter ces principes en étudiant le problème des « ontologies » en IA. Ce problème est le suivant. Nous avons vu que la méthodologie de l'IA consiste à représenter formellement, *i.e.* dans un langage formel de représentation des connaissances, des connaissances pour les opérationnaliser.

Le fait d'être formel pour un langage consiste dans le fait de reposer sur une syntaxe générative pour la construction de ses formules et une sémantique compositionnelle pour leur signification. Une syntaxe est générative parce qu'à partir d'un nombre fini ou dénombrable de primitives logiques et non logiques elle dispose de règles de construction permettant de les assembler en formules. Les primitives non logiques sont les primitives autres que celles fournies par le formalisme lui-même dont on se dote pour exprimer les connaissances du domaine. Par exemple, il faudra se doter d'une primitive *diagnostic* pour représenter le lien existant entre un tableau clinique (signes et symptomes) et une maladie. Toutes les formules du langage formel sont dès lors construites à partir de ces primitives et exclusivement à partir d'elles. Par ailleurs, les règles de composition sémantique permettent, une signification étant assignée aux primitives, de donner un sens au formule du langage. Les primitives non logiques possèdent un sens dans le domaine qu'il faut fixer, les primitives logiques ayant un sens purement formel ou logique indépendant du domaine.

Par conséquent, on voit que, les primitives étant fixées du point de vue de leur expression dans le formalisme et de leur signification dans le domaine, tout est déterminé quant aux autres formules : comment les construires et les interpréter. Mais, ce que le langage formel ne permet pas de faire, puisqu'il en dépend, c'est de déterminer quelles sont les primitives non logiques et quelle doit être leur signification dans le domaine. *Le problème des ontologies en* IA *est de déterminer le choix des primitives non logiques et de leurs significations* :

> « An ontology defines the basic terms and relations comprising the vocabulary of a topic area as well as the rules for combining terms and relations to define extensions to the vocabulary. » (Neches *et al.*, 1991, p.40)

Cela revient à définir quelles sont les notions de base, fondamentales, d'un domaine, à partir desquelles toutes les connaissances peuvent se construire. En un sens, cela revient à définir quels sont les éléments primitifs, ultimes, du domaine, à partir desquels tout existe. C'est l'ontologie de base du domaine, fixée par les primitives non logiques, à partir de laquelle tout ce qui existe dans le domaine se définit comme une combinaison ou construction des éléments qu'elle spécifie.

Ce problème, important pour l'IA classique, en est aussi un pour l'artéfacture. En effet, nous avons choisi de conserver sur le plan des outils un formalisme symbolique pour, à l'instar des approches classiques (voir plus haut section 1-5 et le chapitre 2), disposer d'une représentation syntaxico-formelle qui soit à la fois effective parce que formelle et expressive (une expression des connaissances) parce que symbolique. Par conséquent, puisque notre système construira ses représentations de manière générative à partir de primitives, il nous faut fixer le choix des primitives et de leur signification pour qu'un utilisateur, en face de ces représentations, leur prête le sens qui convient pour qu'elles l'aident dans ses activités.

Mais le fait de choisir l'artéfacture comme cadre de recherche plutôt qu'une approche « classique », entraîne un traitement différent du problème. D'une part en recourant à une étape, ignorée en IA, de spécification descriptive des connaissances du domaine et d'autre part en mettant en place des instruments de contrôle pour l'investissement de sens par l'utilisateur. En d'autres termes, un utilisateur trouvera, avec le système, des principes interprétatifs ou grille de lecture lui permettant de contrôler lui-même le sens qu'il prête spontanément (les représentations étant orthothétiques) aux inscriptions symboliques de l'écran. Ce que veut faire l'artéfacture, c'est contrôler l'orthothéticité des inscriptions symboliques, leur rectitude ou exactitude, étant donné que ce crédit leur est toujours accordé par un utilisateur.

La première partie traite donc de ce problème des ontologies et montre comment l'artéfacture permet d'en proposer une solution possible. Cette solution correspond à d'une part un système notionnel que nous avons élaboré (Zweigenbaum *et al.*, 1995a; Bouaud *et al.*, 1995) dans un domaine médical, la coronarographie, et d'autre part la station de lecture/écriture que nous avons mis en point pour contrôler son utilisation, c'est-à-dire son investissement de sens. Le système notionnel regroupe environ 2000 concepts médicaux et a été effectivement utilisé dans le cadre du projet européen MENELAS (Zweigenbaum & Consortium MENELAS, 1994; Zweigenbaum *et al.*, 1995b) ; la station de lecture a permis la constitution du système notionnel en permettant en particulier de contrôler le sens des représentations symboliques des notions dans le domaine.

8 Herméneutique matérielle

L'artéfacture est la méthodologie que nous proposons pour l'IA à partir du constat selon lequel l'effectivité symbolique physique des systèmes IA produit du sens dans leur domaine d'utilisation bien que le système syntaxico-formel mis en œuvre ne soit pas un modèle des connaissances du domaine, faute d'être capable de rendre compte de leur rapport au contexte et de leur interdépendance.

Par conséquent, tandis que l'artéfacture se fonde d'une part sur un refus de considérer le syntaxico-formel comme un modèle du sens et des connaissances, elle se fonde d'autre part sur le fait que l'effectivité du symbolique fait sens dans le domaine. Mais si le rapport entre cette effectivité et le domaine n'est pas celui de modèle à modélisé comme dans l'approche « classique » de l'IA, on peut s'interroger sur la nature de ce rapport.

En d'autres termes, si l'artéfacture congédie le formalisme comme philosophie du sens dans la mesure où ses principes méthodologiques sont fondés sur ce refus, elle doit préciser quelle peut être une philosophie du sens compatible avec ces principes et permettant ainsi de les justifier. Le problème posé à une philosophie du sens est en effet le suivant :

> comment comprendre que de l'effectivité symbolique produise du sens dont il n'est pas l'explication scientifique ;

Le problème est qu'il existe un lien entre la matière physique telle qu'elle est au principe de l'effectivité computationnelle et la matière au sens où le matériel est opposé au formel. Quel lien unit le matériel-effectif au matériel-régional ?

Prenons l'exemple de la médecine qui nous a déjà permis d'illustrer plus haut les forces et les faiblesses de la formalisation syntaxique. Il n'est pas possible de considérer cet énoncé médical comme étant un support matériel au sens de la matière physique véhiculant une fonction ou forme indépendante de cette matière. Il faut faire place à une autre dimension de la matérialité de l'énoncé que sa simple dimension physique (au sens du physicalisme). La matérialité de l'énoncé et des signes qui le composent est une matérialité *signifiante* dans la mesure où ce qui est signifié ne peut être considéré indépendamment du signifiant qui le manifeste. Le signe concret n'est pas soumis à des lois calculatoires indifférentes à la matière même du signe, mais à des lois conditionnées par cette matérialité signifiante. En d'autres termes, un signe ne signifie pas simplement parce qu'il est le véhicule indifférent d'une forme, mais parce que sa matérialité permet de l'insérer dans le contexte global d'une pratique dans laquelle il fait sens.

Par exemple, un compte rendu médical d'hospitalisation (CRH) ne fait sens que dans le cadre de la pratique qui le produit : il est incompréhensible sinon, ainsi qu'il est facile de le constater dès qu'on le propose à un quidam n'appartenant pas à la pratique hospitalière. Le fait qu'il consiste en quelques feuillets détermine la manière de le transmettre et donc les critères pragmatiques de son interprétation: rédigé en fin d'hospitalisation, il est destiné à s'insérer au dossier de spécialité en vue des hospitalisations futures. Devant être lu à la « prise en charge » du patient lors d'une nouvelle hospitalisation, il se doit d'être relativement court et rédigé de telle manière qu'une lecture rapide (souvent de l'ordre de 30 secondes (Nygren & Henriksson, 1992)) permette d'en saisir le contenu. Par conséquent, une information médicale consignée sur un CRH n'aura pas le même statut si elle est consignée sur une lettre destinée au médecin traitant : dans le premier cas elle doit déterminer quels examens complémentaires effecture ; dans le second, elle détermine ce qu'il faut expliquer au patient et à quel type de signe le médecin traitant se doit d'être vigilant dans son activité de diagnostic et de prévention.

Par ailleurs, le fait qu'il s'agisse de feuilles papiers réunies en dossier par hospitalisation et rubriques par hospitalisation possède des propriétés très importantes (Nygren & Henriksson, 1992) : l'aspect visuel du papier, selon qu'il est plus ou moins jauni signifie l'ancienneté de l'hospitalisation, elle sera négligée lors d'une prise en charge, mais consultée si un problème se pose et qu'une investigation fouillée du dossier s'avère nécessaire ; l'épaisseur des rubriques donne une indication sur leur complexité : une rubrique fine indique une consultation de routine alors qu'une rubrique épaisse indique que des événements inattendus se sont produits ; lors d'une prise en charge, les rubriques fines sont sautées. Ces éléments « matériels » des dossiers de spécialité sont à ce point nécessaires pour la consultation du dossier que les informatisations du dossier tentent, autant que faire se peut, de reproduire ces informations véhiculées par la matérialité du support papier (Nygren *et al.*, 1992) ; ainsi des icônes proportionnels reproduisent l'épaisseur des dossiers et la lecture informatique reproduit le fait de tourner les pages (paginations des documents) pour que la page « se tourne » sous les yeux de l'utilisateur quand il veut continuer sa lecture.

Ce que soulignent ces exemples tirés du monde médical, c'est que matière, dans le sens matériel-régional, ne signifie pas ce qui est étudié par les sciences physiques, mais ce qui correspond au caractère concret et effectif des pratiques au cours desquelles un esprit ne rencontre pas des formes pures sans matières, mais avant tout des choses sur lesquelles il peut investir sa pratique, c'est-à-dire ses actions et ses objectifs. Il ne s'agit donc plus de la matière au sens physique mais de la matière au sens où Husserl l'emploie pour l'opposer à « formel », dans la mesure où ce terme de « matière » appréhende ce qui relève non de la généralité formelle mais de la spécificité concrète et pratique de la réalité. Autrement dit, la notion de matériel-effectif renvoie à l'effectivité de la physique de la nature alors que la notion de matériel-régional renvoie à l'effectivité de la pratique.

Par conséquent, il s'agit de distinguer une conception physique de la matière d'une conception sémiotique de la matière comme matière signifiante. En opposant la *matière physique* à la *matière signifiante*, il est possible de généraliser le concept de matière dans la mesure où la matière physique devient un cas particulier de la matière signifiante : la matière physique est la matière signifiante du point de vue pratique des sciences physiques.

Le fait que les lois formelles du calcul, c'est-à-dire les lois du matériel-effectif, ne puissent pas reconstruire les régularités du matériel-régional incite à penser que les lois formelles du calcul ne sont qu'un cas particulier de l'effectivité, celle du calcul, c'est-à-dire des processus naturels ou physiques compris comme des processus informationnels.

Puisqu'il faut élaborer des lois « matérielles » décrivant comment la pensée s'investit dans les domaines du réel, il ne s'agit donc plus de décrire les phénomènes cognitifs en termes de traitements de l'information, mais de décrire quelles sont les lois du sens découlant de la matière ou pratique spécifique d'un domaine. Ce n'est que dans un second temps que l'on pourra se poser la question de savoir comment considérer la matière signifiante d'une domaine, étudiée à l'aune de la pratique immanente du domaine qui le constitue comme tel, dans le cadre des pratiques formalistes des sciences de l'information : comme une matière physique véhiculant une forme calculatoire indifférente à son support. Il s'agit désormais de *traduire* le sens des phénomènes cognitifs dégagés depuis la pratique du domaine dans le point de vue formaliste. Ce n'est qu'ainsi qu'il sera possible d'entreprendre une formalisation de l'esprit fondée sur une description des phénomènes compatibles avec la « réalité de l'esprit » que l'on rencontre dès qu'il agit de domaines pratiques et techniques.

En d'autres termes, il faut passer du *formalisme*, prônant une universalité des lois du sens et de l'esprit puisque le sens, indifférent à son véhicule matériel physique, ne relève selon lui que du traitement de l'information, à une *herméneutique matérielle* prônant des lois du sens commandées par la matérialité signifiante d'un domaine.

Il faut faire droit à la matérialité signifiante d'un domaine dans la mesure où la manière dont la pratique d'un domaine investit de sens ses supports de signification – énoncés, outils, etc. – n'obéit pas aux principes édictés par le formalisme. Mais cela prouve qu'il existe donc des domaines qui échappent à la normativité du formel et qui ne peuvent donc s'y soumettre sous peine d'être travestis. Par conséquent, le formel n'est qu'un moyen parmi d'autres d'énoncer des lois sur le sens et il ne constitue ainsi qu'une *région* particulière de la signification, tout comme la médecine ou le droit ; il ne constitue pas la manière privilégiée et canonique d'exprimer et d'étudier les lois du sens.

Le fait que le formel ne soit qu'une région du sens s'inscrit dans la constatation plus générale que la pensée ne peut être universelle dans la mesure où elle emprunte toujours la particularité et la régionalité du support matériel signifiant dans lequel elle trouve sa consistance et sa genèse. Ce que signifie le fait que la naturalisation soit une traduction du sens immanent à une pratique dans un domaine à un sens immanent à une pratique formelle et calculatoire, c'est que la naturalisation n'est qu'un projet de *régionalisation* de la pensée dans le domaine du calcul. Il n'existe que des régions du sens : il n'existe donc rien comme de l'« intelligence générale », mais que des intelligences particulières, des intelligences engagées dans des domaines ou pratiques concrètes.

Mais, traduction trahison. Puisque le formalisme n'est qu'une manière parmi d'autres de construire du sens par une pratique fondée sur le traitement de l'information, il ne peut en épuiser toute la teneur et s'en tient à ce qui est exprimable dans son ordre. Une herméneutique matérielle fait donc place au pluralisme des pratiques qui chacune manifeste une dimension particulière du sens.

La régionalisation du formel permet alors de justifier les principes méthodologiques de l'artéfacture. En effet, selon les considérations précédentes, la formalisation syntaxique des connaissances n'est pas tant une modélisation qu'une régionalisation du sens du domaine (matériel-régional) dans le domaine de l'effectivité computationnelle (matériel-effectif). Le problème est en effet de surmonter l'inadéquation du vocabulaire de la description formelle des phénomènes avec la matérialité signifiante des domaines techniques décrits. En étudiant en détail le problème des « ontologies » dans la première partie, nous montrons comment se doter d'une description sémiotique des phénomènes qui rende compte de leur sens immanent dans le domaine qui les produit. Ce n'est que dans une seconde étape qu'il faut en aborder la formalisation, en distinguant d'une part la traduction de la description sémiotique en un langage formel, produisant ainsi une « description formelle », celle dont part directement l'IA formaliste en la considérant comme la description des phénomènes du sens, et d'autre part l'explication formelle de cette description par son opérationnalisation computationnelle en un programme.

9 Plan de la thèse

Le mémoire comprend deux parties, l'une expérimentale, reflétant notre activité de recherche en intelligence artificielle dans le cadre de l'artéfacture, l'autre théorique, reflétant la réflexion philosophique nous ayant permis de définir ce cadre.

9.1 Artéfacture et ontologie en IA

Le but de cette partie est de traiter le problème des ontologies en IA dans le cadre méthodologique proposé par l'artéfacture. Le premier chapitre analyse l'état de la question en IA en partant du problème que rencontrent la plupart des travaux actuels concernant l'élaboration des SBC, à savoir le partage et la réutilisabilité des composants de connaissances. Deux types de travaux se dégagent qui s'attachent chacun à élaborer des composants réutilisables : les méthodes génériques de raisonnement ou de résolution de problèmes d'une part, les ontologies du domaine d'autre part. La réutilisabilité de ces composants reposent sur deux types d'hypothèses : d'une part l'indépendance de la tâche par rapport au domaine, d'autre part l'indépendance du domaine par rapport à la tâche. Il s'avère que ces hypothèses, reposant toutes deux sur une approche formaliste au sens utilisé ici, ne sont pas vérifiées d'emblée par les domaines pratiques où s'élaborent ces travaux et qu'une modélisation préalable du domaine doit en construire une représentation compatible avec ces hypothèses. Une telle modélisation ne peut être que non formelle étant donné qu'elle vise à produire un résultat compatible avec les présupposés d'une approche formelle, la rendant ainsi praticable. En d'autres termes, il faut se donner les moyens d'avoir des composants réutilisables et c'est pour cela qu'il est nécessaire d'effectuer une modélisation non formelle du domaine. Nous démarquant ainsi d'une approche formaliste où le domaine est décrit dans le cadre d'une ontologie formelle qui décrit la forme du domaine sans en décrire le contenu ou matière, nous montrons la nécessité d'élaborer une ontologie matérielle qui puisse appréhender le contenu spécifique du domaine. Or, poser le problème en ces termes, c'est vouloir ajouter à une description formelle et analytique du domaine une description synthétique et *a priori* du domaine.

Le but du second chapitre est de se donner les moyens méthodologiques d'effectuer une modélisation non formelle du domaine et revient sur la nature d'une connaissance du domaine pour dégager une entité théorique qui puisse rendre compte de son contenu et qui soit univoque et non contextuelle. Cette entité est celle d'« essence différentielle » dont l'ensemble constitue l'« ontologie régionale » du domaine : essence car devant déterminer un contenu objectif, différentielle car ce contenu n'est pas posé en soi comme une essence substantielle, mais par l'intermédiaire d'un système de connaissance où le contenu prend sa détermination par les différences qu'il entretient avec ses voisins au sens du système, en comprenant par système un ensemble d'éléments se définissant mutuellement les uns les autres en fonction des relations les unissant. Définie de cette manière, l'ontologie du domaine coïncide pour une large part avec la notion d'ontologie matérielle dégagée par Husserl. Nous explicitons cette coïncidence en développant cette notion husserlienne et la situant par rapport aux approches formalistes.

Le troisième et dernier chapitre de cette partie expose notre méthodologie pour l'élaboration d'ontologies en intelligence artificielle. Il en montre les options théoriques et philosophiques à l'aune du second chapitre et explicite les conséquences pratiques dans le cadre IA présenté par le premier chapitre. Nous explicitons notre méthodologie dans le contexte d'une ontologie de connaissances médicales et retraçons les différentes étapes de constitution : à partir de la verbalisation des connaissances, en d'autres termes, à partir d'un corpus, il faut constituer l'ontologie régionale du domaine, correspondant à la théorie de la spécialité, l'ontologie formelle correspondant à un traitement logique de l'ontologie régionale, et l'ontologie computationnelle correspondant à l'utilisation opérationnelle de l'ontologie formelle dans le programme constituant le SBC. Nous montrons comment établir ces ontologies et marquons leurs différences et oppositions.

9.2 Artéfacture et herméneutique matérielle

La seconde partie traite de l'herméneutique matérielle, notre tentative de refléchir sur le fait fondant l'artéfacture, à savoir que l'ordinateur permet de constituer des connaissances. Cette partie comprend également trois chapitres.

Le premier chapitre a pour objectif d'élaborer une philosophie de l'informatique qui permettent de comprendre ce que signifie l'effectivité que l'on y rencontre. Nous adoptons pour cela une perspective transcendantale en reprenant la philosophie kantienne pour élaborer une épistémologie transcendantale de la logique et de l'informatique, sur le modèle des rapports liant les mathématiques à la physique.

La principale conclusion de ce chapitre est reconnaître que la logique et l'informatique repose sur du synthétique *a priori*. Pour reprendre une distinction husserlienne, cela implique le caractère « régional » et non « formel » de la logique mathématique. Alors que la logique formelle est analytique, la logique mathématique et l'informatique qu'elle fonde sont synthétiques. En d'autres termes, la logique mathématique n'est pas « formelle », mais « matérielle », au sens du matériel-régional que nous avons vu plus haut. L'informatique est dans ce cadre une science naturelle, une physique des signes au sens de calcul effectif sur des symboles, dont la logique mathématique constitue l'outil et contient les principes synthétiques *a priori*. Ce chapitre permet donc de donner un statut épistémologique précis à la notion d'effectivité et à l'informatique en affirmant son caractère de science naturelle. Si l'informatique dévoile un donné, c'est qu'elle n'est pas analytique, mais qu'elle est une science expérimentale. Nous avons ainsi tenté de donner un sens philosophique à la vue profonde de Newell et Simon que nous avons rappelée plus haut.

Qu'il y ait un donné qui surgisse avec l'informatique, qu'il y ait un choc de l'expérience, et que cette expérience possède ses principes synthétiques *a priori*, voila ce que le chapitre 5 a montré pour l'informatique. Mais si cela rend compte du fait qu'il puisse y avoir une élaboration de la science informatique, cela n'explique pas que le nouveau donné donnent lieu à de nouvelles connaissances au niveau du domaine dans lequel il s'interprète. C'est la raison pour laquelle le second chapitre aborde le problème de l'investissement de sens dont nous avons noté l'importance dans notre constat positif. Il s'agit de comprendre ce que peut vouloir dire le fait d'adopter l'ordinateur comme outil de modélisation de connaissances dans un domaine alors que le domaine ne possède en soi une structure formelle. Nous adoptons pour cela une perspective transcendantale en mobilisant la phénoménologie husserlienne qui, nous semble-t-il, nous donne les instruments conceptuels nécessaires pour penser ce problème. Par ailleurs, comme nous le rappelerons, la phénoménologie husserlienne est souvent rapprochée du projet de l'IA comme axiomatisation de l'esprit, si bien qu'elle est en quelque sorte un parcours obligé dès que l'on s'intéresse au problème du sens que peut posséder un artefact informatique.

Le problème est de savoir s'il peut y avoir une « algorithmisation de la phénoménologie » comme Husserl s'est interrogé sur la possibilité d'une « géométrisation de la phénoménologie ». En d'autres termes, la question est de savoir si les processus symboliques dont est le siège un ordinateur ne s'interprètent pas spontanément comme reflétant la démarche d'un esprit tout simplement parce que l'esprit procède réellement ainsi. Assimilant le formalisme husserlien avec ce projet d'une algorithmisation de la phénoménologie, les lectures cognitivistes contemporaines de la phénoménologie husserlienne voient en Husserl le précurseur de la psychologie cognitive et répondent donc affirmativement à cette question. Ces lectures sont discutables dans la mesure où le formalisme husserlien n'est pas une annexion au domaine formel des domaines matériels, dont fait partie la phénoménologie : l'applicabilité du formel comme logique de la pensée aux domaines matériels en tant que contenus pensés ne revient pas à affirmer la formalité des domaines matériels.

Mais, en affirmant la régionalité de la logique, à l'instar de la géométrie, la question peut être posée avec un sens phénoménologique pertinent : en quoi la matérialité de la logique est-elle commensurable avec la matérialité des connaissances matérielles des autres domaines ? Quel rapport y-a-t-il entre la matérialité du formel et la matérialité du vécu constituant les connaissances ? Analysant la structure du vécu à l'aide des *Leçons sur la conscience intime du temps*, nous affirmons leur incommensurabilité et concluons sur le dépassement de la phénoménologie du vécu par une sémiotique.

Le troisième et dernier chapitre de cette partie a pour objectif de tirer partie des conclusions sémiotiques tirées dans le chapitre précédent. Il s'agit de comprendre le fait qu'un signe externe à la conscience est nécessaire pour qu'elle puisse constituer un sens. Cela nous conduit à affirmer la nécessaire médiation de l'objet externe ou transcendant dans la constitution du sens. Affirmer que l'objet externe intervient dans la constitution du sens, c'est affirmer qu'il n'est pas un support neutre mais qu'au contraire il prescrit sans déterminer le sens constitué. Or, il nous semble que cette médiation de l'objet est celle de l'objet technique, dans la mesure où nous comprenons la technicité comme le fait qu'un objet prescrit par sa structure technique l'obtention de tel ou tel résultat. Il s'agit là d'objets techniques particuliers dans le mesure où le résultat obtenu est la constitution d'un sens. En d'autres termes, le signe est un objet technique permettant comme résultat de constituer du sens. On retrouve alors la notion de matière signifiante que nous avions évoquée plus haut (section 1-8). Il s'agit là du thème central de ce que nous appelons « herméneutique matérielle », en comprenant sous ce terme une théorie du sens fondée sur la prescription interprétative que véhiculent les outils techniques du sens, les signes. Nous brossons alors quelques considérations liminaires ce que pourrait être une philosophie du sens formulée dans ce cadre et concluons sur le fait que l'artéfacture est tout simplement l'herméneutique matérielle appliquée au cas où la matière signifiante dont il s'agit en IA est le signe ininterprété de l'informatique. En d'autres termes, l'artéfacture est la théorie de la constitution des connaissances dont la médiation technique est l'informatique.

Première Partie

Artéfacture et ontologies en intelligence artificielle

Le problème de l'interprétabilité linguistique des modèles

CHAPITRE II

ONTOLOGIE ET INTELLIGENCE ARTIFICIELLE

L'objectif de ce chapitre est de présenter l'état de l'art en IA concernant la problématique des ontologies. Cet état de la question nous permettra de préciser en quoi l'approche formaliste de l'IA est inadéquate dans le cas de domaines techniques concrets : supposant en effet que le monde est une extension d'individus, l'approche formaliste se heurte au fait qu'il n'existe rien dans un domaine qui puisse être d'emblée, avant toute modélisation, être caractérisé comme un objet du domaine. Nos propositions dans le cadre de l'artéfacture consisteront à introduire une étape de modélisation explicite des objets du domaine par un travail linguistique sur les objets du discours. Ainsi, les hypothèses du formalisme, illégitimes pour aborder directement la réalité concrète d'un domaine, sont réalisées après cette modélisation et les outils du formalisme peuvent alors être repris de plein droit.

Mais avant d'en arriver aux solutions, exposons le problème des ontologies. Le plan du chapitre est le suivant. Dans un premier temps nous rappelons quels sont les principaux problèmes méthodologiques rencontrés dans la conception des SBC (section 2-1). Puis, nous présentons la méthodologie KADS qui est conçue pour répondre à ces problèmes (section 2-2). Mais, alors que KADS se concentre essentiellement sur les méthodes de résolution, il apparaît qu'un effort méthodologique semblable est nécessaire dans la modélisation non pas des méthodes de résolution, mais dans la modélisation des domaines dans lesquels ces méthodes sont mises en œuvre. Cet effort est celui qui correspond à la problématique des ontologies (section 2-3). Nous énonçons ensuite les hypothèses (section 2-4) sur lesquelles reposent ces travaux pour montrer en quoi un domaine concret comme la médecine ne les vérifie pas. En d'autres termes, le travail méthodologique effectué, certes nécessaire et bien fait, n'est pas encore suffisant. Nous concluons (section 2-5) sur la nécessité de se doter d'un outil non formel permettant de traiter la non formalité du domaine pour produire une caractérisation, le « modèle régional », qui fournit aux travaux formels décrits dans ce chapitre la base compatible avec leurs hypothèses pour pouvoir s'appliquer.

1 LES PROBLÈMES DES SBC

Les travaux actuels concernant le développement des SBC se situent, comme nous l'avons montré dans l'introduction, dans le cadre de l'IA constructiviste : une base de connaissances est un modèle construit du problème. En d'autres termes, l'IA constructiviste se place dans un contexte d'ingénierie où il faut modéliser les problèmes pour automatiser leur résolution.

Dans le cadre d'une ingénierie des connaissances, il est crucial de pouvoir développer des méthodes et des outils génériques, c'est-à-dire réutilisables ou adaptables d'un problème à un autre. Or, et c'est là l'origine de tous les travaux dont nous allons parler dans ce chapitre, il est nécessaire, pour chaque système développé, de construire « *from scratch* » (Neches *et al.*, 1991) les bases de connaissances. Cela a pour fâcheuse conséquence que les bases de connaissances restent de taille modeste : au lieu de consacrer l'effort sur les parties du domaine qui ne sont pas encore modélisées/représentées, il faut refaire un travail déjà fait dans d'autres systèmes. En d'autres termes, il n'existe pas réellement d'ingénierie à base de connaissances qui, pour un problème donné, sait réutiliser et adapter une base et une méthode de raisonnement déjà élaborées. L'ambition est par conséquent de trouver quels sont les principes permettant de réaliser des *ressources* réutilisables pour la conception des SBC.

Cela a pour principale conséquence d'interdire la constitution de très grandes bases de connaissances, très grandes signifiant plusieurs centaines de milliers, voirs de millions de connaissances. Or, tout problème réel, c'est-à-dire tout problème qui n'est pas un cas d'école ou une reconstruction de chercheur, mobilise pour sa résolution par une approche IA de très grandes bases de connaissances (Lenat & Guha, 1990; Hobbs & Moore, 1985). L'enjeu est donc considérable puisque l'IA y joue ni plus ni moins sa capacité à se constituer comme une ingénierie, c'est-à-dire à remplir l'objectif qu'elle s'est assigné selon la caractérisation que nous avons retenue dans notre introduction (section 1-2).

Si les bases de connaissances ne sont pas réutilisables, cela tient la plupart du temps aux mêmes types de problèmes. Cela tient en effet au formalisme de représentation des connaissances : des considérations d'efficacité computationnelle, de disponibilité de langage, etc., impliquent le choix d'un formalisme particulier de représentation. Le problème est alors qu'il n'est pas possible de réutiliser des connaissances représentées dans d'autres formalismes, à moins de disposer de traducteurs permettant de passer d'un formalisme à un autre. C'est ce problème que s'efforce de traiter l'initiative *Interlingua* qui a développé un langage pivot KIF dans lequel tout langage de représentation peut être traduit et réciproquement (Ginsberg, 1991).

Mais si une telle initiative permet de régler le problème au niveau de l'expression des connaissances, elle ne le permet pas au niveau de leur mise en œuvre dans un programme et de leur signification dans le domaine. En d'autres termes, pour réutiliser une connaissance, il faut savoir ce qu'elle peut signifier pour une méthode de résolution de problème, par exemple si elle peut jouer le rôle de « signe » ou de « diagnostic » dans une méthode de résolution développée pour réaliser une tâche diagnostic. Par ailleurs, il faut également savoir ce que peut signifier cette connaissance dans un domaine : le problème est en effet de savoir si, pour une tâche donnée, alors que l'on a besoin de certaines connaissances du domaine, des représentations déjà élaborées ailleurs rendent compte de ces connaissances et par conséquent si on peut les réutiliser.

Ces deux derniers problèmes correspondent à deux types de travaux. Le premier concerne l'élaboration de méthodes génériques de résolution de problèmes : il s'agit, pour une tâche donnée comme le diagnostic, de déterminer une manière standard ou générique de la traiter. Ces méthodes génériques précisent de quelles connaissances du domaine elles ont besoin et quels rôles elles leur font jouer dans le déroulement du raisonnement. Par exemple, la méthode diagnostic établit qu'il lui faut des signes et des connaissances associant signes et maladie pour inférer des premiers vers les secondes. Ce courant est très actif et a donné lieu à de nombreux travaux. Notons par exemple (Chandrasekaran, 1986; Chandrasekaran, 1990; Chandrasekaran & Johnson, 1993), (McDermott, 1988), (Steels, 1990; Steels, 1993) et les travaux relevant de la méthodologie KADS que nous présentons dans la section suivante.

Le second type de travaux motivés par le problème de la réutilisabilité concerne l'élaboration de la définition générique des primitives d'un domaine, c'est-à-dire son ontologie. Alors que les travaux précédents concernant la définition de méthodes génériques de résolution dégagent une *ontologie du raisonnement* ou *ontologie de la tâche*, les travaux de ce dernier type concernent l'*ontologie du domaine*. En d'autres termes, d'une part il faut dégager les objets du raisonnement, leur signification et contenu correspondant à leur rôle dans le raisonnement, d'autre part il faut dégager les objets du domaine, leur signification et contenu correspondant à la réalité ou rôle dans le domaine. Alors que les objets du raisonnement se définissent indépendamment du domaine, les objets du domaine se définissent indépendamment de la tâche. C'est la raison pour laquelle ces deux types de travaux sont naturellement complémentaires et correspondent à deux dimensions de la réutilisabilité des composants des SBC :

> « Many authors in the field of knowledge acquisition for knowledge-based systems (KBS) have emphasized the importance of reusable components to reduce the effort required for KBS development. Two main types of components have been identified as potentially sharable and reusable : (i) problem-solving methods, abstract descriptions of the steps that must be taken for perform a particular task, and (ii) domain ontologies, abstract description of the structure of domaine knowledge in a particular field. » (Van Heijst *et al.*, 1995, p.227)

Le courant « ontologique » est également très actif, bien que plus récent. On peut citer parmi les nombreux travaux de ce champ théorique, outre les travaux de KADS dont nous allons parler, l'initiative CYC (Lenat & Guha, 1990) qui vise à établir toutes les théories sur le monde nécessaire à la compréhension et au traitement d'informations énoncées dans la langue naturelle, c'est-à-dire le sens commun, ainsi que les travaux portant sur l'ontologie formelle, essentiellement dus dans le cadre IA à Nicola Guarino, que nous évoquons plus bas.

Nous allons dans ce qui suit considérer ces deux types de travaux. Nous verrons les méthodes génériques à travers une méthodologie particulière, la méthodologie KADS, pour mieux en cerner les principes et les présupposés. Cela nous est particulièrement utile dans la mesure où cette méthodologie articule en son sein ontologie de la tâche à l'ontologie du domaine. KADS repose sur une hypothèse fondamentale selon laquelle les raisonnements menés dans le cadre d'une tâche sont indépendants du domaine sur lequel le raisonnement porte. Philosophiquement, cette hypothèse méthodologique renvoie à une hypothèse concernant la formalité des raisonnements : la logique formelle contient les lois pures du raisonnement et de la pensée qui valent par conséquent pour tous les domaines où la pensée raisonne. Il s'agit d'un principe philosophique qui affirme qu'un domaine ne contient pas en lui-même de principes de raisonnement, mais seulement des faits sur lesquels s'appliquent les lois formelles du raisonnement. En termes philosophiques, cela signifie que le domaine ne contient pas de principes synthétiques *a priori* mais seulement des vérités empiriques synthétiques *a posteriori* sur lesquelles les lois analytiques (et donc *a priori*) de la logique s'appliquent.

C'est en voulant se dégager du point de vue de la tâche que la problématique des ontologies du domaine s'est constituée de manière autonome. Cette problématique resose sur l'hypothèse méthodologique que le domaine est indépendant de la tâche, ou, à tout le moins, qu'il possède des caractéristiques indépendantes de la tâche. Nous verrons que cette hypothèse méthodologique renvoie à l'hypothèse philosophique selon laquelle le domaine consiste en une extension d'individus dont une axiomatisation peut saisir les relations mutuelles.

Ces deux principes philosophiques, ainsi que leurs conséquences méthodologiques, à savoir l'indépendance de la tâche par rapport au domaine et l'indépendance du domaine par rapport à la tâche, ne tiennent pas dans les domaines pratiques si bien que les méthodes de résolution et structures ontologiques flottent sur les domaines sans trouver les objets sur lesquels prendre prise. Notre travail consistera à montrer comment modéliser un univers d'individus permettant d'appliquer l'acquis méthodologique de l'approche formaliste.

2 Partager les méthodes : KADS

La méthodologie KADS[1] occupe une place de choix parmi les travaux de l'IA comme modélisation constructive dans la mesure où elle met en place une conception homogène et unifiée de la conception de SBC. KADS est une méthodologie conçue pour l'ingénierie des connaissances en général, et l'acquisition des connaissances en particulier. L'ingénierie des connaissances se définit comme « le processus global de construction des SBC » (Schreiber *et al.*, 1993) tandis que l'acquisition des connaissances a le sens plus restreint de caractériser les parties du processus de construction qui constitue la modélisation de la démarche de résolution de problème dont le système sera chargé. L'acquisition des connaissances est l'étape de la conception du SBC au cours de laquelle sont définies les connaissances que le système devra opérationnaliser.

Conformément aux principes de l'IA constructiviste (voir le chapitre précédent), la méthodologie d'acquisition des connaissances KADS repose sur deux hypothèses essentielles : l'acquisition des connaissances est une activité de *modélisation* qui se situe au *niveau des connaissances*. KADS considère en effet que « une base de connaissance n'est pas un récipient rempli avec des connaissances extraites d'un expert, mais un modèle opérationnel qui exhibe un comportement désiré donné observé ou spécifié dans les termes de phénomènes du monde réel » (Schreiber *et al.*, 1993, p.3). En voulant modéliser un comportement défini « dans les termes de phénomènes du monde réel », KADS adopte une vue gouvernée par l'objectivité du domaine et non par la subjectivité de l'expert. S'il est bien évidemment possible de partir des explicitations proposées par l'expert de sa démarche, ce n'est pas elle qui fait l'objet de la modélisation. En effet, ce qui doit être modélisé, c'est le comportement du futur système dans son contexte d'utilisation au sein d'une organisation, qui ne

1. KADS est un acronyme de Knowledge Analysis and Documentation System, plus tard transformé en Knowledge Analysis and Design Support. En fait, le terme « KADS » est utilisé aujourd'hui la plupart du temps comme un nom propre et ne possède plus de signification précise (Schreiber *et al.*, 1993, p. xi, note). On distingue deux générations dans les travaux concernant cette méthodologie : ceux relevant d'une première mouture baptisée KADS-I et ceux relevant d'une mouture plus récente baptisée Common KADS. Etant donné que ces versions sont très proches, nous parlerons indifféremment de KADS, ne précisant que lorsque le contexte l'exige si nous nous référons à KADS-I ou à Common KADS. Cette dernière a été peu publiée et la seule référence à laquelle nous nous reporterons est (Wielinga *et al.*, 1993). Sinon, les autres références citées portent sur KADS-I.

correspond pas forcément à l'activité déjà existante d'un agent humain.

Vouloir modéliser dans les termes du domaine pose le problème de savoir à quel niveau d'abstraction il faut modéliser. En effet, si l'on adopte le point de vue informatique de la conception du SBC, on obtient une spécification de la connaissance à mettre œuvre proche de leur implantation effective dans le système. Le danger lié à une telle manière de procéder est de négliger ce qui importe en termes de connaissances au profit de ce qui relève de la réalisation informatique : on est par exemple incapable de constater que deux programmes différents sont identiques du point de la tâche à réaliser dans le domaine. Or, ce qui importe, est le comportement du système dans les termes du problème posé, et non son comportement du système dans les termes de sa réalisation informatique. Ce qui compte, ce n'est pas comment le programme est écrit, mais ce qu'il fait.

La notion de « niveau des connaissances » est destinée à donner le bon niveau d'abstraction lors de la modélisation des connaissances. Elle permet de caractériser les SBC à construire indépendamment de leur implantation informatique, mais uniquement en tenant compte de ce qu'ils font dans le monde réel. En d'autres termes, au lieu de considérer la syntaxe computationnelle permettant de construire effectivement un système, on commence par s'intéresser à la sémantique de ces systèmes, donc de cette syntaxe, dans le monde réel.

La notion de niveau des connaissances prescrit donc de modéliser de manière sémantique les systèmes pour ensuite seulement considérer l'opérationnalisation syntaxique du modèle sémantique. L'opérationnalisation syntaxique relève de l'ingénierie informatique qui en tant que telle est autonome par rapport au modèle sémantique. Ce dernier est donc la spécification sémantique du système et constitue le cahier des charges de la réalisation informatique.

Comme nous l'avons vu dans notre introduction, c'est à Newell que l'on doit, en 1982, un article fondateur (Newell, 1982) proposant le concept de niveau des connaissances. En spécifiant le système en termes de connaissances et non en termes de programmes, par ce qu'il fait et non par la manière dont il le fait, Newell entendait trouver le niveau d'abstraction où ce qui était différent jusqu'alors pouvait être rendu identique. Mais c'est à Clancey (Clancey, 1985) que l'on doit de voir dans le niveau de connaissances le niveau auquel l'élaboration des SBC et des connaissances qu'ils mettent en œuvre doit s'effectuer. Il fournit le « principe du niveau des connaissances » (*Knowledge-level principle*) sur lequel KADS (et surtout Common KADS) fonde le développement des SBC :

> « The knowledge level is the appropriate level for modelling the competence of knowledge-based systems. It calls for the description of problem solving behavior at a conceptual level that is independent from representation and implementation decisions. » (Wielinga *et al.*, 1993, p.303)

C'est à ce niveau que doit s'effectuer la spécification des connaissances utiles pour le SBC, et la reconnaissance des différents types de connaissances concernés.

2.1 Les modèles de KADS

KADS, pour spécifier la conception d'un SBC, envisage une multitude de modèles qu'il faut construire pour rendre compte de la variété des points de vue qu'il faut mettre en œuvre lors d'une telle conception. KADS distingue en effet :

le modèle organisationnel : comme son nom l'indique, ce modèle rend compte du contexte organisationnel dans lequel se situe l'expertise ; il décrit comment l'introduction d'un SBC dans l'organisation peut l'influencer et la modifier harmonieusement ; en considérant l'impact organisationnel de la conception même du SBC, il s'assure que l'explicitation des savoirs et leur modélisation ne vont pas résoudre d'elles-mêmes la tâche du SBC, ou modifier l'organisation d'une manière telle que le problème ne peut plus se poser dans les mêmes conditions ;

le modèle de l'application : ce modèle définit le problème que le SBC doit résoudre dans l'organisation et la fonction qu'il y aura. Ce modèle doit également prendre en compte les contraintes externes permettant le bon fonctionnement et développement du SBC;

le modèle de tâche : ce modèle spécifie comment la fonction du SBC peut être remplie en exécutant un certain nombre de tâches ;

le modèle d'expertise : ce modèle spécifie les connaissances nécessaires qu'un SBC doit opérationnaliser pour être capable d'avoir le comportement recherché de résolution de problèmes ; ce modèle « remplit le rôle d'une *spécification fonctionnelle* de la partie de l'artefact consacrée à la résolution de problèmes » ; mais « ce n'est pas un modèle cognitif d'un expert humain » ;

le modèle de coopération : ce modèle spécifie la coopération entre les différents agents pour mener à bien la réalisation de la tâche confiée au système ; notamment il spécifie la coopération qui s'installe entre le SBC et son utilisateur : les actions qu'ils doivent respectivement accomplir pour la résolution du problème ;

le modèle conceptuel : c'est la réunion des modèles d'expertise et de coopération ; le modèle conceptuel spécifie le comportement du l'artefact à construire ;

le modèle fonctionnel : ce modèle décrit les techniques de représentation et de calcul qui doivent être utilisées pour formaliser le modèle conceptuel. C'est ce dernier modèle qui est effectivement implémenté pour donner le SBC final.

Le modèle qui a reçu le plus d'attention et celui pour lequel la méthodologie de spécification a été spécialement affinée est le modèle conceptuel, et plus encore sa composante modèle d'expertise. C'est désormais lui qui va retenir notre attention.

2.2 Le modèle d'expertise

Le modèle d'expertise a pour fonction de spécifier les connaissances utiles à la résolution du problème confié au SBC. Un modèle d'expertise comprend classiquement quatre couches, mais seules les trois premières sont explicitées en pratique :

la couche domaine : elle contient les connaissances spécifiques à une application donnée ; elle décrit les concepts du domaine et leurs relations ;

la couche inférence : elle décrit l'utilisation des connaissances du domaine ; elle consiste dans la décomposition d'une tâche en inférences primitives ; l'enchaînement des inférences primitives, appelé « structure d'inférence », décrit les chemins d'inférence possibles dans le domaine ;

la couche tâche : elle contient les buts et les méthodes qu'il convient de mettre en œuvre pour les satisfaire, qui correspondent à la description du SBC comme agent rationnel ; elle consiste en une décomposition tâche / sous-tâches, et en un contrôle sur l'exécution de ces sous-tâches ;

la couche stratégie : jamais définie en pratique, cette couche devrait spécifier comment sélectionner les buts en fonction du contexte courant de résolution du système.

Cette catégorisation en couche des différentes connaissances nécessaires au SBC repose sur le fait qu'il faut distinguer la connaissance du rôle qu'elle peut jouer pendant un raisonnement. Cette distinction renvoie à l'opposition classique en IA entre les connaissances du domaine d'une part et les connaissances de contrôle d'autre part. Mais KADS distingue parmi les connaissances de contrôle différentes catégories, les inférences, les tâches et les stratégies, qui sont suffisamment indépendantes les unes des autres pour être organisées en couche dont on affirme qu'elles sont en interaction limitée.

Chaque couche correspond à une vue abstraite de la précédente dont elle complète la description en spécifiant certaines des caractéristiques qu'elle possède dans le raisonnement que le SBC doit mettre en œuvre. La couche domaine n'étant une vue abstraite d'aucune autre couche, elle ne reflète pas *a priori* la tâche du système et son raisonnement. Dans cette mesure, elle sera dite indépendante de la tâche, neutre par rapport à la forme du raisonnement dont elle sera la matière.

Common KADS affine sensiblement ces distinctions en ne considérant plus le niveau stratégique comme une couche du modèle d'expertise, mais comme une catégorie de connaissances portant sur la résolution de problème proprement dite. Ces connaissances, appelées « connaissances de résolution de problèmes » (*problem solving knowledge*) portent sur la manière d'organiser et de lier les tâches, inférences et structures du domaine dans le cadre d'un environnement donné. Common KADS distingue dans ce type de connaissances deux catégories :

- les *méthodes de résolution de problèmes* (*problem solving method*) qui rendent compte d'une manière possible d'organiser les connaissances du domaine, d'inférences et de la tâche ; son application engendre les structures de connaissances permettant l'usage raisonné des connaissances qui entraîne un comportement pertinent de résolution de problèmes ;

- les connaissances *stratégiques* (*strategic knowledge*) proprement dites qui spécifient quel type de modèle est adéquat pour un problème donné.

Le point important que ces connaissances de résolution de problèmes (méthodes et stratégies) ne sont pas une autre catégorie de connaissance venant s'ajouter aux trois premières couches du modèle d'expertise, mais constituent plutôt des « méta-connaissances » dans la mesure où elles portent sur l'usage des connaissances du modèle d'expertise dans le traitement d'un problème donné. Par conséquent la catégorisation entre méthodes et stratégies est orthogonales à la catégorisation en domaine, inférence et tâche. Il est alors possible d'appliquer cette dernière catégorisation aux connaissances de résolution ; mais cela est rarement nécessaire et la distinction méthodes / stratégies est suffisante.

2.2.1 La couche domaine

Les connaissances du domaine correspondent à une conceptualisation du domaine en vue d'une application particulière qui résulte en une théorie du domaine. La conceptualisation s'effectue à l'aide d'un langage de description construits à partir de *primitives épistémologiques* empruntées à Kl-one (Brachman & Schmolze, 1985). Ces primitives sont les suivantes :

concept : les concepts constituent la pièce centrale des connaissances du domaine. Un concept est essentiellement le nom d'une classe d'individus ;

Propriété / Valeur : les concepts possèdent des propriétés. Les propriétés se définissent par leur nom et une description des valeurs qu'elles peuvent prendre. ;

relation entre concepts : ces relations précisent les rapports entre les classes d'individus du domaine, comme par exemple les relations *est-un* ou *partie-de* ;

relation entre expressions valuées : ces relations précisent les rapports entre des propositions sur les valeurs des propriétés. Ce sont typiquement les relations de causalité :

si la sténose est serrée de 80 %, alors le territoire d'aval est sous-vascularisé

structure : les structures représentent les objets complexes : elles contiennent des concepts d'objets et des relations entre ces concepts et des propositions à leur sujet.

La principale préoccupation des concepteurs de Kads est de fournir aux utilisateurs de la méthodologie des moyens d'expressions suffisants pour expliciter les connaissances du domaine. Les primitives permettent de spécifier *le schéma du domaine* qui correspond à la notion de signature en logique : on donne une liste des symboles tirés du vocabulaire du domaine en précisant de quelle catégorie de primitive ils relèvent : « caillot » est un symbole de concept, « sténose » également, « diamètre » est un symbole de propriété, etc.

Le schéma du domaine correspond à l'un des sens que peut posséder le mot « ontologie » en IA : mais les concepteurs de Kads préfèrent la notion de schéma pour insister sur le fait que, dans leur point de vue, la théorie du domaine résulte d'une activité de conceptualisation constructive dans le cadre de l'ingénierie des connaissances, et ne reflète pas une structure universelle inhérente au domaine, quel que soit le point de vue de modélisation adoptée. Cette prudence terminologique est relativement rhétorique dans la mesure où ces mêmes auteurs affirment que « les connaissances du domaine peuvent être considérées comme relativement neutres par rapport à la tâche, *i.e.* elles peuvent être représentées dans une forme indépendante de son utilisation par de quelconques actions de résolution de problèmes » (Schreiber *et al.*, 1993, p. 25-26).

Common Kads complexifie sensiblement le propos en introduisant de nombreuses distinctions complémentaires dans la définition de la couche du domaine. De manière analogue à la notion de schéma du domaine, Common KADS introduit la notion d'« ontologie », qui est caractérisée ainsi :

> « An ontology for describing a particular domain contains the terms refering to the entities that must be distinguished as categories in the domain and the terms that refer to the relations that can hold between (instances of) the entities. Entities can be primitive, i.e. entities that exist and can be recognized in the domain, or can be abstracted, i.e. terms refer to abstractions of complexes of domain entities or to abstract concepts which have no direct correspondence in the domain. »

On retrouve la définition que nous avions déjà donnée des ontologies : la définition de la signature fonctionnelle et relationnelle du système formel modélisant le domaine avec leur interprétation en termes de « catégories » et de « relations » du domaine. Ce sont nos primitives non logiques. Mais Common KADS introduit une distinction supplémentaire fort intéressante entre entités « primitives » et entités « abstraites » selon qu'elles existent dans le domaine ou en sont abstraites.

Cette distinction est intéressante dans la mesure où elle permet de déterminer ce qui, dans la description du domaine, relève du point de vue de la tâche ou d'une caractérisation intrinsèque du domaine. Pour cela, Common KADS s'appuient sur la distinction introduite dans CYC par (Lenat & Guha, 1990) entre l'*adéquation épistémologique* et l'*adéquation pragmatique* d'une ontologie. Une ontologie est épistémologiquement adéquate si elle effectue les distinctions requises par le problème à résoudre ; elle est pragmatiquement adéquate si elle permet un comportement de résolution de problème efficace. Pour concilier ces deux principes parfois en conflit, une bonne pratique consiste à faire reposer les ontologies épistémologiquement adéquates sur les notions les plus primitives du domaine, les entités « primitives » de la citation ci-dessus, et les ontologies pragmatiquement adéquates sur les entités « abstraites » dont le niveau d'abstraction correspond aux besoins d'une résolution pertinente et efficace du problème posé.

Les entités abstraites constituent par conséquent le point de vue de la tâche sur le domaine. Les entités primitives constituent ce qui existe dans le domaine indépendamment de la tâche. Il est essentiel de bien noter que les points de vue du domaine sont toujours des abstractions de l'ontologie primitive du domaine.

Par exemple, dans le cas des équipements audio, l'ontologie primitive consiste dans une description au niveau des composants électroniques. L'ontologie abstraite pragmatiquement adéquate pour les problèmes relatifs au diagnostic de panne et à la réparation mobilise des notions de moins bas niveau comme « amplificateur », « contrôle du volume », etc. Ces termes sont fondés sur l'ontologie primitive dans la mesure où ils sont construits, assemblés, à partir d'objets existant dans le domaine décrits par l'ontologie primitive. Cela suppose donc deux choses :

- il est possible de définir un niveau primitif; en d'autres termes, on est capable de déterminer quels sont les composants ultimes du domaine ;
- toute entité du domaine est une primitive ou une abstraction de primitives ; il existe donc un réductionnisme de principe stipulant que l'on peut toujours se rapporter à une description en termes des objets existants ultimes du domaine, mais que pour des raisons de convenance pratique on s'en tient à une description abstraite ;

Il s'agit donc d'une conception « atomiste » de la connaissance du domaine puisque l'on suppose qu'il existe des notions primitives, renvoyant à des objets primitifs, à partir desquelles (desquels) toutes (tous) les autres se déduisent par composition.

L'exemple cité (c'est celui donné dans l'article) est à ce titre tout à fait révélateur : il s'agit d'un domaine où par définition ces hypothèses sont vérifiées puisqu'il s'agit d'un domaine d'ingénierie où l'on s'intéresse aux objets *que l'on a construits* : on connaît par définition leurs composants ultimes. Qu'en est-il des domaines où l'on a affaire à des objets que l'on n'a pas contruits ? Nous verrons en section 2-4 que la médecine invalide ces hypothèses dans la mesure où l'on est incapable de définir des composants ultimes et primitifs de l'action médicale : la médecine n'est pas la biologie.

Néanmoins, ces distinctions sont utiles. Elles permettent de dégager précisément ce qui relève de la tâche ou du domaine. Dans la suite, nous appelerons *ontologie primitive* une ontologie fondée sur des entités primitives renvoyant aux composants ultimes du domaine et *ontologie pragmatique* ou *abstraite* une ontologie formulée en des termes adéquats au raisonnement. D'une certaine manière, comme nous allons le voir, une ontologie primitive contient les *primitives du domaine* et une ontologie pragmatique contient les *primitives du raisonnement*.

Recherchant à dégager des méthodes génériques de raisonnement, Common KADS retient donc le point de vue des ontologies pragmatiques. Common KADS distingue donc les notions suivantes :

l'ontologie du domaine qui est l'ontologie pragmatique ;

le modèle du domaine qui contient les propositions exprimant les connaissances du domaine formulées à l'aide des primitives de l'ontologie pragmatique ; le domaine est exprimé à travers le raisonnement que l'on veut y mener ;

l'ontologie du modèle spécifie les termes génériques qu'un modèle de raisonnement peut utiliser pour se référer à des propositions du modèle du domaine ; c'est une méta-description des éléments du modèle du domaine dans la mesure où elle décrit le point de vue du problème ou de la tâche sur le domaine ;

le schéma du modèle décrit la structure des entités représentées dans le modèle du domaine ; la nature des types d'entités requises pour un modèle donné du domaine dépend de l'application, c'est-à-dire du type de problème traité ; le schéma du modèle contient les primitives représentationnelles de l'ontologie du modèle, c'est la composante terminologique fixant son vocabulaire.

En résumé, on a donc le modèle du domaine qui, à partir d'une ontologie du domaine, formule les connaissances du domaine. Ces connaissances et l'ontologie sur laquelle elles s'appuient sont à un niveau pragmatique ou abstrait dans la mesure où elles s'expriment de manière adéquate au raisonnement qui sera mené. L'ontologie du modèle décrit la structure de ces connaissances, c'est-à-dire la structure que l'on a donnée à ces connaissance pour qu'elles soient adaptées au raisonnement. Par exemple, si le raisonnement est un raisonnement tout/ partie, il faut que les connaissances soient formulées en une hiérarchie *partie-de*. L'ontologie du modèle est donc la description générique d'une relation *partie-de*. Enfin, le schéma du modèle contient les termes mis en œuvre dans l'ontologie du modèle.

D'une certaine manière, le point de vue de la tâche absorbe la description du domaine dans la mesure où l'ontologie primitive n'est jamais explicitée mais que seule l'ontologie abstraite ou pragmatique l'est pour formuler les connaissances d'une manière adéquate au raisonnement à mener.

En conclusion, on constate que tant KADS-I que Common KADS reposent sur l'hypothèse que la tâche gouverne la description du domaine sans être déterminée par le domaine lui-même. Le domaine, d'une certaine manière, est neutre dans la description des connaissances dans la mesure où il fournit un vocabulaire (pragmatique) pour exprimer dans le domaine le raisonnement que l'on veut y mener. Mais la nature du raisonnement n'est pas déterminée par le domaine. On retrouve une conception identique dans la définition de la couche inférence.

2.2.2 La couche inférence

La couche inférence contient toutes les inférences primitives que le SBC doit être capable d'effectuer pour mener à bien sa tâche de résolution. Ces inférences sont des vues abstraites des raisonnements accomplis dans le domaine dans la mesure où l'on ne s'intéresse *qu'aux rôles tenus par les connaissances dans le raisonnement indépendamment de leur nature*. Elles sont primitives dans la mesure où elles constituent les briques élémentaires à partir desquelles tout raisonnement devra être construit en les combinant. Ce sera notamment la fonction de la couche tâche que décrire comment des inférences primitives doivent s'agencer pour constituer un raisonnement effectif. Une inférence primitive se définit à l'aide de trois éléments :

la source de connaissance : cette entité est la procédure permettant d'effectuer l'inférence ; prenant des connaissances en entrée, elle en produit en sortie ; pour cela, la source de connaissances utilise des connaissances du domaine qui sont spécifiées par la « vue du domaine » qui lui est associée ;

les méta-classes : une source de connaissance saisit en entrée des données pour en produire d'autres en sorties ; ces données sont appelées « méta-classes » ; une méta-classe est le nom d'un rôle que peut remplir un objet du domaine dans le raisonnement ; pour cela, elle pointe sur les types d'objets qui peuvent tenir ce rôle ;

la vue du domaine : elle spécifie quelles parties de la théorie du domaine peuvent être utilisées par la source de connaissances pour effectuer son inférence.

Pour fixer les choses, prenons un exemple tiré du domaine médical. Une inférence que l'on peut qualifier de primitive peut être celle du diagnostic : à partir de symptômes, le diagnostic comme activité déduit, infère, un diagnostic comme objet, c'est-à-dire une pathologie. Les méta-classes de cette inférence sont donc définies d'une part par le rôle de « symptômes » joués par certains objets du domaine, et d'autre part par le rôle de « pathologies » joués par certains objets du domaine également. La source de connaissance est l'activité de diagnostic.

La vue du domaine attachée au diagnostic correspond aux connaissances du domaine reliant les symptômes aux pathologies. Ce sont essentiellement des relations de deux types : d'une part les symptômes *évoquent* les pathologies, d'autre part les pathologies *causent* les symptômes. Ces relations ne sont pas équivalentes dans la mesure où l'on ne connaît pas toutes les conséquences causales d'une pathologie (ces conséquences ne sont pas nécessairement des symptômes et réciproquement), il existe des symptômes qui ne sont pas causés par les pathologies, non pas qu'il n'y ait pas de rapport causal entre les deux, mais tout simplement parce que ce dernier n'est pas connu !

Enfin, les méta-classes pointent sur les types d'objets du domaine qui peuvent tenir le rôle des méta-classes dans le raisonnement : ce sont donc les objets du domaine qui peuvent tenir le rôle de symptôme ou de pathologie. Ce sera par conséquent, dans notre exemple, les états observables du patient dans le cas des symptômes, et les maladies pour les pathologies.

Il faut remarquer les méta-classes ne se rapportent pas directement aux objets du domaine, mais par l'intermédiaire de leur type, c'est-à -dire le nom qu'il leur a été donné dans le schéma du domaine. Par exemple, la méta-classe symptôme ne pointe pas directement sur les états fiévreux du patient, mais sur tous les états qui ont été catégorisés dans la classe « état clinique » du schéma du domaine.

On peut constater sur un exemple aussi simple combien il peut être agréable et utile de distinguer la connaissance de son rôle inférentiel : supposons que l'on pose comme autre inférence primitive, l'inférence consistant à prédire des observables. Une telle inférence est utile car elle permettra par exemple de prescrire des examens permettant de confirmer ou d'infirmer la valeur prédite, et donc la pathologie incriminée. Une telle inférence s'effectue donc entre la méta-classe « pathologie » et la méta-classe « observables ». On remarque alors que la vue du domaine utilise les mêmes connaissances que l'inférence de diagnostic et que la méta-classe « observables » pointera sur les mêmes types d'objet du domaines que les symptômes, à savoir les états cliniques et paracliniques.

Par conséquent, la séparation de la couche domaine et de la couche inférence permet de déclarer *une seule fois* une connaissance du domaine qui sera utilisée, mobilisées, *plusieurs fois* par des connaissances de contrôle différentes. La connaissance du domaine est en quelque sorte considérée *en elle même, pour elle-même*, indépendamment des *points de vue* que l'on peut avoir sur elle dans le cadre d'un raisonnement, points de vue que traduisent les méta-classes de la couche inférence.

L'ensemble des inférences primitives permet de construire une « structure d'inférence » qui rassemble toutes les inférences en un même graphe bipartite, comprenant d'une part des carrés labélisés par les méta-classes et d'autre part des ronds labélisés par les sources de connaissances. Un tel graphe décrit les inférences possibles, mais ne prescrit pas comment les assembler effectivement pour constituer le raisonnement du système. Spécifiant la *compétence* du système à construire, la structure de tâche doit être complétée par une structure de contrôle permettant d'obtenir la *performance* du système. La couche tâche a pour fonction de spécifier une telle structure de contrôle.

2.2.3 La couche tâche

Cette couche contient les connaissances de contrôle permettant de combiner les inférences primitives pour atteindre un certain but. Elles sont organisées en *tâches*, une tâche permettant d'atteindre un *but*, un but pouvant être atteint par plusieurs tâches, mais une tâche n'est définie que pour un seul but.

Une tâche consiste essentiellement en une décomposition en sous-tâches et une spécification des relations de dépendance entre ces sous-tâches. Il existe plusieurs types de sous-tâches :

- des tâches primitives, qui correspondent aux inférences primitives de la couche inférence ;
- des tâches composées définies dans la couche tâche, donc par une décomposition en sous-tâches également ;
- des tâches de transfert qui requiert une interaction avec l'utilisateur, notamment pour fournir ou demander des données.

Les tâches, du fait de la décomposition en sous-tâches, peuvent être organisées en un arbre de tâche dont toutes les feuilles doivent être des inférences primitives.

Pour une tâche donnée, les dépendances entre les sous-tâches sont exprimées en pseudo-code, au moyen de structures de contrôle comme `tant que`...`répéter` ou `si`...`alors`. Elles mobilisent des variables de contrôle, appelées « termes du contrôle », qui correspondent à des ensembles de méta-classes. Par exemple, un terme de contrôle peut être l'ensemble des pathologies données par les inférences de diagnostic.

Par exemple, la tâche de diagnostic systématique peut être définie de la manière suivante :

tâche *diagnostic systématique*
but :
trouver le plus petit composant dont le comportement révèle un dysfonctionnement, s'il existe ;
entrée :
existence d'un mauvais fonctionnement ;
sortie :
`kaputt-sous-système` : partie du système dont le comportement est anormal ;
termes de contrôle :
`différentiel` : ensemble des hypothèses courantes actives ;
structure de la tâche :
diagnostic-systématique(`problème` ⟶ `kaputt-sous-système`) =
sélection (`problème` ⟶ `système-modèle`)
engendrer-hypothèse (`système-modèle` ⟶ `différentiel`)
`Répéter`
tester-hypothèse (`différentiel` ⟶ `kaputt-sous-système`)
`engendrer-hypothèses(kaputt-sous-système` ⟶ `différentiel`)
`jusqu'à ce que différentiel` = 0

2.3 Les modèles d'interprétation

La notion de « modèles d'interprétation » est née de la volonté de dégager des éléments *réutilisables* d'une modélisation à une autre : il apparaît en effet que dans une application donnée, une large part du modèle d'expertise n'est pas spécifique à l'application mais peut être partagée par d'autres applications. Il paraît donc judicieux de recenser ces éléments génériques, d'en constituer une librairie, et de les fournir comme briques de base à l'ingénieur de la connaissance lors de la réalisation d'une nouvelle application.

Ce souci de généricité correspond à la volonté de donner à l'IA un fonctionnement *cumulatif*, à l'instar des autres disciplines à caractère scientifique, pour faire passer l'IA de l'artisanat à la technologie. La généricité est, pour les auteurs de KADS, la « *conditio sine qua non* pour améliorer l'état de l'art en ingénierie des connaissances ».

Les modèles d'interprétation correspondent à ces éléments génériques réutilisables et l'une des contributions majeures de la méthodologie, de son propre point de vue, est de proposer une librairie de modèles d'interprétation.

Une modèle d'interprétation se définit comme étant un modèle d'expertise auquel est enlevé tout ce qui relève du domaine, c'est-à-dire la couche domaine. Ainsi, ce qui est générique, est le raisonnement, les connaissances de contrôle permettant de mettre en œuvre les connaissances du domaine. Il existe des manières universelles de raisonner, indépendamment de ce sur quoi l'on raisonne, et il est possible d'en établir la liste.

La notion de modèle d'interprétation est donc la conséquence logique de la distinction fondatrice dans KADS entre la matière du raisonnement (couche domaine) et la forme du raisonnement (les trois couches de contrôle). Il s'agit par conséquent d'un fonctionnalisme de la connaissance, dans la mesure où le domaine est *neutre* pour le raisonnement, qu'il peut varier indéfiniment sans modifier la structure du raisonnement.

Ce fonctionnalisme matière/rôle se traduit en pratique par des rapports type/instance entre les rôles fonctionnels dans le raisonnement et les objets matériels du domaine. C'est ainsi que les méta-classes s'instancient dans le domaine en leur associant des types du schéma du domaine sur lesquels elles pointent.

Ceci se confirme par l'utilisation qui est faite des modèles d'interprétation. Il faut pour cela insister sur l'*interprétation* que proposent des modèles du domaine d'application. Le principe est que la bibliothèque de modèles proposée par KADS permet d'interpréter les données du domaine pour construire le modèle d'expertise du SBC. Cela signifie que ce sont les modèles d'interprétation qui permettent de poser les relations et les concepts entre les objets du domaine, ce dernier se contentant de fournir le vocabulaire permettant d'instancier le raisonnement dans le domaine. Ainsi, le raisonnement diagnostic est-il un raisonnement formel donnant lieu à un modèle d'interprétation qui, dans un domaine médical par exemple, instancie ses méta-classes « signes » et « diagnostic » dans des classes d'objets du domaine, comme les « signes cliniques » ou « para-cliniques ».

2.4 L'hypothèse de l'interaction limitée

La méthodologie KADS repose donc sur un *formalisme* des connaissances qui ne peut être justifié que si le domaine est bien cette matière indifférente, neutre, qui véhicule la forme du raisonnement sans contribuer à en déterminer la nature. Comme nous l'avons dit plus haut, cette hypothèse s'énonce dans KADS comme « l'hypothèse de l'interaction limitée entre les couches ». C'est bien sûr sur l'articulation domaine/contrôle que cette hypothèse joue un rôle crucial, bien que valant en droit pour chaque couche. C'est donc à elle que revient le devoir de justifier l'autonomie des modèles d'interprétation et leur constitution en bibliothèques. Il s'agit donc de revenir sur cette neutralité du domaine et examiner quelle peut être la légitimité d'une telle hypothèse.

En présentant l'articulation entre la couche domaine et la couche inférence, celle qui confronte explicitement le domaine à sa mise en œuvre dans un raisonnement, les concepteurs de KADS remarquent qu'il faut distinguer *deux manières de nommer le même objet*, l'une indépendante de son usage, l'autre reflétant son usage. Ils constatent par ailleurs que « toute l'activité de l'acquisition des connaissances est en fait pour une large part une question de donner des noms justes » (Schreiber *et al.*, 1993, p. 28, note). Il semble donc bien que la couche du domaine est indépendante des couches de contrôle, le domaine pouvant être défini indépendamment de son usage dans les raisonnements.

Mais cela semble contredit par des affirmations, apparaissant de manière récurrente dans les exposés de la méthodologie, assertant que « dans beaucoup d'applications, il existe des interactions entre la conceptualisation du domaine et la spécification du processus de résolution », qu'il n'est pas dans leur propos « d'affirmer que la théorie du domaine peut être en général définie complètement indépendamment de son utilisation dans le processus de résolution ».

Si bien que l'on a l'impression que la méthodologie est construite sur l'hypothèse de la neutralité du domaine même si l'on sait que cela n'est pas vrai ! Il semble paradoxal d'avoir d'une part une attitude théorique proche d'un constructivisme conceptuel, où le domaine résulte d'une modélisation effectuée le plus souvent en fonction d'une tâche à accomplir, celle du système en particulier, et d'autre part une attitude méthodologique proche d'un formalisme rigoureux affirmant que la forme du raisonnement est indépendante et indifférente à sa matière.

Mais on sait que, dans KADS, tout est affaire de dénomination, de savoir nommer les choses en fonction du domaine ou en fonction du raisonnement : cela suppose qu'il existe un ensemble de choses, d'individus du domaine qui ne sont jamais remis en cause et que, tel Adam dans le jardin de l'Éden, le problème n'est pas de les créer, Dieu est là pour ça, mais seulement de leur trouver un nom. En d'autres termes, il existe un univers, une ontologie d'individus invariable quelles que soient les applications ; le problème que chaque conception de système doit alors résoudre est de savoir *trouver le nom qui convient pour chacun de ces individus*.

Ce qui doit être construit par la modélisation, c'est la manière de nommer les individus et de les catégoriser, non leur existence. Si bien que l'on peut affirmer à la fois qu'il n'est pas certain de savoir nommer les choses indépendamment de leur implication dans un raisonnement, et que néanmoins la couche du domaine est autonome et neutre vis-à-vis de la tâche puisque l'existence et l'individualité des choses n'est jamais remise en cause.

C'est pourquoi il nous semble nécessaire de tempérer les affirmations et professions de foi « constructivistes » des auteurs de KADS : quand ils parlent de « conceptualisation », il ne s'agit pas d'une construction conceptuelle du domaine, mais une catégorisation des objets du domaine en classes appelées « concepts ». S'il est difficile de catégoriser les objets sans avoir un point de vue relevant de la tâche du système, un tel point de vue ne remet jamais en cause la conception *extensionnaliste* selon laquelle un concept est une classe d'individus, puisés dans une liste inhérente au domaine. Il faut donc affirmer que la méthodologie KADS est bien un formalisme.

2.5 Le formalisme de KADS

Il ressort que notre analyse de la méthodologie KADS qu'elle mobilise une hypothèse reposant sur la neutralité du domaine vis-à-vis de la tâche dans son objectif méthodologique de constituer des méthodes de raisonnement réutilisables. D'une certaine manière, une telle hypothèse est inévitable étant donné l'objectif. Le fait même de parler de réutilisabilité implique une certaine neutralité vis-à-vis du domaine : la méthode se définit indépendamment du domaine.

On pourrait arguer que la structure du domaine est nécessairement prise en compte puisque le modèle du domaine contient les propositions ou connaissances du domaine et que le schéma du modèle du domaine, dégageant la structure générique du modèle du domaine, contient les axiomes dont les connaissances sont les conséquences instanciées. Mais, il faut bien voir que ces axiomes correspondent aux « assomptions that a model makes about the world » (Wielinga *et al.*, 1993, p.319). En d'autres termes, il s'agit de la structure formelle imposée au domaine pour les besoins du raisonnement, et non d'une structure du domaine reflétée par le raisonnement.

On voit que le domaine n'a en fait plus aucune autonomie et qu'il est entièrement absorbé dans la tâche. Le domaine est réduit à n'être que la liste des objets sur lesquels raisonner, et en particulier, comme le soulignent fort pertinemment les auteurs de KADS-I, le domaine est la liste des noms qu'il faut donner aux objets indépendamment de leur usage, sachant que l'usage viendra leur donner de nouveau noms pour les intégrer à des structures relationnelles soutenant le raisonnement. Guarino constate ainsi :

> « the modelling activity must establish a correspondence between a knowledge base and two separate subsystems: the agent's behaviour (*i.e.* the problem solving *expertise*) and its own environment (the problem *domain*). However, current knowledge modelling technologies tend to focus on the former subsystem, viewing domain knowledge as strongly dependent on the particular task at hand; » (Guarino, 1995, p.626)

Nous avons noté plus haut une relative ambiguïté à propos de l'hypothèse concernant l'interaction limitée entre les couches, notamment entre la couche du domaine et celles du contrôle. Les auteurs de KADS professent en effet à la fois un relatif scepticisme quant à l'autonomie du domaine quant à la tâche et une indépendance de la tâche quant au domaine. Au vu de notre discussion, on peut interpréter cette ambiguïté de la manière suivante. Le vocabulaire du domaine est en effet indépendant de la tâche et du raisonnement : de ce point de vue, le domaine est indépendant de la tâche. Mais étant donné que toute connaissance du domaine est un instance d'un axiome postulé par la méthode de raisonnement, il ressort que le domaine n'est compris que depuis le raisonnement que l'on y mène.

Mais cette ambiguïté peut également s'interpréter de manière plus « ontologique ». En effet, nous avons noté la distinction entre ontologie primitive et ontologie pragmatique, la première caractérisant le domaine indépendamment de la tâche, la seconde depuis le point de vue de la tâche. L'ambiguïté pourrait alors se dissoudre si on distingue ces deux types d'ontologies. Si KADS ne parle pas de l'ontologie primitive autrement que pour introduire l'ontologie pragmatique, c'est qu'en fait, étant donné son objectif de dégager des méthodes de résolution générique, seule importe le point de vue de la tâche et de l'ontologie pragmatique.

Mais on peut tenter de renverser la perspective et considérer le domaine en lui-même et étudier les ontologies primitives. En d'autres termes, il faut se tourner non pas vers les primitives du raisonnement mais vers les primitives du domaine. C'est donc à ce type de travaux que nous allons nous intéresser à présent.

3 Partager les domaines : les ontologies

Nous avons déjà rencontré de nombreuses fois la notion d'ontologies. D'un point de vue technique, il s'agit des primitives des langages formels de représentation des connaissances. Une telle définition, techniquement précise, est fort vague du point de vue de la modélisation. Nous avons vu, avec la distinction des ontologies primitives et pragmatiques, combien le rapport avec le domaine pouvait être complexe et différencié. L'enjeu d'élucider ce rapport au domaine réside dans l'objectif de constituer des ontologies réutilisables.

En effet, alors que les méthodologies d'acquisition s'intéressent principalement aux connaissances de résolution d'une tâche, la problématique ontologique s'est autonomisée voilà quelque peu et reçoit une attention soutenue dans la recherche d'un paradigme de recherche guidant l'élaboration d'ontologie et fixant les conditions de leur réutilisabilité d'une tâche à une autre. De la même manière que les tâches sont définies indépendamment des domaines, on entend bien pouvoir définir les domaines indépendamment des tâches.

Il s'agit à présent de considérer le fondement théorique et les réalisations pratiques de la problématique des ontologies du domaine en IA. Après avoir exposé la définition théorique des ontologies telles qu'elle semble se dégager des travaux actuels, nous considérons une proposition d'élaboration effective d'ontologies réutilisables dans le domaine médical.

3.1 Une définition des ontologies

S'il est bien clair pour tout le monde qu'une ontologie se rapporte à un domaine, il l'est beaucoup moins dans quelle mesure et sous quel rapport. La définition qui semble se dégager de la floraison des travaux fort disparates sur ce sujet est celle proposée par Gruber (Gruber, 1993), reprise dans quasiment tous les travaux sur les ontologies (voir par exemple (Van Heijst *et al.*, 1996)) : une ontologie est la « spécification d'une conceptualisation ». Il s'agit à présent de comprendre ce que cela peut signifier.

Une conceptualisation d'un domaine consiste dans la construction d'une théorie conceptuelle où les objets du domaine se voient cataloguer dans des classes conceptuelles et munir des propriétés qualifiant ces classes. Par exemple, dans le domaine médical, une théorie conceptuelle détermine quelles sont les maladies, les traitements, et processus physio-pathologiques, etc., et quelles sont les différentes relations qui les unissent. Chaque objet du domaine est donc catalogué en maladie, traitement ou processus et associé à d'autres objets via les relations.

> « Un corpus de connaissances formellement représentées se fonde sur une conceptualisation : les objets, les concepts et autres entités qui sont présumés exister pour quelque centre d'intérêt ainsi que les relations qui tiennent entre eux. Une conceptualisation est une vue abstraite et simplifiée du monde que l'on veut représenter pour quelque but. Chaque base de

> connaissance, système à base de connaissances, ou agent du niveau des connaissances doit en passer par une conceptualisation, explicitement ou implicitement. » (Gruber, 1993, p.199)

La spécification d'une telle conceptualisation consiste dans la donnée des classes conceptuelles et de leurs propriétés formelles. Il ne s'agit pas de savoir si le diabète est une maladie ou non, ce que fait la conceptualisation du domaine, mais de savoir si la notion de « maladie » fait partie des classes conceptuelles permettant la conceptualisation. De même, la spécification précise quelles sont les relations entre les classes et quelle est leur arité (binaire, ternaire, n-aires, etc.). La spécification n'établit pas quelle relation particulière tient entre le diabète et le glucophage, mais quelle relation existe entre les classes conceptuelles « maladie » et « médicament », et avec quelle arité.

Par conséquent, si une ontologie est la spécification d'une conceptualisation, une ontologie dresse l'inventaire des classes logiques et de leur relation permettant de construire une conceptualisation. C'est donc la même notion que celle de « schéma du domaine » que nous avions rencontrée dans KADS.

Mais on pourrait croire à ce stade qu'une ontologie se borne à être le vocabulaire logique, la signature du système formel permettant d'exprimer la conceptualisation. Dans ce cas, une ontologie ne serait que l'appareillage linguistique de la conceptualisation et n'exprimerait rien de particulier quant à la nature ou au contenu du domaine. Mais, il n'en est rien dans la mesure où il est nécessaire de décrire en langue naturelle ce qui est dénoté par les termes formels du langage et de le spécifier par des axiomes contraignant l'interprétation et l'usage formels de ces termes :

> « Pour les systèmes à base de connaissances, ce qui « existe » est exactement ce qui peut être représenté. Quand la connaissance d'un domaine est représentée dans un formalisme déclaratif, l'ensemble des objets qui peuvent être représentés est appelé l'univers du discours. Cet ensemble d'objets, et les relations formalisées entre eux, se réfléchissent dans le *vocabulaire* représentationnel avec lequel le programme à base de connaissances représente la connaissance. Ainsi, on peut décrire l'ontologie d'un programme en définissant un ensemble de termes représentationnels. Dans une telle ontologie, les définitions associent les noms des entités de l'univers du discours (*e.g.* classes, relations, fonctions ou d'autres objets) à du texte lisible par des humains décrivant ce que ces noms sont censés dénoter, et à des axiomes formels qui contraignent l'interprétation et l'usage bien formé de ces termes. » (Gruber, 1993, p.199)

Dans cette perspective, une ontologie doit donc tenir compte, dans la spécification de la conceptualisation, de la signification attendue des termes conceptuels utilisés, et lui faire correspondre des prescriptions sémantiques. Il faut être capable de spécifier le sens des termes du domaine et traduire en logique formelle le sens spécifié. A ce moment, deux problèmes se posent. Il s'agit d'une part du fait de l'interprétabilité linguistique des primitives ou descripteurs, rendant délicate l'explicitation du sens des primitives, et d'autre part du fait qu'une syntaxe logique est en elle-même *ontologiquement neutre*, rendant problématique la traduction formelle de l'explicitation du sens des termes.

Nous avons déjà souligné dans notre introduction que la signature des formalismes de représentation des connaissances adoptait comme expressions symboliques les unités linguistiques permettant d'exprimer les connaissances dans la langue naturelle. Ce fait est lourd de conséquence dans la mesure où il faut se confronter à la charge sémantique que prennent *ipso facto* ces descripteurs. En effet, le choix d'un descripteur repose en général sur l'intention du concepteur de signifier telle ou telle notion du domaine : pour rendre ce choix explicite, il n'a à sa disposition que la possibilité d'emprunter à la langue du domaine l'unité linguistique lui paraissant la plus propre à exprimer le lien entre le descripteur et le domaine. Mais reste néanmoins implicite ce qui doit être exactement compris par cette unité, car il n'est pas spécifié comment interpréter l'unité pour qu'elle renvoie plutôt à tel ou tel aspect de la notion visée. C'est un problème fort délicat que tout modélisateur rencontre dans la conception des SBC. Il est donc nécessaire de pouvoir disposer des moyens d'exprimer ce qui est exactement visé par l'emploi de tel ou tel descripteur dans la signature d'un formalisme. En effet, ces moyens

> « allow the knowledge engineer to make clear the *intended meanding* of a particular predicate symbol. This is especially important since we are constantly using natural language words for predicate symbols, relying on them to make our statements readable and to convey meanings not explicitly stated. However, since words are ambiguous in natural language, it may be important to tag these words with a semantic category, in association with a suitable axiomatisation, in order to garantee a consistent interpretation. » (Guarino, 1995, p.626)

Comme on le voit à travers cette citation de Guarino, ces moyens consistent à spécifier formellement l'interprétation visée. Mais, on rencontre alors le deuxième fait que nous avons énoncé comme problème, à savoir que la logique du premier ordre est ontologiquement neutre.

En effet, on constate que les formalismes de représentation des connaissances en IA repose sur des structure syntaxiques dont la motivation repose sur la volonté de signifier telles ou telles structures dans le domaine. Cependant, le formalisme ne peut par lui-même contraindre cette signification si bien que l'usage formel de ces structures syntaxiques peut ignorer voire contredire ce qui a présidé à leur élaboration. Par exemple, les concepts et rôles dans les logiques de description renvoient, par leur label même (« concept » et « rôle ») à certaines interprétations privilégiées, alors que le formalisme, les caractérisant comme des prédicats unaires et binaires respectivement, ne dit rien sur ces interprétations. Il est donc nécessaire de disposer de moyens spécifiant le formalisme de manière à ce qu'il reflète la signification visée, pour qu'il ne soit pas possible d'exprimer des formules syntaxiquement correctes et sémantiquement vides de sens dans le domaine :

> « What is needed is a way to constrain and to make explicit the *intended models* of a KR [Knowledge Representation] language, in order to facilitate large-scale knowledge integration and to limit the possibility to state something that is reasonable for the system but not reasonable in the real world. » (Guarino, 1995, p.632)

C'est pour éviter ces problèmes qu'il est donc nécessaire de renforcer la spécification de la conceptualisation d'« engagements ontologiques » contrôlant l'utilisation des classes conceptuelles et leur *interprétation sémantique*. La spécification ne consiste donc pas seulement en un vocabulaire, mais aussi dans les contraintes sémantiques liées à l'utilisation de ce vocabulaire. En d'autres termes, outre la spécification syntaxique stipulant qu'une relation est binaire ou ternaire, qu'une classe conceptuelle est un ensemble d'atomes ou un atome, il est nécessaire d'introduire des spécifications sémantiques stipulant comment interpréter une classe conceptuelle dans le domaine modélisé. Ces spécifications sont formulées au « niveau ontologique », niveau de spécification des langages de représentation des connaissances.

3.2 Engagement ontologique et niveau ontologique

La notion de niveau ontologique renvoie à une célèbre distinction en IA établie par Brachman (Brachman, 1979) entre les différentes manières de caractériser les primitives d'un langage de représentation. A un premier niveau, le niveau *logique*, les primitives sont des prédicats et des fonctions, qui possèdent une sémantique formelle en termes de relations entre objets du domaine. A un second niveau, le niveau *épistémologique*, les primitives sont des primitives structurant la connaissance (« *knowledge structuring primitives* ») : ces primitives sont génériques et expriment certains types de relation entre des notions spécifiques d'un domaine. De telles notions spécifiques correspondent à un niveau *conceptuel* où les primitives sont des concepts spécifiques du domaine. Enfin, un niveau *linguistique* clôt cette hiérarchie de niveaux en caractérisant les primitives comme des unités linguistiques du domaine.

Le niveau ontologique (Guarino *et al.*, 1994; Guarino, 1994) s'intercale entre les niveaux conceptuel et épistémologique pour expliciter comment une primitive épistémologique assume la signification d'un primitive conceptuelle du domaine. En effet, le niveau épistémologique est introduit pour souligner le fait que les structures logiques ne doivent pas être neutres quant à leur interprétation possible, mais doivent renvoyer à des structures effectives du domaine. Le niveau ontologique permet d'expliciter cette non neutralité du formalisme en spécifiant l'interprétation visée.

Cette explicitation s'effectue en restreignant le nombre des interprétations possibles, cela ayant pour effet de caractériser la signification des catégories ontologiques de base retenues pour décrire le domaine. Le niveau ontologique est donc le niveau de la *signification* alors que le niveau épistémologique est celui de la *structure* : le niveau ontologique spécifie la signification de la structure épistémologique qui s'efforce de rendre compte dans son formalisme des concepts spécifiques du domaine (niveau conceptuel).

Il s'agit à présent de savoir comment expliciter la signification visée des primitives. Cette explicitation repose sur la notion d'« engagement ontologique ». Cette notion trouve son origine dans les travaux de Quine (Quine, 1977) et stipule que la signification d'une notion consiste dans le fait qu'elle implique l'existence dans le domaine d'un ou plusieurs individus correspondant à cette notion.

Le principe est que, par exemple, une classe conceptuelle comme « pommes » ne pourra pas s'interpréter de la même manière que la classe conceptuelle « rouge ». En effet, ce qui est une pomme dans le domaine est fixé une fois pour toutes : l'interprétation de la classe conceptuelle « pomme » est rigide et pointe toujours sur les mêmes objets. Autrement dit, la classe des pommes est toujours la même quel que soit le point de vue adopté. Si, au contraire, ce qui est une pomme n'en est plus une si je change de point de vue, cela signifie que mes points de vue sont ontologiquement incompatibles. En revanche, la classe conceptuelle « rouge » peut recevoir différentes interprétations sans que l'intégrité du domaine soit remise en cause : un objet peut ou non être rouge. Autrement dit, selon un point de vue donné, une pomme peut être rouge, selon un autre, être verte, sans que cela remette en cause l'ordre ontologique du monde. Le fait d'être rouge n'est pas ontologiquement équivalent au fait d'être une pomme. Si bien qu'être rouge n'engage pas ontologiquement aux mêmes conséquences que le fait d'être une pomme. Quand je mange une pomme, il existe un objet de moins sur terre ; quand je peints la pomme rouge en vert, l'ensemble ces choses existantes n'est pas modifié.

Le principe est donc qu'une conceptualisation du domaine doit être spécifiée par des prescriptions sémantiques qui prennent l'allure de règle d'interprétation de sémantique formelle : ce sont les règles sémantiques qu'il faut respecter pour qu'un système formel utilisant le vocabulaire logique stipulé par la spécification syntaxique puisse passer pour une ontologie de tel ou tel domaine. Prenons un exemple de spécification pour fixer les idées en suivant (Guarino, 1992).

Le principe est qu'il faut restreindre les interprétations possibles de la représentation formelle du domaine pour que ne soit accepté que ce qui est possible dans le domaine. On pose que chaque interprétation possible est un monde possible au sens de la sémantique des mondes possibles de Kripke. On veut que ne soient retenus que les mondes ontologiquement compatibles avec le monde réel dont la représentation formelle est la modélisation. Ces mondes sont *ontologiquement compatibles* avec le monde réel, et par force entre eux, dans la mesure où ils respectent les mêmes engagements ontologiques, les mêmes prescriptions sémantiques traduisant ces engagements ontologigues. La notion de compatibilité ontologique définit une relation d'équivalence entre les mondes possibles et dote ainsi la théorie formelle du domaine d'une sémantique de type $S5$ (Chellas, 1980).

Par exemple, les types rigides comme la classe des pommes doivent posséder la même extension dans chaque monde possible acceptable par la théorie. Cela signifie en particulier qu'ils doivent vérifier une propriété du genre suivant :

$$\forall x(x \in \alpha \supset nec(x \in \alpha))$$

Cette formule signifie que si, dans un monde possible, un objet x appartient au type α, alors il en est de même dans tous les autres mondes possibles ontologiquement compatibles avec lui. Cette prescription sémantique traduit l'intuition ontologique concernant l'identité des substances : « si x est un α dans un situation particulière, il doit rester un α dans toute situation possible pour conserver son identité. Par exemple, un animal peut cesser d'être un chiot tout en restant un chien : *animal* et *chien* sont sémantiquement rigides, et *chiot* non. » (Guarino, 1992, p.252).

Dans un telle théorie, il est facile de voir que si le prédicat *pomme* est rigide, *rouge* ne l'étant pas, les formules $Pomme(a) \wedge Rouge(a)$ et $Pomme(a) \wedge \neg Rouge(a)$ sont ontologiquement compatibles alors que les formules $Pomme(a) \wedge Rouge(a)$ et $\neg Pomme(a) \wedge Rouge(a)$ ne le sont pas. En effet, si a est une pomme dans un monde possible de la théorie, il en est une dans tout les mondes possibles, alors qu'il peut être rouge dans l'un mais pas l'autre. Par conséquent, $Pomme(a) \wedge Rouge(a)$ et $\neg Pomme(a) \wedge Rouge(a)$ affirme simultanément que a est une pomme dans tous les mondes possibles puisque $Pomme(a) \rightarrow nec(Pomme(a))$ et qu'il existe un monde dans lequel $\neg Pomme(a)$, c'est-à-dire un monde où a n'est pas une pomme. La contradiction résultant correspond donc au fait que ces deux formules ne sont pas ontologiquement compatibles. Ces deux formules spécifient des univers différents malgré une identité de vocobulaire ou de schéma du domaine.

De telles distinctions permettent de rendre compte des intuitions selon lesquelles il est « naturel » d'écrire $\forall x : Pomme\ Rouge(x)$ alors que $\forall x : Rouge\ Pomme(x)$ sonne faux : l'embarras provenant de la quantification ; autant il semble naturel de balayer l'ensemble des objets qui sont des pommes, *attendu que les pommes sont des objets du monde réel*, autant il semble contre intuitif de considérer tout ce qui est rouge pour établir si c'est une pomme ou non. C'est que certains prédicats renvoient directement à des objets du monde, alors que d'autres en sont seulement des propriétés plus ou moins fluentes d'ailleurs. Alors que la rougeur ne suffit pas à déterminer un objet dans le monde, le fait d'être une pomme le permet. La notion sémantique de rigidité rend compte de cette dissymétrie.

En spécifiant les engagements ontologiques inhérents à l'établissement du schéma d'un domaine, le travail ontologique consiste à rendre compte formellement, logiquement, de la signification attendue des concepts. Comme l'affirme Guarino, « au niveau ontologique, ce que nous voulons est rendre compte par une sémantique formelle restreinte les engagements ontologiques inhérents à l'usage d'un prédicat donné comme d'une espèce (*sort*) » (Guarino, 1994).

3.3 L'ontologie sous-jacente du niveau ontologique

Le travail ontologique de Guarino permet d'approfondir notablement la représentation des connaissances telle qu'elle s'élabore selon la démarche formaliste en montrant comment surmonter la neutralité ontologique de la logique : en engageant le formalisme vis-à-vis du domaine en établissant des axiomes, tels ceux spécifiant les types rigides, dont le respect restreint les possibilités d'interprétations à celles possibles ou pertinentes dans le domaine, le niveau ontologique lie le raisonnement à la réalité ontologique du domaine.

La logique du premier ordre est en effet syntaxique et fondamentalement *neutre* vis-à-vis de l'interprétation que l'on peut lui donner. Il ne faut pas y voir une faiblesse de la logique, mais bien plutôt son essence : en tant que formelle, elle s'interroge sur la possibilité d'inférer sur la seule foi de la forme, *vi formae*. Si la logique est neutre par rapport aux domaines d'interprétation, c'est qu'elle a été conçue pour ça. Mais, comme le constate fort justement Guarino, la formalisation logique d'un domaine ne peut rester neutre vis-à-vis de ce dernier : s'il est logiquement équivalent de quantifier sur les pommes ou les objets rouges dans les formules du types $Pomme(x) \wedge Rouge(x)$, il ne l'est pas dans le domaine des fruits et légumes. En particulier, parmi deux inférences logiquement possibles, $Pomme(a) \wedge Rouge(a) \supset \exists x Pomme(a)$ et $Pomme(a) \wedge Rouge(a) \supset \exists x Rouge(a)$, une seule sera reconnue correcte dans le domaine, $Pomme(a) \wedge Rouge(a) \supset \exists x Pomme(a)$, l'autre paraissant indirecte : il existe quelque chose de rouge parce qu'il existe une pomme rouge.

Ce travail formel sur les engagements ontologiques des prédicats s'inscrit dans la conception logique du monde dans la mesure où il exprime par une *sémantique formelle* l'engagement ontologique. En mobilisant explicitement une conception extensionnelle du sens comme celle des mondes possibles, ce travail suppose que le monde est une extension d'individus et que les connaissances qui les visent sont des structures ensemblistes construites à partir de cette extension. On peut même dire que ces hypothèses sont encore renforcées dans la mesure où, par exemple, la rigidité sémantique assure que dans des mondes ontologiquement compatibles, la liste des individus, au sens des objets dénotés par les types spécifiques (« *sort* ») ne peut être que la même : les mondes possibles sont de mulptiples points de vue possibles sur un monde réel fondamentalement identique à lui-même. Ce qui peut changer par conséquent d'un monde à un autre, ce sont des prédicats construits sur cette extension d'objets dont la signification ne remet pas en cause l'existence de ces objets (*e.g.* « chiot » possède une signification dont ne dépend pas l'existence des objets : un chien qui n'est plus un chiot est toujours le même chien).

Par ailleurs, l'adoption d'un cadre intensionnel comme en fait foi le choix d'une sémantique de mondes possibles ne doit pas illusionner : l'intensionalité logique est de nature extensionnelle, étant donné qu'elle ne fait que rabattre l'interprétation extensionnelle d'un prédicat valable globalement dans l'univers d'interprétation à la localité d'un monde possible de cet univers. Mais on entend bien conserver les avantages inhérents à l'extensionalité, à savoir la compositionalité des classes et la possibilité de définir la négation [2]. En d'autres termes, l'ontologie primitive, au sens vu plus haut, c'est-à-dire le domaine tel qu'il est indépendamment des tâches, postulée par ce travail, est que le monde est constitué d'individus à partir desquels construire les interprétations ensemblistes des termes du modèle du domaine. Le mobilier ontologique ultime du monde, ce à partir de quoi l'on peut construire les ontologies pragmatiques dont les systèmes de résolution de problèmes ont besoin, est une base extensionnelle d'individus.

Cela signifie que toute structure construite à partir de cette base ne caractérise pas un domaine mais le raisonnement l'on veut y faire et l'interprétation qu'il reçoit dans ce domaine. Cela paraît confirmé si l'on rapporte le travail de Guarino au projet auquel il se réfère de concevoir une ontologie formelle. En effet, la détermination formelle des engagements ontologiques inhérents à l'usage de certains termes dans la modélisation logique des domaines renvoie au problème de la définition d'une *ontologie formelle*. Par ontologie formelle, il faut comprendre le développement systématique, formel et axiomatique d'une logique de toutes les formes et modes d'être (Cocchiarella, 1991). Le projet d'une ontologie formelle repose sur la possibilité de définir formellement des engagements ontologiques partagés par tous les domaines : par exemple, l'axiome défini plus haut de la rigidité est la définition formelle de ce que tout domaine doit comprendre par *espèce* (*sort*). L'ontologie formelle se donne donc pour tâche d'axiomatiser la signification supposée des notions permettant de spécifier la conceptualisation d'un domaine, c'est-à-dire ce que l'on entend pas concepts, relations, propriétés, rôles, etc. On obtient alors une terminologie structurée des primitives de représentation munie de la formalisation de leur signification ontologique.

Ce projet repose sur l'hypothèse que les modes d'être et les formes d'être sont universels et transcendent les domaines. Ce postulat se justifie dans la mesure où les modes d'être et les formes d'être ne se déterminent pas tant par le domaine que par la manière de le penser, si bien que ces distinctions ne relèvent pas du domaine, mais de la pensée du domaine et de ses lois : l'ontologie formelle dépend des lois de la pensée, c'est-à-dire de la logique. Nous retrouverons cette notion d'ontologie formelle dans le chapitre suivant dans la mesure où elle reprend fidèlement le projet husserlien d'une détermination des lois de la pensée et de ses implications nécessaires sur ce qui est pensé, à savoir le domaine.

Dans la mesure où les lois de l'ontologie formelle et les objets qu'elle pose découlent des lois de la pensée, il devient facile de comprendre pourquoi toute ontologie du domaine en IA reflètent nécessairement le point de vue de la tâche, c'est-à-dire le point de vue du raisonnement que l'on veut mener à propos des objets du domaine. Les primitives de conceptualisation reçoivent une sémantique via les lois formelles qu'elles vérifient pour

2. Il est en effet impossible, comme l'avait bien vu Aristote de définir la négation en termes de propriété : qu'est-ce qu'un non-homme en termes de propriétés partagées par tout ce qui n'est pas un homme, et niées par tous les hommes ? Rien dans la mesure où il n'existe pas de telles propriétés : qu'y-a-t-il de commun au nombre deux, à une table ou une couleur ? Rien en dehors du fait qu'ils sont des quelques choses. Mais un homme est aussi un quelque chose : il n'est donc pas possible d'opposer ces objets aux hommes par une propriété. Le point de vue extensionnel permet de définir la négation comme le complémentaire d'une classe et non comme l'union de propriétés, si bien que, pour peu qu'il y ait un univers du discours défini comme un liste d'individus, la négation est toujours définie. Renoncer à l'extensionnalité est donc fort coûteux puisque cela interdit en définitive de disposer d'un *calcul* logique.

rendre les raisonnements possibles. Par conséquent, le point de vue ontologique propre au domaine semble se réduire à n'établir que la liste des individus existants. L'ontologie du domaine consiste dans le catalogue d'individus dans laquelle l'ontologie formelle viendra puiser ses extensions. En d'autres termes, l'approche relevant de l'ontologie formelle spécifie l'ontologie primitive du domaine comme étant une liste des individus existants du domaine, celle-ci fournissant la base pour l'instanciation des raisonnements sur le domaine. Autrement dit, l'ontologie formelle détermine comment raisonner en termes d'espèce ou de propriété par exemple, stipulant ce que cela implique en termes d'individus, en considérant comme acquis que l'on sait ce qui est un individu. L'ontologie formelle se donne l'ontologie du domaine sous forme d'individus et détermine les contraintes formelles que cela implique pour le raisonnement. Il s'agit de traduire une caractérisation du domaine en contraintes formelles sur le raisonnement. Mais il faut se donner la caractérisation du domaine, on la considère comme acquise comme base d'individus existants ou ontologie primitive.

D'une certaine manière, cela correspond à ce qui est visé dans ces travaux, c'est-à-dire non pas la caractérisation du raisonnement, mais bien la caractérisation du domaine indépendamment de la tâche. En d'autres termes, on veut dégager les axiomes propres au domaine indépendamment du raisonnement dont ils seront l'objet. Comme le souligne Guarino, le projet n'est pas de montrer en quoi la description du domaine est influée par l'objectif de la tâche, comme nous l'avons vu avec les ontologies pragmatiques introduites dans Common KADS, mais bien, en adoptant une perspective complémentaire, de montrer que la description de la tâche doit se conformer à la structure intrinsèque du domaine : « les distinctions ontologiques peuvent avoir un impact sur les possibilités de raisonnement offertes par un formalisme de représentation des connaissances ». Mais, là où l'objectif est manqué, c'est que l'on ne sait pas comment trouver la structure du domaine : on la considère comme acquise ; seule sa transcription en contraintes formelles pour le raisonnement est étudiée.

Nous avons établi nos conclusions en nous appuyant sur le fait que les travaux cités ici mobilisent comme relation au domaine la relation de dénotation telle qu'elle est mise en œuvre en sémantique formelle comme dans la sémantique des mondes possibles. Cela suggère qu'il existe en quelque sorte une tension entre l'outil logique utilisé et le projet : alors que le projet est de caractériser le domaine en lui-même, on en arrive à élaborer les axiomes d'une ontologie formelle universelle. La raison est simple : en utilisant l'outil logique, on veut formuler des prescriptions sémantiques sur le domaine de manière *a priori* ; on veut exprimer non pas ce qui est propre au domaine, mais ce que tout domaine, en tant que domaine, doit vérifier. Le projet est, en quelque sorte, de définir des règles du type suivant : « si x est une espèce, alors x doit vérifier tel ou tel axiome ; s'il est une propriété, alors il doit vérifier ces axiomes », etc. Il s'agit en effet du projet même de l'ontologie formelle, comme nous le reverrons avec Husserl au chapitre suivant : décrire de manière *a priori* ce que tout domaine, en tant que domaine d'objets, vérifie. La logique permet d'exprimer ces contraintes *a priori* dans la mesure où, selon Guarino, elle seule permet d'établir de manière *non ambiguë* des contraintes :

> « the use of modal logic, as a tool to constrain the intended semantics of the underlying non-modal theory, seems unavoidable if we wish to express ontological constraints unambiguously. In the perspective of formal ontology previously mentioned in the introduction, these constraints should also be related to *a priori* distinctions amoung the entities of the world [...]. » (Guarino, 1994)

Mais, en exprimant les contraintes vérifiées par tout domaine, on exprime pas ce qui est vérifié par ce domaine-ci et non par ce domaine-là. En d'autres termes, on reporte ces distinctions à l'expression empirique *a posteriori* des connaissances du domaine dans le cadre ontologique axiomatisé de l'ontologie formelle. En voulant surmonter la neutralité de la logique du premier ordre, Guarino rencontre la neutralité due à l'analyticité *a priori* de la logique : l'axiomatique de l'ontologie formelle prend effectivement en charge un certain rapport au domaine. Mais ce rapport n'est pas spécifique à un domaine particulier, mais valable pour tout domaine.

Or, on peut comprendre le projet de manière plus approfondie en voulant caractériser un domaine indépendamment de la tâche, c'est-à-dire en caractérisant ce qui est vrai de manière *a priori* dans le domaine et seulement dans ce domaine. Ce doit en effet être *a priori* dans la mesure où la caractérisation effectuée doit être valable pour toute application dans le domaine. Par conséquent, au lieu de s'appuyer sur une élaboration logique qui par définition ne peut que livrer des principes analytiques, il faut dégager une caractérisation *synthétique* a priori du domaine et donc en mobilisant pour cela des outils non formels.

L'enjeu de vouloir caractériser le domaine indépendamment de la tâche, de cerner ce que toute application dans le domaine doit vérifier renvoie au projet d'élaborer une ontologie synthétique *a priori*, non pas formelle mais *matérielle*, pour reprendre la terminologie husserlienne.

Le projet de l'ontologie formelle permet d'expliciter certains aspects que toute application doit respecter : elle explicite les axiomes que doit vérifier toute connaissance portant sur « substance », une « propriété », etc. Elle s'intéresse à ce qu'est une « propriété » en général, en spécifiant ce que cela implique que de penser quelque chose à travers la catégorie de « propriété ». Ainsi, considérer que « rouge » est une propriété et non une substance implique que l'on doit respecter certains principes dans notre manière de penser le rouge. Le projet d'une ontologie matérielle permet d'expliciter ce que signifie le « rouge » dans tel ou tel domaine. Non pas en tant que propriété, c'est-à-dire en tant que catégorie formelle, mais en tant que contenu vérifiant des lois *a priori* dans le domaine.

Mais d'un point de vue pratique, quel est l'enjeu d'avoir une ontologie « matérielle » ? Le projet de l'ontologie formelle, du fait de la formalité de l'outil logique utilisé, renvoie à une conception du monde comme extension d'individus. Or, si une telle conception n'est pas vérifiée par les domaines, cela implique qu'il existe une dimension intermédiaire entre l'ontologie formelle et les connaissances empiriques qu'elle permet d'exprimer. Cette dimension, c'est celle du synthétique *a priori*. L'enjeu pratique est donc de dégager la dimension synthétique *a priori* des domaines de telle manière à pouvoir décrire le domaine comme une extension d'individus. Autrement dit, ce sont les principes synthétiques *a priori* du domaine qui permettent de décrire le domaine comme une base extensionnelle par laquelle les contraintes d'interprétation de l'ontologie formelle pourront s'appliquer.

On doit mieux à présent comprendre le sens de notre travail : il ne s'agit pas de remettre en cause l'intérêt d'une approche formelle, mais de montrer qu'elle est incomplète par nature. Compléter le formalisme, c'est le dépasser en réintroduisant la perspective du synthétique *a priori*.

3.4 Des librairies d'ontologies médicales

Les travaux que nous venons de voir sont théoriques. Il est par conséquent utile de vérifier si nos conclusions tiennent toujours dans le cadre des travaux plus pragmatiques et empiriques menés dans le cadre des ontologies, qui consistent à développer concrètement des ontologies et à les organiser de manière à pouvoir les réutiliser. Dans cette perspective, ces travaux prolongent le travail qui a été commencé avec la constitution de bibliothèques de méthodes de résolution, *e.g.* les modèles d'interprétation de KADS. Comme nous l'avons souligné plus haut, la généricité des méthodes de résolution de problème doit être corrélée à la généricité de la représentation du domaine, et en particulier de son ontologie. Nous nous intéresserons dans cette section à la tentative de (Van Heijst *et al.*, 1995) de constituer une libraire d'ontologies médicales.

Si comme nous l'avons vu avec KADS, la constitution de méthodes réutilisables de résolution de problème a déjà fait l'objet de nombreux travaux, les ontologies réutilisables de domaines particuliers de connaissances a reçu moins d'attention. Ce peut être en partie imputable à deux problèmes qui se posent dès que l'on aborde la problématique du domaine : le problème de l'immensité (*hugeness problem*) et le problème de l'interaction (*interaction problem*) (Van Heijst *et al.*, 1995).

Nous connaissons déjà le second problème, mais dans sa version « limitée » qui permet de fonder l'autonomie des modèles d'expertise de KADS vis-à-vis des domaines d'application. Pris dans sa complète généralité, le problème de l'interaction consiste dans le fait que « les connaissances du domaines ne peuvent être représentées indépendamment d'hypothèses portant sur la manière dont elles vont être utilisées dans le raisonnement ». Le problème de l'immensité consiste simplement dans le fait que les connaissances propres à un domaine sont *vraiment* nombreuses. Pour qu'il soit néanmoins possible d'obtenir des ontologies réutilisables, il est nécessaire de se donner un cadre méthodologique permettant de pallier au moins localement ces problèmes.

Mais, avant d'examiner les hypothèses de travail permettant de réaliser la constitution d'une librairie d'ontologies réutilisables, il est utile de préciser ce que l'expression d'« ontologie » est censée recouvrir. Les auteurs des travaux auxquels nous faisons ici référence distinguent 4 types d'ontologies :

les ontologies du domaine qui expriment des conceptualisations spécifiques à un ensemble d'application ;

les ontologies génériques qui contiennent des concepts tenus pour génériques dans de nombreux domaines : par exemple, les *états, processus, action*, etc. ; les concepts des ontologies du domaine sont des spécialisations de concepts génériques ;

les ontologies de représentation qui contiennent les conceptualisations sous-jacentes aux formalismes de représentation des connaissances ; ce sont donc des ontologies de descripteurs linguistiques ;

les ontologies d'application qui correspondent l'ontologie nécessaire à une application particulière.

Les ontologies d'application, propres à une application, ne sont pas réutilisables. En revanche, les ontologies de domaine et les ontologies génériques sont censées l'être. Une ontologie de domaine contient la structure des connaissances du domaine, leurs types et leurs contraintes d'utilisation. Une ontologie n'est pas une base de connaissance dans la mesure où les connaissances décrivent des états du monde alors que l'ontologie du domaine ne contient que les descripteurs permettant de déclarer ces états du monde.

Le principe de l'approche est donc de constituer une librairie d'ontologies domaniales et génériques à partir desquelles construire une ontologie de l'application. Ce qui permet de surmonter les problèmes d'immensité et d'interaction consiste en plusieurs principes :

- une taxinomie de spécialisation des domaines médicaux qui permet de hiérarchiser les ontologies domaniales ;
- une hiérarchie de méthode de résolution ;
- une réseau d'inclusion de théories génériques ;

Les théories génériques contiennent des concepts qui sont réutilisables dans tous les domaines médicaux et pour toutes les méthodes de résolution. Ce sont par exemple les concepts de *patient*, *maladie*, *traitement*, etc., qui reflètent la terminologie de base partagée par tous les praticiens de la médecine. Ces catégories sont dites « naturelles » dans le mesure où elles reflètent un consensus s'imposant naturellement à tout un chacun. Ces concepts s'organisent selon des liens plus ou moins forts de systématicité : *patient* est solidaire de *soignant*, *maladie* de *traitement*. Il y aura donc autant de théories de types que de systématicité recensés. Ces théories peuvent entretenir cependant des liens d'inclusions : la terminologie des maladies et des traitements peut faire appel à des notions plus générique comme *état* et *processus* définies dans une théorie d'abstraction plus élevée.

Il existe donc une première hiérarchie d'ontologies génériques, organisées selon des liens d'inclusion, stipulant par exemple que si l'on veut utiliser des connaissances relatives aux maladies, il faut disposer de la terminologie concernant les états et les processus.

Pour les autres concepts ne relevant pas des ontologies génériques, ils sont organisés selon leur spécificité en termes de domaines ou en termes de méthodes de résolution. Une taxinomie de spécialisation des domaines médicaux, reflétant l'activité médicale effective, permet de définir la spécificité des concepts : par exemple, l'infarctus du myocarde est plus spécifique que la notion d'infarctus dans la mesure où le premier relève des accidents cardiaques alors que le second relève de la catégorie plus générale des accidents vasculaires. Le principe est que « quand un concept possède un domaine particulier comme sa valeur de spécificité du domaine, il est spécifique pour ce domaine, mais est réutilisable sur tous ses sous-domaines ».

Par ailleurs, les méthodes de résolution impliquent des « engagements ontologiques » dans la mesure où les méthodes mobilisent des rôles qui doivent correspondre à des objets du domaine, comme nous l'avons vu avec KADS. En d'autres termes, les objets et concepts du domaine se définissent par les méthodes qui les utilisent. Il paraît dès lors difficile d'en considérer la réutilisabilité d'une méthode à une autre. En effet, « pris littéralement, le problème de l'interaction exclut la réutilisation des concepts à travers les méthodes ». Mais, en adoptant une solution semblable à celle proposée pour la réutilisabilité d'un domaine à un autre, « le problème de l'interaction ne se pose pas avec la même acuité pour chaque concept, et il va être montré que le degré de spécificité des concepts en termes de méthodes peut être utilisé comme un index pour organiser la bibliothèque d'ontologies ».

La solution est donc d'établir une hiérarchie de méthodes de résolution telle que les méthodes possédant un même père partagent les rôles définis par la méthode correspondant à ce nœud.

Chaque concept reçoit ainsi deux attributs qui déterminent sa spécificité en termes de domaine et de méthode de résolution. Pour chaque combinaison d'un nœud de la taxinomie des méthodes de résolution et un nœud de la taxinomie des domaines doit correspondre une théorie dans la bibliothèque. L'idée est que cette théorie, spécifique à un domaine et une méthode, est réutilisable pour toutes les méthodes et tous les domaines qui sont plus spécifiques dans les taxinomies.

Ces structures taxinomiques permettent de résoudre les problèmes de l'interaction et de l'énormité : la hiérarchie de spécificité des méthodes résout le problème de l'interaction, la hiérarchie de spécificité des domaines résout le problème de l'énormité.

Ces structures permettent la constitution d'une bibliothèque d'ontologies dans la mesure où, pour une application donnée existante, il suffit de recenser les concepts utilisés et leur définition en les considérant sous les trois aspects suivants : leur généricité (*e.g. maladie, état, etc.*), leur appartenance à un domaine spécifique (*e.g. infection nosocomiales*), leur existence comme rôle dans une méthode de résolution (*e.g. diagnostic, hypothèse, etc.*). Cela fait, il suffit de reporter les définitions dans les théories appropriées selon les index de spécificité établis. De la même manière, pour utiliser la bibliothèque, il suffit de recenser les concepts utiles déterminés par l'analyse du problème et au moyen des scores établis selon les hiérarchies de spécificité, de réutiliser les théories existantes de la bibliothèque.

En conclusion, il apparaît dans ce travail que le domaine se traduit par un vocabulaire structuré selon des relations d'hyperonymie et d'hyponymie sans que le contenu des notions ne soit explicité autrement que par leur rôle dans les raisonnements mis en œuvre dans les méthodes. Une théorie de la bibliothèque consiste en effet dans l'application d'une méthode sur le vocabulaire d'un sous domaine, réitérant ainsi la démarche de KADS. Si bien que nos commentaires à propos de KADS s'appliquent également ici et pour les mêmes raisons. La notion de bibliothèques d'ontologie structure la couche domaine en une taxinomie de vocabulaire et permet de la croiser avec une hiérarchie de modèles d'interprétation. En d'autres termes, la librairie d'ontologies médicales est une librairie d'ontologies pragmatiques. Par conséquent, ce travail ne remet pas en cause les conclusions que nous avons tiré de KADS à propos des ontologies pragmatiques et de travaux sur l'ontologie formelle à propos des ontologiques primitives. Il est donc temps à présent que nous avons dégagé leurs hypothèses de travail, d'examiner la validité de ces hypothèses.

4 Logique et ontologie

Les travaux que nous avons rapidement parcourus dans ce chapitre ont pour principale motivation de rendre possible le partage et la réutilisabilité des composants de connaissances dans les SBC. Ce n'est en effet qu'à cette condition qu'une véritable ingénierie de la connaissance est possible, c'est-à-dire une pratique qui capitalise ses acquis et ses savoirs en composants génériques et outils méthodologiques.

La constitution de composants génériques s'effectue, on l'a vu, selon deux dimensions : les ontologies du raisonnement, qui capitalisent les manières de raisonner selon les problèmes traités, les ontologies du domaine, qui capitalisent les différentes structures objectives selon les domaines traités. Pour que de tels travaux soient menés à bien, on a vu avec KADS qu'une hypothèse sur l'indépendance du domaine et de la tâche est nécessaire : l'hypothèse de l'interaction limitée.

En effet, les travaux de KADS reposent sur l'hypothèse qu'il est possible de caractériser une méthode de résolution indépendamment du domaine. Réciproquement, les travaux sur les ontologies du domaine reposent sur l'hypothèse qu'il est possible de déterminer la structure objective d'un domaine indépendamment de la tâche traitée. Ces hypothèses sont nécessaires dans la mesure où ce sont elles qui assurent la généricité et la réutilisabilité des composants construits.

D'un point de vue plus fondamental, ces hypothèses renvoient à des options philosophiques que l'on peut formuler ainsi : l'indépendance de la tâche vis-à-vis du domaine suppose qu'il existe des lois du raisonnement qui sont indépendantes de ce sur quoi l'on raisonne ; l'indépendance du domaine vis-à-vis de la tâche repose sur l'hypothèse qu'un domaine consiste en une extension d'individus dont l'ontologie décrit la structure ensembliste. La première hypothèse suppose qu'il existe des lois *formelles* du raisonnement, et que tout raisonnement concret dans un domaine se rapporte à un moule formel. La seconde hypothèse suppose que les individus du domaine constituent l'univers de référence dans lequel interpréter les lois formelles du raisonnement.

La question qui se pose à présent est de savoir dans quelle mesure ces hypothèses sont tenables et donc si les méthodologies qu'elles fondent sont utilisables d'emblée dans un domaine de la pratique. Il semble que la réponse doive être négative : dans les domaines pratiques, ce qu'on appelle un domaine correspond non pas à la description d'objets particuliers indépendants des actions que l'on entreprend sur eux, mais à un ensemble de pratiques dont le critère de démarcation est plus ou moins flou, relevant de normes sociales, culturelles, professionnelles, techniques, etc. En d'autres termes, séparer raisonnement et domaine est une idéalisation qui ne renvoie pas à une distinction réelle. La question est alors de savoir à quelles conditions de telles hypothèses sont applicables et les résultats qu'elles fondent utilisables.

Nous examinons ces deux hypothèses dans un domaine particulier, le domaine médical. Nous montrons quels problèmes se posent dès que l'on veut savoir si elles sont vérifiées.

4.1 Le monde comme ensemble d'individus

Pour que le monde puisse être compris comme une nomenclature, il est nécessaire que seule la dénomination des objets soit soumise à la décision du concepteur, mais que l'existence de ces objets soit hors de doute. Or on s'aperçoit que c'est précisément l'existence sur laquelle il faut décider lors de la modélisation de connaissances médicales. En effet, selon le point de vue adopté, ce ne sont pas seulement les noms des objets et les relations qui les définissent qu'il faut déterminer, mais si l'objet en question existe ou non. En d'autres termes, le problème n'est pas simplement de trouver le bon niveau d'abstraction pour exprimer une ontologie pragmatique adéquate aux objectifs que l'on s'est fixé pour la conception d'un SBC ; le problème est qu'il n'existe pas d'ontologie primitive recensant les existants primitifs dont les ontologies pragmatiques ne seraient que l'abstraction.

Par exemple, les objets s'individuent en médecine selon des critères qui relèvent du contexte pragmatique. C'est ainsi qu'un objet n'existe que s'il peut faire l'objet d'un examen ou d'une intervention : les innovations technologiques font évoluer l'ontologie médicale dans la mesure où les méthodes d'investigations s'affinent. Mais, si l'on considère par exemple les comptes rendus de scanner du cerveau destinés à la neuro-chirurgie, on s'aperçoit que les objets sont mentionnés seulement s'il est possible d'intervenir chirurgicalement : ils n'existent en tant qu'objets que s'ils font l'objet d'une action chirurgicale. Et l'on constate que les nouvelles méthodes de micro-chirurgie enrichissent le catalogue des objets qu'un scanner met en évidence. Que dire dans un tel cas? Que ces objets existaient bien mais n'étaient pas mentionnés parce qu'inutiles, ou bien qu'ils n'existaient pas auparavant, du moins pas en tant qu'objets, mais seulement à titre de traces graphiques sur un cliché, qu'ils deviennent objets seulement quand le contexte pragmatique s'y prête?

De même, ce qu'est un objet n'est pas stable selon le point de vue. Il l'est pour des raisons de granularité : un histologue ne considère pas le cœur comme un objet mais comme un composé, de la même manière qu'un physicien des particules ne considère pas les objets macroscopiques comme de véritables objets. Chaque spécialiste considère un niveau d'objectivité particulier sans préjuger de sa compatibilité avec les autres types d'objectivité définis dans les autres disciplines. Il n'existe pas de niveau privilégié détenant la vérité objective ultime duquel tous les autres devraient découler. L'histologue n'a pas davantage raison que le clinicien : ils traitent tous deux d'objets différents et communiquent leur savoir respectif pour définir un enrichissement réciproque de leurs connaissances.

Par ailleurs, une « sténose » peut être définie comme un objet d'un point de vue clinique, alors que d'un point de vue radiologique il ne s'agit que d'une déformation morphologique. Dans un cas, il s'agit d'un individu, dans l'autre d'une propriété morphologique d'un individu. Les travaux logiques de l'ontologie formelle nous apprennent que cela ne veut pas dire la même chose, et ils ont bien raison. Mais ils ne nous disent pas comment déterminer dans quel cas on se trouve en pratique et comment s'y retrouver avec un même terme qui recouvre des statuts objectifs différents.

Pour être capable de disposer du monde comme d'une liste d'individus, il faudrait pouvoir disposer d'un point de vue universel intégrant tous les autres, des critères d'objectivation possédant un primat absolu de sorte que tous les autres dussent s'y rapporter comme à leur condition de légitimité tant théorique que pratique. Or ces objets, nous ne les avons pas construits, ni créés : nous ne les connaissons que depuis le point de vue pragmatique qui nous les fait rencontrer comme le support ou l'objet de notre action. Relatifs aux actions, ils ne les transcendent pas par une individualité qui serait établie depuis une instance qui les considérerait en eux-mêmes, indépendamment des actions pratiques qui les mobilisent, comme tente de le faire la science.

Pratique thérapeutique et donc commandée par la finalité de son action, la médecine aborde un objet dont la structure est mal définie car nul ne l'a pensé, nul ne l'a construit : le corps humain. Puisqu'il est impossible de se rapporter à l'ontologie de la construction scientifique, mathématique ou physique, la consistance de l'objet ne se définit qu'à partir de la pratique : l'objet est ici le corrélat de l'activité pratique et non de la pensée.

Par ailleurs, la science permet rarement de reconstruire un objet de la pratique : elle en construit d'autres, ceux que l'on rencontre dans la pratique scientifique. C'est la raison pour laquelle la théorisation scientifique d'un domaine correspond à un bouleversement complet de son objectivité, ce qui était objet ne l'est plus et réciproquement. Il n'existe pas de point de vue unique permettant de considérer dans une même perspective universelle les objets du sens commun, les particules élémentaires, les systèmes biologiques, etc. La raison essentielle étant que les conditions d'objectivité des uns, *i.e.* ce qui fait qu'un objet est un objet, sont incompatibles avec celles des autres. Si l'on considère sérieusement ce qui constitue l'objectivité en physique, il ne peut y avoir de place pour les objets de la vie quotidienne comme une banale table.

Considérons un dernier exemple, les médicaments. Un individu pourrait être un médicament tel qu'il est commercialisé dans une pharmacie, ou tel qu'il est prescrit. Un médicament se définit par sa classe thérapeutique et son principe actif. Si par exemple, le principe actif est modifié, mais que l'indication thérapeutique reste identique, c'est, pour le prescripteur, toujours le même médicament, le même individu dont il est question. Mais il ne peut en être de même pour le pharmacologue : si l'on change de principe actif, on modifie le médicament. Cela implique que pour le prescripteur, c'est toujours le même objet, alors que pour le pharmacologue il s'agit de deux objets différents.

En conclusion, l'individu comme tel, l'objet tel qu'il est en soi-même, est inaccessible. Il n'existe aucun domaine concret où il est possible de se donner le monde comme un ensemble d'individus, les pratiques amenant à effectuer des découpages incompatibles entre eux. L'objectivité ne procédant pas d'un vérité absolue, mais que de points de vue particuliers, il ressort qu'il n'est possible de déterminer ce qu'est un objet que par son intérêt pour une pratique donnée. En d'autres termes, il est impossible de définir extensionnellement les objets, mais seulement à travers les propriétés par lesquelles ils existent pour une pratique donnée. Un objet n'existe pas comme membre d'une liste d'individus, mais comme le corrélat, défini par les propriétés qu'il doit vérifier, d'une pratique donnée. Il faut donc être capable de déterminer ce qu'est un objet indépendamment des méthodes fournies par la logique puisque les hypothèses fondatrices de cette dernière ne sont pas vérifiées. Un objet défini en termes de propriétés non formelles se définit non référentiellement, les propriétés ne sont pas des traits caractérisant l'objet dénoté, mais différentiellement, les propriétés le distinguent de ses voisins. Est un objet non pas ce qui existe dans un catalogue d'individu mais ce qui peut être distingué de ce qui l'entoure, ce qui peut faire l'objet distinct de l'action ou de la pensée.

Le basculement qu'il faut donc effectuer consiste à considérer les ontologies comme des structures différentielles conçues dans le cadre d'une tâche définie au sein d'un domaine d'activité.

4.2 Le raisonnement comme logique formelle

La seconde hypothèse formaliste est celle mise en œuvre dans KADS, qui stipule l'indépendance de la tâche vis-à-vis du domaine. Cette hypothèse semble motivée par un constat empirique indubitable : un domaine se constitue d'un ensemble de pratiques de raisonnement. Elle fait appel néanmoins à une généralisation contestable : ces pratiques de raisonnement sont universelles, et indépendantes des domaines.

Cela signifie qu'il faut considérer le domaine comme n'étant qu'une instanciation, avec le vocabulaire du domaine, des termes de contrôle du raisonnement. Cela signifie donc que le vocabulaire du domaine, c'est-à-dire le fait d'adopter la sémantique du domaine et non plus seulement celle de la spécification de la tâche, n'introduit aucune variation par rapport à ce qui est signifié par la tâche elle-même. Autrement dit, une méthode générique

Diagnostic systématique

```
find(diagnosis)
select(system-model)
WHILE (no conclusion) DO
  decompose(system-model)
  WHILE(nb-hypos-in-differential > 1)
  select(variable-value)
  specify(norm)
  compare(variable-value, norm)
```

Classification heuristique

```
obtain_data
transform(variables)
match(abstracted_data)
specify(solution_abstraction)
```

FIG. 2.1 –*Structures de tâche du diagnostic systématique et de la classification heuristique formulées dans le cadre de la méthodologie* KADS-I

de résolution formalisant une tâche diagnostic pourrait décrire tant le diagnostic médical que le diagnostic automobile puisque l'usage des termes de ces domaines ne modifient en rien la nature de la tâche diagnostic.

Une telle affirmation n'est pas évidente. En effet, considérons ce que peut être une méthode générique de résolution. Soit on dispose d'une méthode spécifiée de manière très générale comme la classification heuristique (Clancey, 1985). Pour la mettre en œuvre dans un domaine, il ne s'agit pas simplement de l'instancier, mais il faut aussi la développer et en terminer la spécification dans le domaine. La classification heuristique, méthode conçue pour le diagnostic, doit donc s'interpréter et se développer dans le domaine si bien que rien n'assure par principe que l'on obtiendra une méthode de résolution identique selon que le domaine est le diagnostic médical ou le diagnostic automobile.

Soit on dispose d'une méthode spécifiée de manière précise, et donc en des termes formels assez proches de l'opérationnalisation informatique. Dans ce cas, il s'agit d'un même programme que l'on veut appliquer dans deux domaines différents. Le travail porte alors que la caractérisation du domaine dans les termes du programme et on retrouve alors le problème vu précédemment de savoir ce que signifie un terme du domaine, ce qu'est un objet du domaine, pour qu'il soit enrôlable dans le programme. Or, (Duursma, 1992) montre qu'il existe bien de telles différences entre les méthodes génériques de résolution. Notamment, il montre comment la méthode du diagnostic systématique ou structurel (*systematic diagnosis*) (vue plus haut) est différente dans la précision de sa formulation de la classification heuristique (*heuristic classification*) (Clancey, 1985) (voir également la figure 2.1) :

> « It is not surprising that heuristic classification seems so different from the structural diagnosis procedure and that no elaborated control structure nor hidden assumptions can be found. Heuristic classification is in fact no procedure, but a "knowledge level" interpretation of search behaviour. Therefore the application of the interpretation model for heuristic classification is different from the one for systematic diagnosis. Systematic diagnosis provides a procedure for analysis and a method for solving the problem; while heuristic classification as such provides a framework for analyzing the problem, but is not specific enough as a problem solving method. »

Ainsi, la notion de méthode générique ou de modèle d'interprétation peut renvoyer à de différents niveaux de spécificité. Il s'agit donc davantage de « modèle d'interprétation », c'est-à-dire de méthodes pour interpréter les données du domaine et les enrôler dans une méthode de résolution, que de méthodes génériques, réutilisables parce que génériques. Ainsi, les modèles d'interprétation ne sont pas indépendants du domaine dans la mesure où ils ne s'instancient pas dans le domaine, mais dans la mesure où ils s'interprétent dans le domaine. En d'autres termes, les modèles d'interprétation ne peuvent contribuer en rien à l'affirmation que les lois du raisonnement peuvent être spécifiées indépendamment des domaines où elles sont mises en œuvre.

Si parfois on peut rencontrer une telle affirmation – que les méthodes sont indépendantes des domaines, c'est que l'on mobilise l'hypothèse selon laquelle il doit exister quelque chose comme une *logique formelle* théorisant les lois de la pensée indépendamment des contenus pensés. La logique formelle tire son appellation, comme on le sait, du fait qu'elle ne s'intéresse qu'à la forme du raisonnement en faisant abstraction de sa matière. Cette indépendance à la matière implique qu'elle vaut pour toute matière, et donc que les lois du raisonnement sont universelles. Une version formaliste de KADS serait qu'il faut recenser ces lois formelles pour en constituer une bibliothèque à partir de laquelle construire le modèle d'un domaine par instanciation. En fait, nous venons de voir qu'il ne s'agit pas d'instanciation, mais d'interprétation.

Nous verrons dans la partie suivante que la logique formelle qui est mise en œuvre en IA ne reflète pas les lois du raisonnement, et que de ce fait, il est faux que toute pensée, pour peu qu'elle soit rationnelle doive s'y plier. La logique utilisée en IA n'est pas une discipline « formelle » parce qu'indifférente à son contenu : elle possède son propre contenu. Ce contenu est le calcul : les lois de la logique sont les lois mathématiques du calcul effectif et non celles de la pensée possible. Par conséquent, la logique formelle théorise un domaine particulier, le calcul, de la même manière que toute autre discipline scientifique théorise son domaine d'investigation. Il ne peut donc exister de lois transcendant tous les domaines possibles. Il n'existe que des théories régionales de domaine, des théories *matérielles* car dépendantes de leur contenu, et les théories dites *formelles* sont des théories régionales dont le domaine est le formel, c'est-à-dire le calcul.

Mais, avant d'en arriver là, reprenons le travail méthodologique de cette partie pour conclure ce chapitre sur les problèmes rencontrés par les travaux actuels et sur le sens de nos propositions.

5 Artéfacture et ontologie

Les exemples que nous venons de voir dans le contexte médical montrent que les travaux méthodologiques vus plus hauts sont incomplets dans la mesure où ils se donnent des hypothèses qui ne sont pas vérifiées. Notre objectif est de formuler une étape méthodologique supplémentaire dont le but est de fournir une caractérisation du domaine d'application (c'est-à-dire une tâche à réaliser dans un domaine) qui vérifient les hypothèses formalistes.

En fait, quel est le problème, d'un point de vue pratique ? Le problème est de savoir ce que signifie un terme dans un domaine, de savoir à quelle notion il renvoie pour être capable de décrire d'une part quels objets du domaine sont mobilisés et d'autre part quels raisonnements il autorise. Il faut en outre noter que les notions n'existent pas dans un univers où elles seraient directement accessibles, mais qu'elle ne peuvent être déterminées qu'à partir de leurs expressions linguistiques. En d'autres termes, comme le verrons dans le chapitre 5, les notions sont des invariants sémantiques que l'on reconstruit dans le contexte d'une tâche et d'un domaine.

L'enjeu est par conséquent, si l'on veut profiter de l'acquis méthodologique lié aux hypothèses formalistes, d'être capable de caractériser le sens des notions de manière à ce qu'il soit compatible avec son exploitation formaliste. Par exemple, on a vu que l'objectif des travaux ontologiques est de compléter les formalismes de représentation de connaissances d'appareils formels permettant de restreindre les interprétations possibles des représentations formelles aux significations visées. Le problème savoir si l'on est capable de déterminer ce qu'est la signification visée. Ce que des travaux tels que ceux de Guarino permettent de faire, c'est de traduire en contraintes formelles sur l'interprétation la signification visée des termes du domaine. Ainsi, si l'on veut caractériser une pomme comme étant une espèce naturelle, c'est-à-dire comme un type rigide, on peut énoncer l'axiome correspondant à la rigidité du type. Mais il est fort délicat de savoir ce que doit vouloir signifier des termes aussi simples et banals d'apparence que « sténose » ou « estomac ». Ce que cela signifie, c'est que la formalisation de contraintes sémantiques sur l'interprétation ne peut pas faire l'économie d'une description du sens visé que la formalisation traduit en contraintes formelles.

Le chapitre suivant s'attache à déterminer la nature du sens visé et les moyens de le caractériser : le sens visé pour une notion est ce que nous appelons une « essence différentielle ». Cette notion a pour fonction d'assurer la transition entre d'une part la non formalité du domaine et d'autre part la formalité des outils dont on dispose. Le fait pour une essence d'être « différentiel » provient de ce qu'elle se construit à partir d'une description du domaine en un réseau d'identités et de différences entre unités de sens. La détermination du sens des notions à partir de ce réseau obéit à des principes non compositionnels et contextuels. Mais, à partir d'une description du domaine comme un réseau d'identités de différences, il est possible de le *normer* pour fixer un *système* d'identités et de différences dans lequel les positions ou valeurs dans le système sont des « essences » en un sens husserlien, c'est-à-dire des invariants éidétiques permettant de définir un objet. En d'autres termes, ce travail de normalisation sémantique permet de se donner le sens visé à partir duquel une formalisation des contraintes sémantiques que ce sens induit est faisable. La notion d'essence différentielle est ce qui permet de terminer l'avancée méthodologique que constitue la conception de composants ontologiques réutilisables. Elle sera mise en œuvre dans le chapitre 4.

C'est cette notion d'essence différentielle qui prend en charge la dimension synthétique *a priori* du domaine qui échappe par principe à une approche formelle. La normalisation qui fixe la sémantique du domaine en un système d'identités et de différences est l'étape méthodologique qui, à partir du matériau linguistique, dégage les principes synthétiques *a priori* du domaine. Il est temps à présent d'expliciter des notions philosophiques en faisant un parcours rapide de quelques théories philosophiques dans le chapitre suivant.

Chapitre III

Ontologie et philosophie

L'objectif de ce chapitre est d'adopter un point de vue épistémologique sur les distinctions ébauchées dans le chapitre précédent. En particulier, il s'agit de préciser la notion de synthétique *a priori* et de montrer comment la notion d'essence différentielle est celle qui permet de décrire les connaissances d'un domaine.

Le plan du chapitre est le suivant. Nous présentons (section 3-1) le problème que constitue la notion de synthétique *a priori* d'une part dans le cadre de l'IA et des SBC (section 3-1.1) et d'autre part d'un point de vue plus philosophique (section 3-1.2). Cette discussion conclue sur l'adoption d'une perspective transcendantale au sein de laquelle nous distinguons d'une part l'attitude critique, héritée de la philosophie transcendantale kantienne, qui permet de traiter en priorité les connaissances dites scientifiques, et d'autre part l'attitude phénoménologique, héritée de la phénoménologie husserlienne qui, reposant sur une conception plus large de la connaissance, nous permet de dégager les principes qui nous sont nécessaires pour traiter le problème du synthétique *a priori* en IA. Notamment, la phénoménologie husserlienne développe une problématique explicitement ontologique en opposant ontologie formelle et ontologie matérielle, problématique qui nous fournit les principes de notre méthodologie, principalement la notion d'essence différentielle.

1 Le problème du synthétique a priori

1.1 Le synthétique a priori en IA

L'IA a pour objectif d'opérationnaliser des connaissances exprimées en langue naturelle pour résoudre automatiquement des problèmes. L'IA doit donc disposer d'un principe, que nous avons appelé principe de modélisation, lui permettant de caractériser d'une part en quoi les connaissances en langue naturelle sont des connaissances, et d'autre part de dégager, pour le représenter formellement, le savoir qu'elles véhiculent.

En particulier, ce qui nous intéresse ici, d'un point de vue épistémologique, c'est de caractériser en quoi consistent les connaissances du domaine, et en particulier en quoi consiste une notion ou un concept du domaine. Pour cela, nous distinguons pour caractériser une notion

son contenu conceptuel, qui correspond à ce qui est visé, ce qui est dit ou compris dans cette notion ;

son contenu gnoséologique, qui correspond à la manière dont le contenu conceptuel possède une réalité dans le monde, à la manière dont le contenu conceptuel est rapporté à une existence dans le monde ;

son contenu prescriptif, qui correspond au fait qu'une notion possède un contenu conceptuel et un contenu gnoséologique.

Pour reprendre des catégories philosophiques classiques (Gilson, 1981), le contenu conceptuel correspond à l'essence, et le contenu gnoséologique à l'existence de l'essence. Le contenu gnoséologique précise en quoi le contenu conceptuel détermine l'existence et s'y rapporte. Le contenu conceptuel correspond d'une certaine manière au sens de la notion : une notion dépourvu de contenu conceptuel est une notion absurde. Par exemple, c'est la notion de « cercle carré ». Le contenu conceptuel correspond donc au « sens » husserlien (voir *infra*). Le contenu gnoséologique correspond à la vérité de la notion, le fait qu'elle véhicule une vérité sur le monde. Une notion peut avoir du sens, posséder un contenu conceptuel sans avoir de contenu gnoséologique, c'est-à-dire sans être une connaissance sur le domaine.

Le but, dans la modélisation des connaissances d'un domaine, est d'être capable de disposer d'une représentation formelle des connaissances qui conserve leur contenu prescriptif : le fait qu'elles signifient quelque chose et le fait que cela est vrai dans le domaine.

1.1.1 Le problème des ontologies primitives

On a vu dans le chapitre précédent que le principal problème rencontré dans les travaux formels consiste dans la notion d'ontologie primitive qu'il fallait supposer donnée avec le domaine. Autrement dit, les objets existants du domaine sont supposés connus d'emblée, si bien qu'il n'est pas nécessaire d'élaborer une modélisation explicite de ces objets.

Il faut en effet distinguer, dans les ontologies du domaine deux types d'ontologie, les ontologies primitives et les ontologies pragmatiques ou abstraites. Les premières contiennent les objets primitifs qui d'une part existent, ce sont les existants du domaine, et d'autre part fondent l'existence de tous les objets que l'on peut rencontrer dans le domaine. Les secondes, les ontologies pragmatiques, contiennent une représentation du domaine dans les termes qui sont adaptés au raisonnement qui va être mis en œuvre pour résoudre le problème. Les ontologies pragmatiques sont fondées sur l'ontologie primitive dont elles sont une abstraction effectuée du point de vue de la méthode de raisonnement. L'ontologie pragmatique est la vue que la méthode impose sur le domaine.

Les ontologies pragmatiques mobilisent les termes du domaine dans les méthodes de résolution en supposant d'une part que leur contenu conceptuel est déterminé par leur rôle logique dans le raisonnement et d'autre part que leur contenu gnoséologique consiste dans le fait de faire référence à un objet existant du domaine, ou à une abstration d'objets existants. Les termes de l'ontologie pragmatique se fondent sur les objets primitifs existants du domaine représentés dans l'ontologie primitive.

La signification d'un terme de l'ontologie pragmatique est déterminée, étant donné la manière de constituer ce type ontologie, par le raisonnement dans lequel il est enrôlé. Ce que signifie un terme dans une hiérarchie tout/partie, est déterminé par le raisonnement méréologique effectué sur cette hiérarchie. Le contenu conceptuel d'un terme est donc fixé par les relations formelles établies par le raisonnement.

Par ailleurs, un terme de l'ontologie pragmatique contient un contenu gnoséologique dans la mesure où il est fondé sur les objets primitifs de l'ontologie primitive. Le terme renvoie à une connaissance puisque d'une part il renvoie à des objets existants du domaine et d'autre part il en affirme les prédicats formels posés par le raisonnement.

Or, les inférences établies par le raisonnement doivent préserver la vérité dans le domaine. Cela signifie que le contenu conceptuel, c'est-à-dire le potentiel inférentiel du terme dans le cadre du raisonnement, préserve la vérité dans le domaine. Or, la vérité dans le domaine, ce sont les objets primitifs du domaine, par définition de l'ontologie primitive. Par conséquent, il est nécessaire que le contenu conceptuel reliant les termes entre eux reste constamment en accord avec leur contenu gnoséologique, *i.e.* avec les objets qu'ils dénotent.

Le principe de la solution adoptée par les travaux formels que nous avons vus consiste d'une part dans le fait que le domaine consiste en une base de référents et d'autre part dans le fait que les termes ou concepts du domaine dénotent ces référents. Les inférences explicitant le contenu conceptuel des termes préservent la vérité dans la mesure où elles sont formulées de manière à construire la dénotation de terme inféré à partir de la dénotation des termes prémisses, si bien que le contenu conceptuel du terme inféré renvoie toujours à un contenu gnoséologique. Par conséquent, tout repose sur la capacité de la notion de dénotation à articuler contenu conceptuel à contenu gnoséologique.

1.1.2 La dénotation

L'épistémologie logiciste s'est bâtie autour du renouveau que Frege imprima à la logique à la fin du siècle dernier lorsqu'il fonda la logique dite mathématique par son projet d'une *Begriffsschrift* et la distinction cardinale prédicat/argument. Mais, fondamentale pour cette épistémologie, la notion de dénotation est sans doute le legs ontologiquement le plus important.

La relation de dénotation est celle que l'on trouve dans la théorie logique des modèles où la valeur sémantique d'une construction syntaxique est un objet déterminé dans un univers de référence. Cette conception dénotative où les expressions renvoient à un objet pris dans un univers de référence, souvent le monde réel, est en effet due à Frege :

> « Or, il est naturel d'associer à un signe (nom, groupe de mots, caractères), outre ce qu'il désigne et qu'on pourrait appeler sa dénotation, ce que je voudrais appeler le sens du signe, où est contenu le mode de donation de l'objet. » (Frege, 1971b)

Ainsi, la signification d'un terme, sa *Bedeutung* que l'on traduit habituellement par référence ou dénotation, est l'objet du monde qui lui correspond. Pour Frege, l'univers du discours, l'univers de référence, est unique et ne peut être que l'univers réel, dans sa globalité, comprenant des objets physiques et d'autres non physiques, comme le Vrai et le Faux, qui sont, comme on le sait, les dénotations possibles des propositions. La vérité comme valeur sémantique n'est donc pas relative à un univers particulier de référence, on dira relative à un modèle, mais absolue. A ce titre, la logique devient une théorie du monde :

> « Pour Frege [...] il ne peut être question de changer d'univers. On ne peut même dire qu'il se restreint à un univers. Son univers, c'est l'univers. Pas nécessairement l'univers physique, évidemment, car pour Frege certains objets ne sont pas physiques. » *in* (van Heijenoort, 1967, p.325)

Le projet frégéen d'une théorie du monde provient de la volonté de réformer la langue naturelle pour obtenir des raisonnements sûrs et fiables sur les objets du monde. D'une langue ambiguë et imprécise, on obtient une langue artificielle univoque et précise. Ce dont la langue naturelle parle, elle en parle mal et confusément. L'universalité de la langue logique permet de parler de la même chose, l'univers complet, mais de façon précise et satisfaisante :

> « Les langages de Frege et Peano sont réellement des fragments du langage naturel dont la signification et la grammaire ont été rendus plus rigoureuses et avec des symboles particuliers à la place des mots conventionnels. » *in* (Hodges, 1985 1986, p.143) *apud* (Gochet & Gribomont, 1990)

La logique frégéenne est donc un langage universel dont la dénotation est l'univers complet, comprenant tous les objets physiques, scientifiques et phénoménologiques. La logique est pour Frege une théorie du monde dont elle exprime les structures intelligibles. C'est sans doute pour cela que Frege n'a pas abordé la question sémantique du point de vue formel, intra-mathématique, car la sémantique n'est pas pour Frege une mise en rapport de symboles dans le cadre des mathématiques, mais bien plutôt le considération du monde depuis une langue parfaite pour l'exprimer.

L'univers de référence devint relatif quand il fallut distinguer la réforme de la langue naturelle de l'utilisation formelle de signes sans signification susceptibles de recevoir différentes interprétations au gré des différents modèles de référence adoptés selon le cas :

> « [...] on observe deux courants : l'un va de Frege à Quine en passant par Russell. Les représentants de ce courant de pensée voient dans la logique un langage (interprété) et même une théorie qui traite des aspects les plus abstraits de la réalité. L'autre va de Boole à Hintikka en passant par Skolem. Ses représentants voient dans la logique un calcul. Cette dernière conception de la logique permet de traiter les systèmes axiomatiques non comme des collections de formules *vraies absolument*, mais comme des collections de formules *vraies relativement à un modèle*, c'est-à-dire relativement à un choix de *domaine* pour les variables et d'*interprétation* pour les constantes. » (Gochet & Gribomont, 1990, p.77)

Il faut donc thématiser la vérité relative de ces collections de formules en explicitant la relation qu'elles entretiennent avec le modèle servant de support à leur interprétation. Cette relation consiste en l'ensemble des règles de correspondance entre symboles de constantes de la théorie logique (T) et objet de l'univers (U), symboles de fonction de T et fonctions de U, symboles de prédicat de T et relations dans U.

La dénotation peut posséder un double statut. Elle peut être dans un premier cas une relation intra-mathématique entre structures formelles : il s'agit de la vérité d'une structure relativement à un modèle, lui-même objet mathématique construit. Dans un second cas, elle peut être une relation extra-mathématique entre une structure formelle et le monde, à la manière d'un rapport langage / monde, exprimant en quoi la structure formelle se rapporte au monde et en est une connaissance.

Dans le premier cas, la relation de dénotation est une notion mathématique comme une autre, et ne relève pas de l'épistémologie mais de la logique mathématique dans la mesure où elle ne prétend pas plus qu'une autre notion mathématique détenir une légitimité quelconque à décrire le réel, ou un rapport privilégié au réel permettant à des structures mathématiques de devenir, par la dénotation, des théories scientifiques du monde. De ce point de vue, la dénotation est une théorie particulière interne au mathématique, et non une relation méta-mathématique permettant de penser la valeur gnoséologique de la science.

Si l'on considère en effet que la dénotation est une relation théorie / monde, elle s'appliquerait à tout formalisme, logique et mathématique, dans la mesure où, s'exprimant dans un système formel, logique et mathématique prendraient leur signification par la dénotation qui les renverrait au monde. D'une certaine manière, les mathématiques se trouveraient incluses dans la logique puisque leur signification serait explicitée par une relation établie par la logique : la dénotation. Autrement dit, la dénotation explicitant la signification de tout langage, la logique étant la théorie de la langue parfaite, les mathématiques en tant qu'elles sont un langage seraient soumises à la logique en tant que théorie du langage. Mais, comme le souligne J. Petitot :

> « Il ne faut plus penser la question de la réalité et de l'applicabilité des idéalités mathématiques selon l'analogie d'un rapport langage-monde. Dans leur rapport au réel les théories mathématiques ne dénotent pas, pas plus d'ailleurs que les théories physiques où elles s'impliquent. Elles déterminent - légalisent - des données phénoménales, ce qui est tout à fait autre chose. Certes, la théorie des modèles internalise dans la méta-mathématique un rapport apparemment de type « langage - réalité ». Mais celui-ci est intra-mathématique et ne met donc en jeu aucune extériorité. Il reste par conséquent étranger aux questions de réalité et d'applicabilité. Poser ces questions en voulant redoubler ce rapport méta- mathématique par un rapport "ontologique" de type « langage (mathématique) - monde (réel) » revient à penser la connaissance en termes de prédication et à dénier la coupure essentielle qui existe entre la science et le sens commun. Penser la connaissance en terme de prédication, c'est en rester à une tradition métaphysique classique qui ne possède plus aucune valeur. » in (Petitot, 1991a, p. 225)
>
> « Les mathématiques ne se réduisent pas à la logique. Ce sont des sciences d'*objets* et le rapport syntaxe/sémantique est *interne* aux mathématiques elles-mêmes. » (Petitot, 1992b, p.54)

Dans le second cas, l'usage fait de la dénotation excède les mathématiques et relève de la philosophie, et même de l'ontologie car il faut bien se donner l'univers d'interprétation si ce dernier ne résulte pas de la construction mathématique elle-même. La sémantique est en effet devenu un objet de science lorsqu'elle fut internalisée dans les mathématiques *via* la théorie des modèles. En dehors de ce cadre mathématique, elle dépend du statut philosophique de la logique et des implications ontologiques qui en découlent. Or, c'est dans cette perspective que se meut le formalisme comme épistémologie, en adoptant la notion de dénotation pour rendre compte du rapport unissant la connaissance à son objet. C'est donc dire que toute science se retrouve subordonnée à la logique.

Mais une telle exploitation épistémologique de la dénotation ne va pas sans poser de nombreux problèmes. Un premier problème porte sur la nature de l'univers d'interprétation, dès lors que sa donation ne relève pas de l'investigation scientifique proprement dite, mais la précède. C'est une question d'ordre ontologique. Un deuxième problème porte sur le fait qu'une théorie formelle munie de son interprétation dénotationnelle devient une *mimesis* de la réalité qu'elle formalise : on donne au monde la forme discursive de la syntaxe et du concept.

L'approche formelle n'est donc légitime que s'il est possible d'affirmer que le monde consiste en une base de référent permettant de donner une signification gnoséologique aux concepts de la théorie. Il faut donc s'assurer que la matériau empirique à partir duquel s'édifie l'élaboration théorique remplit ces conditions : le travail scientifique consistant à alors à ajuster la structure du système formel qu'est la théorie à la réalité référentielle du domaine.

Or, nous savons que, dans notre cas, le domaine consiste en une réalité de pratiques exprimées en langue naturelle. Autrement dit, le matériau empirique à partir duquel construire une théorie du domaine ne correspond pas à une base extensionnelle d'objets dont la théorie doit décrire la structure ensembliste.

Par conséquent, la dénotation ne fournit pas la relation qu'il nous faut pour rendre adéquats contenus conceptuels et contenus gnoséologiques, pour que le raisonnement stipulé par les méthodes de résolution conserve la vérité dans le domaine parce qu'il mobilise des *connaissances* du domaine et des représentations qui, certes, possèdent un sens dans la mesure où elles sont interprétables, mais qui ne renvoyent pas nécessairement à des connaissances.

La dénotation ne pourra remplir sa fonction gnoséologique qu'à partir du moment où, à partir du matériau empirique du domaine, pourra être dégagée la base référentielle dont elle a besoin.

1.1.3 Connaissance et expérience possible

Il nous faut conserver, en modélisant et représentant le domaine, ce qui fait que les connaissances en langue naturelle sont des connaissanes. Or, nous avons vu dans l'introduction que les connaissances en langue expriment notre rapport au monde, notre expérience dans le monde, bref notre pratique. Par conséquent, une représentation conservera la teneur gnoséologique des expressions linguistiques des connaissances si elle conserve leur renvoi à une expérience possible dans le domaine, à une pratique possible dans le domaine. Au lieu de faire des inférences qui conservent la vérité en termes de relations entre classes de référents, il faut inférer de manière à rester dans le cadre d'une expérience possible.

Le problème est alors de caractériser ce qu'est une expérience possible. Un domaine se présente par définition comme étant ce qui regroupe un ensemble de connaissances en vue d'un type donné d'expérience possible dans le monde. Par exemple, la médecine correspond à l'expérience et la pratique thérapeutique. La physique mathématique à l'expérience et la pratique de la mesure et la théorisation des phénomènes, etc. Il faut donc être capable de caractériser l'expérience possible définissant un domaine, en tant qu'elle caractérise ce domaine-ci et non ce domaine-là.

Or, nous avons vu dans le chapitre précédent que c'est précisément ce que l'approche formelle est incapable de faire. En effet, les méthodes de résolution représentent le domaine en tant qu'il instance les raisonnements : ce sont les ontologies pragmatiques. Par ailleurs, l'ontologie formelle décrit le domaine en tant qu'il instancie des catégories formelles de la pensée ou du raisonnement, des catégories génériques comme « propriétés », « espèces », etc., et pas seulement des catégories formelles spécifiques à un raisonnement comme « hypothèse », « signe », « diagnostic ». L'ontologie formelle spécifie ce qui doit être vrai dans le domaine pour user de telle ou telle catégorie générique, mais elle ne spécifie pas comment vérifier si le domaine remplit vraiment ces spécifications. On reste au niveau d'un domaine *en général*, des objets *en général*, sans spécifier en fonction du domaine.

Par conséquent, il est nécessaire de se tourner vers un cadre épistémologique où est pris explicitement en compte le problème de l'expérience possible propre à un domaine. Ce cadre est celui de la philosophie transcendantale.

1.2 Le synthétique a priori en philosophie

Le transcendantal consiste par définition à articuler le conceptuel au gnoséologique, en montrant que le conceptuel n'a de valeur gnoséologique, *i.e.* ne vaut comme connaissance, que s'il peut se construire à partir d'une existence donnée dans le phénomène, ou s'il correspond à une construction possible selon les critères de l'expérience réelle. Cela signifie que le gnoséologique possède une dimension indépendante du conceptuel : il ne correspond plus au logiquement non contradictoire, mais à une expérience possible. Le gnoséologique, dans les philosophies transcendantales, se définit en termes d'expérience possible et d'intuition où l'être comme existence se donne à la réceptivité de la pensée, pour que l'être comme essence puisse bâtir sur sa base ses constructions conceptuelles.

Dans son principe, le transcendantal est l'attitude consistant à dire que le contenu gnoséologique d'un concept consiste dans son rapport à une expérience possible. Un concept correspond à une connaissance et pas seulement à un contenu pensé s'il renvoie à une existence possible dans le cadre de l'expérience. Il ne s'agit plus d'une possibilité logique reposant sur la non contradiction, mais d'une possibilité transcendantale reposant la manière dont le monde réel nous est donné dans l'expérience. Un concept logiquement possible, n'impliquant pas de contradiction, peut ne pas être transcendantalement possible, c'est-à-dire qu'il peut ne correspondre à aucune réalité expérimentale.

Le contenu gnoséologique d'un concept, dans une perspective transcendantale, est donc de nature prescriptive : il renvoie à une expérience possible en prescrivant d'une part quelle expérience permet de produire la réalité correspondant au concept, et d'autre part en permettant de juger si le cours de l'expérience effective confirme ou infirme le concept et son contenu essentiel.

C'est dire qu'il faut comprendre le concept, pour qu'il soit une connaissance, comme *une structure d'anticipation*. Le concept est une loi prédisant le cours futur de l'expérience et prescrivant les actions possibles à adopter. Prenons des exemples. Si l'on considère le concept d'électron, il ne constitue pas en lui-même une connaissance. Mais, dans le cadre d'une théorie physique, il lui correspond des lois physico-mathématiques déterminant quelles données expérimentales permettent de déterminer que dans une chambre à bulle, c'est bien un électron qui a été produit. Par ailleurs, ces mêmes lois prescrivent l'action possible en déterminant quelles expériences, quels montages expérimentaux, sont susceptibles de produire des phénomènes dans lesquels sera connu le concept d'électron.

Sur un plan plus cognitif, considérons par exemple notre perception d'un maison. Mais, plus exactement, nous *avons la sensation visuelle* des faces visibles de la maison, sans voir effectivement les autres, mais cependant nous *percevons* la maison, avec ses faces cachées et sa structure interne de maison. Ce qui fait que nous *sentons* les faces visibles sans les percevoir et *percevons* une maison sans la sentir, c'est que le concept de maison est une connaissance nous informant d'une part que les faces visibles senties correspondent à une maison et d'autre part, puisqu'il s'agit une maison, qu'il faut s'attendre à ce qu'elle ait des faces cachées, car elle est un objet à trois dimensions. Cela signifie que le concept de maison est une connaissance dans la mesure où il est une structure d'anticipation sur le cours futur de l'expérience, prescrivant que si je fais le tour de la maison, je devrais trouver quelque chose comme la face arrière de la maison. Ce concept de maison se construit donc à partir des sensations, au fur et à mesure que les sensations attendues sont confirmées par les sensations effectives. Par ailleurs, le concept, en prescrivant ce qui est attendu, indique quelles expériences mener pour le confirmer ou l'infirmer. En l'occurrence, pour savoir s'il s'agit d'une véritable maison ou d'un décor peint en trompe-l-œil, il suffit de contourner les faces visibles [1].

Le concept comme structure d'anticipation traduit la nécessité logique de l'essence en légalisation du cours de l'expérience, comprise comme mise en concordance du donné empirique phénoménal avec les anticipations perceptives. L'anticipation n'est possible que si le concept se comprend comme la méthode de construction de l'objet perçu à partir des données perceptives. C'est en effet la méthode de construction qui prescrit quelles données lui permette d'opérer et quelles données sont nécessaires pour la poursuite de la construction. Si l'objet est donné d'emblée, que ce soit par son concept comme contenu pensé (conceptualisme dogmatique), ou comme référent possible (empirisme logique), il ne peut réguler le cours de l'expérience dans le mesure où soit il en est totalement séparé, soit il en épuise la donation d'un seul coup, par la dénotation.

Il faut donc que l'ontologie de l'essence pense ses concepts comme des méthodes de construction ou de constitution. Ce n'est que dans ce cadre qu'il est possible de redonner au concept un contenu gnoséologique qu'il semblait avoir définitivement perdu. Notre perspective sera donc résolument transcendantale et nous considérons à présent les philosophies kantiennes et husserliennes. Nous ne citons la position critique que pour mémoire dans la mesure où nous y reviendrons longuement dans notre quatrième chapitre et où la construction transcendantale kantienne, élaborée exclusivement pour les concepts scientifiques, n'aborde pas les concepts de la langue naturelle, fût-elle de spécialistes. C'est en revanche ce que permet de faire la phénoménologie husserlienne. Mais, on peut rétorquer que le formalisme est en lui-même une construction, et qu'il n'est pas légitime de lui opposer une perspective transcendantale sous prétexte que cette dernière est constructive. C'est qu'en fait, il ne s'agit pas du même type de construction.

1. S'il s'agit d'un décor et non d'une maison, il n'y aura pas de confrontation explicite à la conscience d'une sensation visuelle brute et d'une attente conceptuelle déçue, mais le remplacement de la perception de la maison par la perception d'un décor. Le doute perceptif se traduit par la concurrence des deux concepts. Il faut se rappeler que l'on ne perçoit jamais des sensations, mais qu'elles ne sont que senties. En revanche, ce qui est perçu, c'est l'objet transcendant maison ou décor.

1.2.1 Constructions logique et transcendantale

Passer de la construction logique à la construction transcendantale, ou de la dénotation à la construction comme accès à l'objet connu, est la démarche que nous proposons pour résoudre le problème que nous avons rencontré dans la perspective formaliste. Comme nous venons de le voir, notre solution consiste à renouer avec un thème souvent controversé dans la littérature épistémologique, le synthétique *a priori*.

L'empirisme logique est un constructivisme logique où toute notion de synthétique *a priori* est bannie. En effet, selon cette approche, toute proposition scientifique se construit à partir des énoncés d'observation. C'est ainsi que chez Carnap, toute proposition est dotée d'une signification, *i.e.* est une connaissance, si elle se construit logiquement (est une conséquence logique) d'énoncés dits protocolaires décrivant le faits observés. Il s'agit donc d'un empirisme logique puisque tout dérive de l'expérience (empirisme) et que cette dérivation est logique.

Mais, pour qu'il s'agisse d'un authentique empirisme, il faut bien, pour reprendre une expression connue, que toute connaissance commence avec l'expérience, et qu'elle résulte toute de l'expérience [2]. Si la logique permet de dériver toute connaissance de l'expérience, la connaissance ne sera à proprement parler empirique que si la logique n'altère en rien le contenu de l'expérience, qu'elle n'y ajoute ou retire aucun contenu gnoséologique. En d'autres termes, il faut que les outils formels de la science, la logique et les mathématiques, soient analytiques et non synthétiques.

> « Les jugements de l'arithmétique, de la géométrie, certains principes de la physique tels qu'ils sont donnés par Kant comme exemples de connaissance *a priori* seront discutés par la suite. C'est justement dans le refus de la possibilité d'une connaissance synthétique *a priori* que réside la thèse fondamentale de l'empirisme moderne. La conception scientifique du monde ne connaît que des énoncés d'expérience sur des objets de toutes sortes, et les énoncés analytiques de la logique et des mathématiques. » (Hahn *et al.*, 1985, p.118)

L'outil logico-mathématique ordonne les faits, mais ne leur ajoute rien : il est analytique. Cela signifie les mathématiques et la logique ne recèlent aucune vérité en elles-mêmes: elles ne véhiculent que des tautologies (toujours vraies, quel que soit le monde réel, elles n'en disent rien) ou des contradictions (toujours fausses, quel que soit le monde réel, elles n'en disent rien non plus). Pour disposer d'une science, il faut obtenir des vérités caractérisant ce monde-ci *i.e.* le monde réel. Puisque les langages de la science sont incapables de nous donner de telles vérités (tautologies et contradictions traitent autant du monde réel que des univers irréels), la seule source de vérités scientifiques est l'expérience. La science, issue de l'expérience, est donc dite synthétique *a posteriori*. Elle ne peut recéler du synthétique *a priori* : dans la science, ce sont les mathématiques et la logique qui sont *a priori* car elles permettent d'organiser logiquement les données de l'expérience pour en déduire les énoncés théoriques. Mais cette organisation de l'expérience ne lui ajoute rien, elle est analytique et non synthétique. Ce qui est *a priori* est nécessairement analytique. La science est donc un langage qui sert à exprimer le monde sans rien lui ajouter.

C'est donc précisément ce qu'il faut remettre en cause. En effet, si la construction logique des objets est purement analytique, il est nécessaire qu'il y ait une mimésis entre

2. « Mais bien que toute notre connaissance commence *avec* l'expérience, elle ne résulte pas pour autant toute *de* de l'expérience. » (Kant, 1980, p.63).

le monde et la théorie qui la vise : en effet, d'où pourrait venir un écart quelconque entre les deux, si la science ne fait qu'ordonner les faits scientifiques ?

Si la construction doit porter également sur les objets, il faut donc que l'outil de la science possède des principes de construction, qui par définition seront *a priori* et synthétiques. *A priori*, car le principe de construction précède nécessairement le construit. Synthétique, car une construction doit nécessairement être synthétique pour que l'objet construit ait une unité, une cohésion propre, et ne soit pas simplement la somme de ses éléments constitutifs. Or nous avons vu que le problème est précisément d'être capable de construire, de déterminer ce que peut être l'objet dans un domaine. Alors que la construction logique se donne l'objet pour en faire la référence de ses théories, construisant alors la référence des expressions symboliques à partir des référents élémentaires ou primitifs, la construction transcendantale permet de constituer l'objet que l'on veut viser dans une connaissance. Mais, comme nous l'avons dit, rien n'empêchera, ce travail effectué, de renouer avec la construction logique : mais cette dernière fonctionnera alors à l'interieur même de la modélisation, entre le système formel et un univers d'objets construits qu'il dénote, et non entre la théorie et le monde. En d'autres termes, on fait de la dénotation une relation intra-théorique, et non une relation extra-théorique, c'est-à-dire entre la théorie et le monde.

C'est donc à une conception de la science comme construction d'objet qu'il faut revenir, et non comme langage, à la science comme *organon* construisant les objets à partir des faits, synthétisant les faits en objet grâce à des principes constructifs de nature logico-mathématique, et non à la science comme *canon* formel ordonnant, analysant, les faits en classes appelées objets.

1.2.2 La solution critique

La persective transcendantale remonte bien évidemment à Kant[3]. Dans la philosophie kantienne, un concept, contenu pensé, reçoit un contenu gnoséologique quand il se schématise, c'est-à-dire qu'il entre en correspondance avec une structure spatio-temporelle déterminant le cadre *a priori* de la sensibilité, faculté réceptive à laquelle les phénomènes, c'est-à-dire le matériau empirique, sont donnés. Les phénomènes correspondent à ce qui est donné : c'est l'existence. Comme ils ne sont pas donnés n'importe comment, mais toujours dans le cadre de formes *a priori*, l'espace et le temps, un concept correspond à une expérience sensible possible, à une donation possible de phénomène, que s'il correspond à une détermination possible de l'espace et du temps.

Le rapport de la connaissance conceptuelle à la réalité effective comprend deux moments fondamentaux : la modélisation et la schématisation. La modélisation correspond à la détermination mathématique de la forme des phénomènes donnés à la réceptivité sensible, à savoir la forme spatio-temporelle des phénomènes. En effet, temps et espace sont des entités continues permettant l'application du calcul différentiel. Un modèle donne donc les équations qui légalisent la forme des phénomènes. Il correspond au modèle d'observation

3. Nous présentons dans le chapitre 5 un exposé de la philosophie kantienne où on trouvera une bibliographie. Nous n'exposons ici que quelques principes généraux pour illustrer notre propos. De plus, nous énonçons ici des principes d'inspiration kantienne, mais qui ne sont pas strictement kantiens. En effet, dans notre opposition entre modèle d'observation et modèle d'explication, nous utilisons la réactualisation du motif transcendantal par J. Petitot (Petitot, 1991a). Nous y revenons également dans le chapitre 5.

que l'on rencontre dans les disciplines de mathématiques appliquées comme l'automatique, par lequel le comportement d'un système est décrit par une équation d'observation.

Mais un modèle d'observation n'est pas un modèle explicatif dans la mesure il ne correspond pas des concepts d'objet possédant des rapports de causalité entre eux. Le modèle d'observation décrit, il n'explique pas. Pour obtenir une explication, il faut faire correspondre aux équations un concept d'objet dont le contenu physique soit précisément des phénomènes dont la forme est celle que légalise le modèle d'observation. Le concept est dit alors se schématiser dans ces équations. La schématisation permet de voir dans le modèle d'observation une explication objective et causale des phénomènes observés.

Prenons comme exemple la chute des corps. Les phénomènes donnés sont l'observation de la chute de certains corps. La forme de ces phénomènes correspond à la mesure des espaces parcourus et des temps écoulés. La modélisation, *i.e.* l'établissement d'un modèle d'observation, correspond à la légalisation de ces mesures en une équation fixant leur rapport. Typiquement ici, une telle équation sera $x = 1/2gt^2$. Enfin, la schématisation, l'établissement d'un modèle d'explication, correspond à l'introduction de la notion de masse pesante et de la loi de la gravitation: $F = kmm'/r^2$. Le modèle d'explication mobilise des lois portant sur des objets physiques, alors que le modèle d'observation mobilise des équations portant sur des mesures.

La solution critique permet d'envisager le cas de la connaissance scientifique, en l'occurrence la connaissance scientifique mathématisée comme la physique. Mais, ce faisant, elle ne rend pas compte des connaissances non scientifiques qui reposent sur une rationalité non mathématique. Se saisissant des phénomènes, la solution critique montre comment des objets sont construits et qui, dès lors, ne correspondent plus aux objets de la perception empirique sensible, ceux du sens commun. Si la solution critique permet de comprendre comment un contenu conceptuel peut être élaboré de manière à ce qu'il possède toujours un contenu gnoséologique, cette solution ne vaut que pour l'expérience scientifique, sans montrer comment la connaissance non scientifique possède un contenu conceptuel valant également pour l'expérience, quoique non scientifique. Il convient d'élargir la perspective pour considérer comme le concept s'ouvre à la réalité.

2 Phénoménologie et ontologie

La solution phénoménologique consiste à considérer ce qui dans la connaissance, de manière générale, lui confère sa valeur gnoséologique. Le principe fondamental est de considérer les choses elles-mêmes, telles qu'elles se donnent, telles qu'elles se manifestent.

La phénoménologie transcendantale husserlienne[4] aborde explicitement la probléma-

4. Il n'est pas question de d'entreprendre ici un exposé de la phénoménologie. Les études d'E. Lévinas (1988; 1989) font habituellement références, ainsi que celles de R. Schérer (1962) consacrées aux *Recherches Logiques*. (Bachelard, 1957) aborde la logique transcendantale de Husserl. (Granel, 1968) étudie le temps et la perception chez Husserl ainsi que (Ricoeur, 1987). Les traductions récentes des œuvres de Patočka (1988; 1992) donne une introduction originale à phénoménologie par son grand exégète tchèque. (Dreyfus & Hall, 1982), (Miller, 1984) et (Fisette, 1994) illustrent l'interprétation que reçoit la phénoménologie husserlienne dans sa confrontation avec la philosophie analytique d'inspiration frégéenne. La revue *Études philosophiques* a consacré un numéro spécial (1991-I) sur la phénoménologie et la psychologie cognitive où se confrontent ces interprétations d'inspiration analytique avec les interprétations plus continentales. Nous

tique ontologique et propose des distinctions cardinales qui nous permettront de fonder et élaborer notre méthodologie. La phénoménologie oppose la logique transcendantale à la logique formelle, où il ne s'agit pas seulement de prendre en compte la forme de la connaissance mais également son contenu, sous l'aspect de ses conditions de possibilité. C'est pour cette raison que la tradition phénoménologique paraît plus féconde dans une réflexion sur les ontologies et la modélisation des connaissances. Pour préserver les avantages que procure une telle approche, il faudra discuter de manière plus approfondie les tentatives de formalisation de la phénoménologie husserlienne, qui rapportent la logique transcendantale à la logique formelle, faisant de celle-là le moule formel analytique de la connaissance qu'est celle-ci, renvoyant le rapport au contenu et au réel à une théorie de la dénotation.

La phénoménologie husserlienne se caractérise par sa volonté de faire « retour aux choses mêmes ». La pensée ne peut se comprendre qu'à partir des objets qui lui sont donnés, et non l'inverse. Abolissant le préjugé et le recours à des catégories prédéfinies, la phénoménologie se fonde sur la mise en présence de l'objet lui-même, de sa donation en personne. La phénoménologie se fonde sur la vérité de l'objet car elle requiert sa présence pour construire la connaissance phénoménologique. Il s'agit donc d'une philosophie de l'intuition.

L'intuition correspond à la mise en présence directe de l'objet, à sa présence « en chair et en os » : elle vient « remplir » à travers sa donation originaire la visée « en creux » qui correspond à la pensée de l'objet sans que celui-ci soit effectivement présent. On pense ainsi à l'arbre derrière la fenêtre sans que nous le voyons effectivement. Notre pensée possède un objet, l'arbre derrière la fenêtre, mais elle le pense « à vide » : l'objet est posé comme corrélat de la pensée et sa consistance objective procède la pensée qui le vise. Quand je regarde à travers la fenêtre et que je vois l'arbre, il m'est donné effectivement et sa perception remplit ma pensée à vide (Husserl, 1959; Husserl, 1950).

Selon nos catégories, la pensée à vide possède un contenu, le contenu conceptuel, mais ne constitue pas une véritable connaissance car elle ne correspond pas à une mise en présence effective avec l'objet. La perception possède un contenu qui est le contenu gnoséologique, car en mettant en présence un objet donné en chair en os, la pensée se trouve confirmée et mise en adéquation au réel.

Husserl considère trois types d'objets et trois types d'intuitions correspondantes mettant en présence l'esprit avec ces objets. Ce sont le fait empirique, l'essence matérielle et l'essence formelle ou catégoriale. Les intuitions correspondantes sont l'intuition sensible, l'intuition des essences, et l'intuition catégoriale.

L'intuition sensible donne le fait empirique. Ce dernier se caractérise par sa contingence : il est ainsi et pourrait être néanmoins autrement.

Il pourrait être autrement notamment par rapport à son *essence* qui permet de déterminer les modifications ou *variations* qui altéreraient sa nature ou au contraire ne l'affecteraient pas. Par exemple, la situation spatio-temporelle d'un objet empirique conditionne rarement sa nature, et il s'agit du *même* objet selon qu'il est placé sur une table ou sur un chaise s'il s'agit d'un ordinateur portable (version moderne de l'encrier husserlien). Cette situation spatio-temporelle est par conséquent contingente car non essentielle : la modifier

reprenons par aileurs certains aspects de la phénoménologie husserlienne dans les chapitres 6 et 7.

ne modifie pas l'objet, il s'agit toujours du même objet. En revanche, l'objet empirique doit posséder une et une seule localisation. C'est une connaissance qui vaut pour tout objet empirique : elle conditionne notre expérience de ce type d'objets. En d'autres termes, il s'agit d'une connaissance que l'on établit *a priori* par la variation et qui vaut pour notre expérience. Elle est donc synthétique et *a priori*.

Par conséquent, tout fait empirique renvoie à une essence qui détermine les conditions de son identité. L'objectivité empirique, factuelle et contingente renvoie donc à un deuxième type d'objectivité, l'objectivité essentielle. Toute objectivité, pour être telle, ne doit pas être seulement *pensée*, mais aussi *donnée*, pour valoir comme le corrélat objectif d'une *connaissance*. Cela signifie qu'il doit exister, à côté de la perception des faits empiriques ou intuition sensible, une intuition des essences :

> « *L'essence (Eidos) est un objet d'un nouveau type. De même que dans l'intuition de l'individu ou intuition empirique le donné est un objet individuel, de même le donné de l'intuition éidétique est une essence pure.* » *Ideen I* (Husserl, 1950, p.11)

Mais, de même que tout objet empirique contient le principe de l'intuition de son essence, il renvoie également à une autre possibilité tout aussi fondamentale, l'intuition de sa structure formelle ou intuition catégoriale. Tout objet empirique peut en effet être considéré sous l'aspect de sa structure formelle : ses caractéristiques instancient des formes générales, des catégories logiques, correspondant aux « concepts de propriété, de qualité relative, d'état de chose, de relation d'identité, d'égalité, de groupe (collection), de nombre, de tout et de partie, de genre et d'espèce, etc. ». Ce sont donc des objets formels et en tant que tels ils possèdent leur propre mode de donation, l'intuition catégoriale, « catégoriale » car Husserl appelle « catégories » les objets formels :

> « Nous définissons désormais comme *catégories logiques* ou *catégories de la région logique constituée par l'objet en général, les concepts fondamentaux de caractère purement logique impliqués* dans ces axiomes et par le moyen desquels l'essence logique d'objet en général reçoit ses déterminations dans le système total des axiomes, ou qui expriment les déterminations inconditionnellement nécessaires et constitutives d'un objet en tant que tel, c'est-à-dire d'un quelque chose, – dans la mesure où absolument parlant il doit pouvoir être « un quelque chose ». » *Ideen I* (Husserl, 1950, p. 40)

Il existe donc trois types d'objets recensés par la phénoménologie dont l'objectivité se distingue essentiellement par le mode de donation, impliquant la distinction de trois modes intuitifs comme mise en présence de l'Etre.

Mais, à proprement parler, si toutes les intuitions procèdent de l'intuition sensible, l'intuition des essences reposant sur la variation des intuitions particulières, l'intuition catégoriales étant fondées sur l'intuition sensible, cette dernière ne permet pas de poser au sens strict une objectivité dans la mesure où le fait empirique qui s'y donne reste contingent. C'est pourquoi seules les essences et les objectivités catégoriales méritent de plein droit l'appellation d'objet et peuvent faire l'objet de considérations ontologiques. C'est pourquoi Husserl considère les *ontologies matérielles d'essences* et les *ontologies formelles d'objectités catégoriales*.

Pour Husserl, est un objet tout ce que la pensée peut viser dans son activité noétique. Il reprend à son compte pour définir ce qu'est un objet la définition des logiciens : est objet tout ce qui peut être le sujet d'une assertion vrai. Un objet, c'est par conséquent ce qui peut être le substrat du jugement.

Mais tous les objets ne donnent pas lieu aux mêmes jugements : l'objet « rouge » ne donne pas lieu aux mêmes jugements que l'objet « vert », ce qui est compréhensible puisqu'ils n'ont pas le même contenu objectif. Mais cela n'est pas la seule manière possible de distinguer les objets ; par exemple, « rouge » et « être rouge » ne donnent pas lieu aux mêmes jugements alors qu'ils ont le même contenu objectif ; mais ils diffèrent dans leur manière d'être ce contenu : à la manière d'être une propriété (« être rouge ») ou à la manière d'être une essence (« le rouge »). Il faut donc distinguer dans un objet le contenu pensé de la manière d'y penser.

En effet, on pourra formuler le jugement « le rouge est une couleur » ainsi que « cet oiseau est rouge ». Mais on ne pourra pas écrire « être rouge est une couleur » ni « cet oiseau est le rouge ». Alors qu'il s'agit à chaque fois de penser le rouge, on y pense dans le premier cas selon le mode d'être d'une essence « rougeur » et dans le second cas selon le mode d'être d'une propriété, « être rouge ». Ces manières différentes d'être le même contenu se manifestent donc par le fait que les jugements logiques construits à partir de ces objets ne seront pas identiques. Ils conditionnent de manière différente la *forme* des jugements qui les contiennent. Cela signifie que le mode d'être d'un objet conditionne sa combinatoire dans les jugements.

Mais, le plus remarquable est que cette combinatoire des modes d'être est indépendante, jusqu'à un certain point, des contenus objectifs. En d'autres termes, la détermination des jugements qui peuvent être formulés est indépendante des contenus objectifs et ne prend en compte que les manières d'être de ces objets : il suffit de savoir qu'un objet à la forme d'une propriété pour être capable de construire un jugement le contenant, indépendamment de fait de savoir de quelle propriété il s'agit. Si l'on est capable de recenser toutes les formes possibles d'objets ainsi que les formes de jugements auxquels ils peuvent donner lieu, on obtient la possibilité d'une logique formelle, se décomposant d'une part en une *apophantique*[5] *formelle* ou *logique apophantique*, c'est l'étude de la forme possible ou forme catégoriale des jugements, et d'autre part en une *ontologie formelle* ou *logique ontologico-formelle*, c'est l'étude de la forme possible, *i.e.* forme catégoriale, des objets.

> « Le double sens corrélatif d'« évidence » et de « vérité » que nous avons mis en lumière implique aussi manifestement un *double sens corrélatif de « logique formelle »* : en partant de l'*orientation* traditionnelle *vers les jugements* en tant qu'opinions apophantiques, donc en privilégiant l'orientation de la critique, nous obtenons une *logique apophantique* qui, pleinement élargie du côté des formes catégoriales des sens apophantiques de façon à comprendre même les formes apophantiques des théories en tant que sens, atteint la *mathesis universalis*. Si nous privilégions l'*orientation vers les objectités catégoriales possibles* elles-mêmes ou plutôt vers leurs formes, alors de prime abord et d'une manière conséquente nous mettons en mouvement un *logique ontologico-formelle* mais qui, comme il va de soi, sera cependant astreinte, pour des raisons de méthode, à prendre pour objets les jugements en tant que sens, quoique seulement comme moyens, tandis que le dessein final concerne les objets. » *Logique formelle et Logique transcendantale* (Husserl, 1957, p. 198)

L'ontologie formelle structure les objets formels : ces objets sont toutes les formes possibles des objets en tant que sujet possible d'une assertion vraie. Or on peut porter

5. De *apophansis* proposition.

un jugement sur n'importe quoi : mais « n'importe quoi » c'est déjà quelque chose. Juger sur n'importe quoi, c'est juger sur quelque chose : l'ontologie formelle s'intéresse donc aux formes possibles de « quelque chose » dans un jugement ; ce sont toutes les formes possibles du « quelque chose en général », *i.e.* les objectités catégoriales. Le « quelque chose en général » est la forme la plus générale qui soit des objets : toute objectité catégoriale sera alors une détermination formelle de ce « quelque chose en général ». L'ontologie formelle rassemble donc toutes les manières possibles de penser un objet quelconque : un objet formel, c'est une manière possible de se rapporter à un quelque chose. Or, ces manières possibles ne sont pas quelconques ; elles constituent une classe close dont on connaît la structure :

> « Ce qui est objet [...] a, à l'égard de ses formes catégoriales possibles (des formes qui appartiennent à l'*objet* !), une typique tout à fait déterminée. C'est la typique des *modes du quelque-chose-en-général* : « propriété », relation, état des choses, pluralité, individualité, suite, ordre, etc. Nous les nommons *formes dérivées* du quelque-chose, de la catégorie formelle fondamentale « objet » ; » *Logique formelle et Logique transcendantale*

Ces formes s'engendrent à partir de la catégorie suprême « objet en général » qu'elles déterminent et par itération les unes sur les autres : par exemple, on peut considérer une relation de relation, une propriété de relation, etc. Si bien que l'étude de l'ontologie formelle devient une science à part entière, possédant ses lois et ses objets, et une science au statut particulier car, étudiant la forme possible des objets quelconques, elle est *universelle* :

> « De là naît une idée universelle de science, celle d'une *mathématique formelle prise dans toute son ampleur*, dont le domaine universel se délimite nettement comme l'extension du concept formel suprême : *objet en général* ou comme l'extension du quelque chose en général pensé dans la généralité la plus vide qui soit, avec toutes les formes dérivées engendrables *a priori* (et donc concevables) dans ce champ, formes qui dans une construction itérative toujours nouvelle produisent des formes toujours nouvelles. Au nombre de ces dérivations, à côté des formes : ensemble et nombre (finis ou infinis) on trouve les formes : combinaison, relation, suite, liaison, tout et partie, etc. Alors on est tout près de considérer toute cette mathématique comme une *ontologie* (doctrine apriorique de l'objet), mais comme une ontologie *formelle*, rapportée aux modes purs du quelque chose en général. On aurait ainsi acquis l'idée directrice pour déterminer dans des examens structurels aprioriques les domaines particuliers de cette ontologie, de cette mathématiques des objectités en général. » *Logique formelle et Logique transcendantale* (Husserl, 1957, p. 107-108)

Mais l'idée d'ontologie formelle est inséparable de l'idée de logique formelle. Les objectivités catégoriales ne sont posées que dans le cadre des jugements logiques, si bien que toutes les objectités catégoriales se dérivent de la catégorie suprême « objet en général » au moyen du jugement logique lui-même :

> « ces modes [les objectités catégoriales] sont dérivés de cette catégorie fondamentale [objet en général]... *par le moyen du juger lui-même*, et aussi par toutes les activités doxiques qui constituent le sens le plus large de jugement. La propriété en tant que forme prend naissance originellement dans le simple jugement catégorique de même que la collection prend naissance originellement dans l'activité de colligation, naturellement pas comme *datum* psychique réel mais comme corrélat intentionnel de la position. » *Logique formelle et Logique transcendantale* (Husserl, 1957, p. 154-155)

Alors que la logique formelle systématise toutes les manières de penser et constitue donc une logique du jugement ou apophantique formelle, l'ontologie formelle systématise tous les modes possibles de ce qui est pensé, et constitue donc une analytique ontologico-formelle. Il est donc impossible de comprendre le projet d'une ontologie formelle si l'on ne considère pas en même temps celui d'une logique formelle. Chez Husserl, ce projet s'articule en trois étapes. La première étape s'intitule la *morphologie pure des significations*, qui correspondrait aux règles de construction des formules bien formées dans les systèmes formels modernes. La morphologie pure répertorie les formes possibles du jugement comme « S est p » et ses formes dérivées : « Sp est q », etc.

La seconde étape s'intitule *logique de la conséquence* ou *analytique de la non-contradiction*. Cette logique a pour objet le jugement *distinct* qui un jugement complètement explicité et articulé dans expression logique : il n'est pas formulé de manière *passive* mais de manière *active* dans la mesure où on a *distinctement* conscience de sa structure formelle et de ce qu'elle implique :

> « [...] d'une part la conscience vide peut s'écouler de la manière suivante : elle n'a pas de séparations internes et elle *n*'est *pas* une organisation de pensées vides particulières ; au contraire elle peut s'écouler sous le mode d'une conscience vide organisée, qui se manifeste comme organisée. [...] Ainsi une conscience vide non-organisée peut être transformée en un conscience vide organisée « correspondante » et dans cette transformation le sens contenu dans l'opinion confuse (étant soumis à l'identification qui amène à la coïncidence avec le contenu organisé et qui est du type de l'« explicitation ») se « déploie » en tant que *résultat de l'explicitation*, en tant qu'opinion proprement dite avec pour contenu le contenu qui auparavant était confus de manière globale. » *Logique formelle et Logique transcendantale* (Husserl, 1957, p. 81-82)

L'analytique de la non-contradiction renvoie donc à ce qui correspond, dans notre moderne théorie des systèmes formels, à la caractérisation de ce qu'est une démonstation et un théorème, dans la mesure où la conception distincte de la structure organisée des jugements permet de déduire les jugements qu'ils impliquent *vi formae*. La troisième partie de la logique s'intitule *logique de la vérité* ou *analytique de la vérité possible*. Correspondant à notre moderne sémantique ou théorie des modèles, l'analytique de la vérité s'en distingue dans la mesure où elle se trouve concernée par le contenu pensé et pas exclusivement par la forme de ce qui est pensé.

Entre l'analytique de la conséquence ou de la non-contradiction et l'analytique de la vérité, Husserl réactive la distinction entre penser et connaître. Alors que la logique de la conséquence consiste dans la possibilité de la pensée, la logique de la vérité consiste dans la possibilité de la connaissance. Il ne suffit pas pour cette dernière que soit conceptuellement distinct le jugement, il faut que soit donnée à travers lui la chose jugée « en chair et en os » : il faut qu'il y ait donation de l'être pour que le jugement soit vrai :

> « Ici se séparent *deux évidences*, celle dans laquelle le *jugement lui-même, en tant que jugement*, vient à être donné en personne : le jugement s'appelle alors aussi jugement distinct, provenant de l'effectuation de jugement réelle et proprement dite. Deuxièmement, nous avons affaire aussi à cette *évidence* dans laquelle vient à être donné en personne *ce à quoi veut arriver « à travers » son jugement celui qui juge*, c'est-à-dire celui qui veut *connaître* – comme se le représente toujours la logique.
> [...]
> mais c'est seulement un *juger dans la plénitude de la clarté* qui peut être *connaissance*

> *actuelle* et il existe alors l'*évidence nouvelle* que sont données « elles-mêmes » les choses elles-mêmes, qu'est donné « lui-même » l'état des choses lui-même, choses et état des choses qui sont visés dans le juger tendant vers la connaissance, même là où le juger manquait encore complètement de clarté et n'était pas rempli par l'intuition. » *Logique formelle et Logique transcendantale* (Husserl, 1957, p. 85-86)

L'objet d'étude de la logique de la vérité n'est donc plus le jugement distinct, mais le jugement vrai. La logique formelle est en effet la logique de la pensée pure, la logique de la raison pure :

> « *La raison pure s'élève non seulement au-dessus de tout ce qui est fait empirique, mais aussi au-dessus de toutes les sphères essentielles se rapportant à l'hylé, au concret.* La raison pure désigne le système fermé sur soi des principes purs qui précèdent tout *a priori* se rapportant à l'hylé (et de toutes les sciences s'occupant de cet *a priori*) et qui dominent pourtant d'autre part – c'est-à-dire quant à la forme – ces sciences elles-mêmes en tant que formations de la raison. » *Logique formelle et Logique transcendantale* (Husserl, 1957, p.41-42)

Indépendante donc de toute donation de la vérité, ne consistant que dans la pure conception du sens, la logique formelle ne peut consister que dans les deux couches que nous avons distinguées, la morphologie des significations et la logique de la conséquence :

> « Une théorie systématique *pure* de la région du *sens* – dans son acception stricte, dans son acception propre, – est constituée par l'analytique uniquement si l'on considère les deux premières couches : la *morphologie pure des sens* (ou significations) et l'*analytique pure de la non-contradiction.* » *Logique formelle et Logique transcendantale* (Husserl, 1957, p. 186)

En revanche, la logique de la vérité doit refléter dans ses lois les possibilités de la connaissance et pas seulement celles de la pensée. Elle est donc authentiquement transcendantale dans la mesure où il faut tenir compte de ce qui est donné dans l'expérience de manière *a priori*. Il faut donc définir ce qui, dans le jugement logique, renvoie à l'expérience. Or, cela ne fait pas de doute : c'est par les substrats non syntaxiques du jugement que se fait le contact avec l'expérience réelle. En effet, tout jugement possède des variables tenant lieu des objets sur lesquels s'applique le jugement. Ces objets peuvent être eux-mêmes des composés syntaxiques et posséder une structure catégoriale; par exemple le jugement « Sp est q » définit comme valeur de la variable sujet l'objet catégorial « Sp », obtenu par nominalisation du jugement « S est p ». Mais, si l'on remonte le cours des compositions catégoriales des objets-substrats du jugement, on est assuré d'arriver à des objets substrats exempts de toute composition syntaxique ou catégoriale. Ces substrats sont dont purement *matériels*, c'est-à-dire totalement indéterminés d'un point de vue formel.

> « Mais on doit voir *a priori* avec évidence que *tout jugement possible aussi bien que réel* nous *ramène*, si nous nous occupons de ses syntaxes, *aux noyaux derniers*, on doit donc voir que le jugement est une construction syntaxique, même si éventuellement elle a une médiateté de haut degré, faite de *noyaux élémentaires qui ne contiennent plus de syntaxe.* » *Logique formelle et Logique transcendantale* (Husserl, 1957, p. 274)

Si les lois de la logique de la conséquence doivent valoir pour la logique de la vérité, il est donc nécessaire qu'elles valent pour les matériaux non syntaxiques des jugements. En effet, dans une telle logique de la vérité,

> « Les lois de la conséquence et de la non-contradiction deviennent des lois de la vérité matérielle possible. Bien que dans ces lois les contenus matériels (les noyaux) restent des indéterminés, ils sont maintenant pensés dans cette généralité comme rapportés pourtant à une objectité possible. La vérité possible en tant que justesse veut dire en effet possibilité de l'adéquation aux choses possibles elles-mêmes. Ainsi par exemple dans les lois formelles de la justesse possible des jugements prédicatifs sont incluses *eo ipso* les lois de la possibilité des états des choses. Ainsi donc le système mathématique de la logique en entier acquiert une relation à l'objectité possible en général. » *Logique formelle et Logique transcendantale* (Husserl, 1957, p. 193)

Il est donc nécessaire que les lois formelles de la logique de la conséquence soient définis dans le cadre de *l'expérience possible*, cadre assurant la possibilité de réunir les objets matériels non syntaxiques dans l'unité du jugement. Il est nécessaire que ces objets puissent être donnés ensemble pour que la logique du jugement puisse juger avec *vérité* sur leur être ensemble :

> « *Avant* tout juger existe une base universelle de l'expérience ; elle est constamment présupposée *comme unité concordante d'expérience possible*. Dans cette concordance, tout « a affaire » avec tout matériellement. Mais l'unité de l'expérience peut être aussi *discordante* et cependant conforme à l'essence *de telle sorte que ce qui apporte une contradiction a une communauté d'essence avec ce qu'il contredit,* de telle sorte que dans l'unité de l'expérience qui a une cohésion et qui a encore une cohésion même quand il s'agit de contradictions, tout se tient avec tout dans une communauté d'essence. Ainsi *dans son contenu tout juger originel* et, de même, tout juger qui se poursuit avec cohésion a de la *cohésion grâce à la cohésion des choses dans l'unité synthétique de l'expérience* sur la base de laquelle il repose. » *Logique formelle et Logique transcendantale* (Husserl, 1957, p. 294-295)

C'est dans le cadre de cette unité de l'expérience que le jugement peut avoir une valeur de vérité : le vrai affirmant l'adéquation positive du prédicat au sujet, le faux affirmant l'adéquation négative. En particulier, hors de cette unité de l'expérience, il n'y a aucune raison pour que le principe du tiers exclu soit vérifié. Ainsi, si la logique formelle n'a rien à dire *a priori* sur la vérité de ses jugements, affirmant seulement les lois syntaxiques de la déduction, elle n'en présuppose pas moins, pour être capable de « passer la main » à la logique de la vérité que les substrats du jugement ont « matériellement affaire les uns avec les autres » :

> « Les considérations et la théorie de la *logique formelle* n'ont, dans leur orientation objective, rien à dire sur ce point mais chacune de leurs formes logiques avec ses S et ses p, avec tous les symboles littéraux qui interviennent dans l'unité d'une connexion formelle, *présuppose d'une manière cachée* que dans cette connexion les S, p, etc., *ont matériellement « affaire les uns avec les autres »*. » *Logique formelle et Logique transcendantale* (Husserl, 1957, p. 296)

Comment donc s'assurer de cette compatibilité matérielle assurant l'unité d'une expérience possible dans laquelle le jugement de la logique de la vérité viendra trancher pour une adéquation positive ou négative? Chaque jugement doit porter en lui, comme marque de sa genèse, le fait qu'il procède *in fine* de jugements empiriques sur des individus. L'évidence formelle qu'il possède n'est valide qu'à partir de jugements d'évidence empirique donnant les choses en elles-mêmes:

> « Le niveau le plus base auquel nous arrivons lorsqu'en suivant le fil conducteur de la genèse du sens, nous revenons en arrière, nous conduit, comme nous le savons déjà, aux *jugements portant sur des individus* et donc, pour les jugements évidents au sens de la vision des états des choses eux-mêmes, à des évidences individuelles de la forme la plus simple ; ce sont les *jugements d'expérience* purs et simples, jugements sur les données de la perception et du souvenir possibles, fournissant des normes pour l'exactitude des jugements-opinions catégoriques du niveau le plus bas relatif à l'individu. » *Logique formelle et Logique transcendantale* (Husserl, 1957, p. 282)

Cela implique que le formel procède de l'individuel, qu'il en dépend quant à sa validité. Mais, le rapport entre le formel et l'individuel est particulier. Le formel ne se fonde pas sur un parcours exhaustif des individus, mais sur le fait d'en choisir des quelconques, laissés dans une généralité indéterminée. Cela signifie que le formel n'est pas l'essence du matériel mais seulement la forme à laquelle nécessairement il se soumet pour être pensé dans des jugements. Les lois du formel laisse le matériel indéterminé. Car, par ailleurs, l'individuel permet de fonder d'autres objectités que les objectités formelles ou catégoriales, les objectités matérielles. En effet, l'individuel se prête à un examen au travers duquel il permet de constituer non pas une généralité formelle, mais une généralité matérielle:

> « Il se manifeste alors une distinction fondamentale dans les diverses sortes de *généralisations* essentielles, selon que, d'une part, elles sont effectuées au sens de l'*a priori matériel* ou qu'elles le sont, d'autre part, au sens de l'*a priori formel*. Là, dans le premier cas, nous tirons de l'individuel qui est pris à titre d'exemple *la teneur qui lui est essentiellement propre* et nous obtenons les genres, les espèces et les lois essentiels se rapportant au concret ; mais dans la *généralisation formalisante* tout ce qui est *individu* doit être *vidé* de toute teneur concrète *pour arriver au quelque chose en général.* » *Logique formelle et Logique transcendantale* (Husserl, 1957, p. 287)

Il y donc des vérités aprioriques valant pour tous les quelques choses en général, ce sont les *a priori* formels valant de façon principielle pour tout ce qui peut être pensé, et les vérités aprioriques valant pour tous les objets matériellement déterminés. Mais, s'ils sont matériellement déterminés, il est impossible d'obtenir des vérités matérielles valant pour tous ces objets, ce serait alors des vérités formelles, mais seulement pour une partie d'entre eux, pour une *région* du réel matériel. Les ontologies matérielles sont donc des ontologies régionales.

Husserl structure donc en ontologies régionales ou matérielles ce qui peut être pensé en instanciant l'ontologie formelle. Une ontologie régionale est une taxinomie, une hiérarchie d'objets matériels, *i.e.* d'essences, rassemblés sous un concept générique suprême. Par exemple, il y aura une ontologie régionale pour la cuisine, les animaux domestiques (connaissances phénoménologiques) et une ontologie régionale pour la géométrie ou la nature (connaissances scientifiques), bref, pour tout ce qui peut être pensé. A chaque essence matérielle sont subordonnés les objets et faits empiriques correspondant aux variations contingentes de l'essence. Une ontologie régionale est par conséquent la théorie d'une *ré-*

gion du réel dont elle systématise l'empiricité à travers les relations eidétiques qu'elle établit à travers ses structures.

> « Toute objectivité concrète de caractère empirique s'intègre ainsi que son essence matérielle, à un genre matériel *suprême*, à une « *région* » d'objets empiriques. A l'essence régionale pure correspond alors une *science éidétique régionale*, ou, pourrait-on dire, une *ontologie régionale*. » *Ideen I* (Husserl, 1950, p. 19)

L'ontologie régionale est une théorie du domaine et contient les vérités premières vérifiées par tous les objets empiriques. Ces vérités sont les conditions *a priori* de la possibilité de connaître empiriquement les objets du domaine. Ces vérités sont par ailleurs synthétiques dans la mesure où elles ne découlent des propriétés formelles des objets mais procèdent du domaine particulier considéré. Ces vérités sont donc synthétiques et non analytiques puisque, ne valant pas pour toutes les régions du réel, c'est qu'elles reflètent la nature d'un domaine particulier.

Mais un domaine particulier d'objets est également soumis aux lois valant pour toute objectivité, c'est-à-dire les lois de l'ontologie formelle. Lois du quelque chose en général, elles valent nécessairement pour un quelque chose en particulier, pour peu qu'il soit effectivement quelque chose, ce dont nous assure la notion d'essence. C'est ainsi que

> « L'investigation des faits empiriques ne peut manquer d'être influencée par l'abondance des connaissances qui ont un rapport pur et de validité *inconditionnée* à tous les objets possibles de la région, dans la mesure où ces connaissances dépendent pour une part de la forme pure de l'objectivité en général [c'est l'ontologie formelle], pour une part de l'Eidos de la région, cet Eidos représentant pour ainsi dire une forme matérielle nécessaire pour tous les objets de la région [c'est l'ontologie matérielle]. » *Ideen I* (Husserl, 1950, p. 19)

Valant pour tous les domaines possibles, les lois de l'ontologie formelle sont analytiques. Définissant la forme ce qui peut être pensé, elles n'en donnent pas le contenu.

> « Tandis que ceux-là [concepts tels que grandeur, partie, tout] se groupent autour de l'idée vide du quelque chose ou de l'objet en général, et sont reliés à lui par les axiomes ontologiques formels, ceux-ci [concepts tels que maison, arbre, couleur, son, espace, sensation, sentiment, etc.] s'ordonnent autour des différents genres concrets les plus généraux (*catégories matérielles*) dans lesquels sont enracinés des *ontologies matérielles*. Cette division cardinale entre la sphère ontologique « formelle » et la sphère d'essence « concrète » ou matérielle nous livre la véritable différence entre *disciplines* ou, respectivement, *lois et nécessités analytiques* a priori et celles qui sont *synthétiques* a priori ; » *in* (Husserl, 1959) *Recherche Logique* III p. [252]

Les ontologies matérielles sont donc synthétiques. Mais il faut distinguer le synthétique propre aux idéations, produisant les essences, du synthétique propre aux idéalisations, produisant l'exactitude scientifique. Alors que le synthétique des idéations correspond à une vision du monde correspondant à ce que nous délivre une description linguistique, le synthétique des idéalisations correspond à une vision mathématique. Les idéalisations correspondent au passage à la limite des essences obtenues par idéation. L'idéation est le principe de la constitution des essences à partir de l'individuel et de leur intuition. Elle repose sur la technique phénoménologique de la variation éidétique :

> « Le procédé, par lequel nous pouvons créer méthodiquement cette conscience en nous [intuition des essences], est la variation. On partira d'un objet quelconque comme « modèle »; on le « variera » de façon totalement arbitraire. Il se révèle alors que la liberté ne saurait être absolue, qu'il y a des conditions sans lesquelles les « variantes » ne seraient plus des variants de ce modèle, des « exemples » de la même espèce. Cet « invariant » identifié à travers les différences, définit précisément l'essence des objets de cette espèce, ce sans quoi ils seraient inimaginables, au sens d'impensables. » in (Tran-Duc-Thao, 1992)

La variation permet l'obtention d'un noyau invariant constituant l'essence : ce noyau fixe les limites de l'identité à soi et de la différence à autrui. L'essence est obtenue par une variation fondée sur l'opposition et la différence aux plus proches voisins. Le principe de voisinage est fourni par l'imagination :

> « Elle [l'opération de dégagement d'une essence] repose sur la modification d'une objectivité d'expérience ou d'une objectivité imaginée, qui en fait un exemplaire arbitraire recevant du même coup le caractère d'un « modèle » conducteur, d'un point de départ pour la production d'une multiplicité de variantes ouvertes à l'infini : elle repose donc sur une *variation*. En d'autres termes, nous nous laissons conduire par le fait pris comme modèle, et nous suivons sa métamorphose dans l'imagination pure. Il faut pour cela que des images analogues toujours nouvelles soient obtenues dans l'imagination comme images dérivées qui, dans leur concrétion totale, soient analogues à l'image originelle. Nous produisons ainsi des variantes par un libre vouloir, chacune intervenant dans le mode subjectif de « l'arbitraire », ainsi que le processus total de variation. Il apparaît alors qu'à travers cette multiplicité de figures successives il y a une unité, que dans ces variations libres d'une image originelle, par exemple d'une chose, un *invariant* reste nécessairement maintenu comme la *forme générale nécessaire*, sans laquelle quelque chose comme cette chose, prise comme exemplaire de son espèce, serait d'une manière générale impensable. Cette forme se détache à l'intérieur de l'exercice de variation volontaire, et, la nature de ce qui différencie les variantes nous restant indifférente, elle se donne comme comme un contenu absolument identique, un quid invariable selon lequel se recouvrent toutes les variantes : *une essence générale*. » *Expérience et Jugement*, p. 411. (Husserl, 1991, p.414)

La variation opérée par l'imagination pure à partir d'un modèle fourni par l'expérience ou l'imagination pose donc un invariant qui est une essence. Cette essence correspond à ce qu'est la nature même de l'objet : « Elle se présente comme ce sans quoi un objet de cette sorte ne peut être pensé, c'est-à-dire sans quoi il ne peut être imaginé intuitivement comme tel. Cette essence générale est l'eidos... » p. 411-412.

La technique de la variation permet de découvrir l'invariant en le caratérisant comme un pôle d'identité. Husserl insiste sur le fait que l'essence définie ainsi n'a rien de « métaphysique » et que l'essence est simplement une nécessité éidétique, une nécessité de la pensée. Husserl insiste également sur le fait que les différences intervenant dans la découverte de l'essence sont indifférentes : Husserl définit en quelque sorte l'essence par un contenu positif, par des propriétés qu'elle posséderait dans toutes ses réalisations individuelles, indifféremment des différences individuelles.

Une telle conception n'est pas entièrement satisfaisante dans la mesure où on se demande ce que peut être un tel contenu positif s'il est indifférent aux différences qui l'opposent à ce qui n'est pas lui. En effet, l'essence se détermine comme étant une « identité », s'opposant à ce titre à tout le reste. Et c'est même ainsi qu'elle se révèle par la variation : le parcours de l'imagination n'est pas totalement libre, il doit respecter cette identité de l'essence sinon elle se dissout. Ce qui revient à affirmer que l'identité se constitue sur la différence, et qu'elle n'est accessible seulement par ces différences avec ce qui n'est pas elles. En d'autres termes, nous aimerions soutenir que les essences matérielles ne sont pas indifférentes aux différences avec ce qui n'est pas elles dans la mesure où ce sont seulement ces différences qui permettent de constituer leur identité. En d'autres termes, nous voulons soutenir qu'un essence matérielle n'est pas une substance, mais est de nature différentielle.

Par exemple, si j'effectue une variation sur le concept de chien, je fais librement varier dans mon imagination une image de chien. Je fais défiler dans mon imagination toute sorte d'animaux familiers. Husserl affirme que l'essence de chien sera donnée par le noyau identique de toutes mes images correspondant à des chiens. Ce qui reste commun, c'est ce que je ne peux modifier sans perdre le fait qu'il s'agit d'un chien, et non par exemple d'un chat. Par conséquent, ce qui reste commun, ce sont les propriétés qui, si elles sont modifiées, je perds l'identité de l'objet que je considère. Par conséquent, ces propriétés sont celles qui constituent l'essence en opposition aux autres objets, étant donné que c'est le fait qu'elles soient vérifiées qui constituent l'identité de l'essence. Si bien que le contenu de l'essence est bien une identité, mais une identité construite à partir des différences avec ce qui n'est pas l'essence visée. Ce que la variation met en évidence, ce sont les axes de détermination sur lesquels toute modification altère l'identité de l'essence.

C'est pourquoi nous interprétons le fait que les essences matérielles ne sont pas métaphysiques, c'est-à-dire qu'elles ne renvoient pas à une Idée transcendante au sens platonicien, par le fait qu'elles sont différentielles, puisque, si elles n'ont pas une identité par transcendance, elles ne peuvent en avoir une que par différence. Nous appellerons donc *essences différentielles* les essences obtenues par idéation, c'est-à-dire par la technique de la variation eidétique.

Par ailleurs, en sémantique différentielle, dont les principes sont exposés dans le chapitre suivant, le signifié d'une unité linguistique se construit comme la position que prend cette unité dans la langue comprise comme système. Cela signifie que le signifié se définit comme étant l'union des différences qui l'oppose à ses voisins dans le système de la langue. On ne peut être que frappé de l'identité entre la variation éidétique et la détermination différentielle du signifié linguistique. Il faut simplement noter que c'est l'imagination qui tient lieu chez Husserl du corpus permettant d'établir le système de la langue et les plus proches voisins de l'unité linguistique ou matérielle à définir. Si bien qu'il est tentant d'affirmer que les connaissances véhiculées par la langue correspondent aux essences matérielles husserliennes en faisant l'analogie entre variation éidétique dans l'imagination et commutation sémantique en corpus.

L'analogie est limitée dans la mesure où, comme nous le verrons dans le chapitre suivant, une analyse linguistique menée selon des principes différentiels ne délivre pas de connaissances, c'est-à-dire d'essences matérielles. Il lui manque ce qui est nécessaire pour qu'il y ait connaissance, c'est-à-dire un rapport à l'expérience possible. Cependant, si on ajoute aux descriptions linguistiques un tel rapport, il est alors possible de retrouver comme résultat de la variation des essences, à l'instar de la variation husserlienne.

Enfin, il faut distinguer ces essences de celles que l'on obtient en prenant leur idéalisation, c'est-à-dire en les conformant à un idéal d'exactitude mathématique/géométrique. Il faut en effet distinguer les essences anexactes obtenues par idéation, comme le rond dans la nature, des essences exactes obtenues par l'idéalisation de ces dernières, comme le cercle géométrique :

> « Les concepts géométriques sont des *concepts idéaux* ; ils expriment quelque chose qu'on ne peut « voir » ; leur « origine », et donc aussi leur contenu, diffèrent essentiellement de ceux des *concepts descriptifs* en tant que concepts exprimant des essences issues sans intermédiaire de la simple intuition, et nullement de essences « idéales ». Les concepts exacts ont pour corrélat des essences qui ont le caractère « *d'idées » au sens kantien* du mot. A l'opposé de ces idées ou essences idéales nous trouvons les *essences morphologiques* qui sont les corrélats des concepts descriptifs. » *Ideen I* (Husserl, 1950, p.138)

Les essences morphologiques sont l'objet des sciences descriptives : vagues de façon intrinsèques, ces essences ne sont pas en attente de mathématisation, elles sont par nature non mathématiques. Les essences exactes font l'objet des sciences mathématiques géométriques qui, grâce à l'exactitude, peuvent en faire une étude systématique et les appréhender de manière prescriptive. Par conséquent, ces deux types de science, descriptive ou prescriptive, sont irréductibles l'un à l'autre :

> « Il faut bien voir en outre que les sciences exactes [phénoménalité] et les sciences purement descriptives [phénoménologie] ont bien entre elles une liaison, mais qu'elles ne peuvent jamais être prises l'une pour l'autre et que quel que soit le développement d'une science exacte, c'est-à-dire opérant avec des infrastructures idéales, elle ne peut résoudre les tâches originelles et autorisées d'une description pure. » in Ideen I, p. [139]

Si nous considérons à présent toutes ces remarques du point de vue de l'IA, il est facile de remarquer que la phénoménologie transcendantale husserlienne, tout en adoptant un formalisme qu'il nous faudra critiquer plus loin, tempère néanmoins les ardeurs logicistes de l'IA comme modélisation qualitative. En effet, si Husserl reconnaît bien une espèce de primat à l'ontologie formelle, qui domine de la légalité vide toutes les régions matérielles, il affirme néanmoins une relative autonomie de ces régions puisqu'elles justifient l'élaboration d'ontologie régionale.

Par conséquent si, à première vue, les considérations husserliennes semblent entre en résonnance avec celles de l'IA, l'ontologie régionale constituant l'ontologie du domaine, l'ontologie formelle constituant son modèle pour le raisonnement, il semble que le parallèle s'arrête là dans la mesure où l'ontologie du domaine n'est en IA que la nomenclature des noms donnés dans le domaine aux structures du raisonnement, *i.e.* aux structures formelles, alors que les ontologies régionales sont des ontologies d'essences.

Ces essences ne sont pas des noms ou des termes : ce sont des concepts dont le contenu est obtenu par la variation éidétique. Manifestant une position dans un système de connaissances, les essences sont différentielles et manifestent donc un contenu qui par définition échappe à la détermination formelle.

L'opposition entre l'approche husserlienne et l'approche IA devient encore plus manifeste si l'on se souvient que l'ontologie formelle est analytique alors que l'ontologie régionale est synthétique : cette dernière contient des vérités qui lui sont propres dans la mesure où elles ne valent que pour les objets de la région. Elles sont synthétiques à double titre : d'une part parce qu'elles ne valent pour toutes les régions possibles, c'est-à-dire qu'elles ne sont pas valables pour un « quelque chose en général », autrement dit qu'elles sont pas formelles ou analytiques, et d'autre part parce qu'elles contiennent des vérités *a priori* sur le domaine. Or, dans la tradition logiciste dont procède l'IA qualitative, il ne peut exister de synthétique *a priori*, mais seulement des vérités empiriques *a posteriori*. Cela signifie qu'en tant que structure apriorique, l'ontologie du domaine ne peut manifester aucune vérité sur le domaine autres que celles provenant de l'instanciation des vérités formelles dans le vocabulaire du domaine.

A un primat de l'ontologie formelle, Husserl vient apporter la contradiction en posant une dualité du formel et du matériel, ce dernier accédant à la dignité ontologique grâce à la reconnaissance du synthéticité *a priori* de ses déterminations. Mais Husserl accorde néanmoins une subordination absolue du matériel au formel :

> « Or, cette subordination du matériel au formel se déclare en ceci que *l'ontologie formelle contient en soi en même temps les formes de toutes les ontologies possibles* (entendons de toutes les ontologies « authentiques », « matérielles ») et qu'elle prescrit aux ontologies matérielles *une législation formelle commune*; ...» *Ideen I* (Husserl, 1950, p. 22)

La philosophie husserlienne est donc un formalisme. Husserl ne met pas sur le même plan les ontologies formelle et matérielles. Le formel n'est pas un région supplémentaire comme une autre. Elle est la forme de la région, sa structure logique. Cela présuppose qu'il est possible de définir ce qu'est une forme logique indépendamment de sa matière, qu'il est possible d'établir les lois de la pensée pure indépendamment des contenus pensés.

> « L'ontologie formelle semble d'abord être sur le même plan que les ontologies matérielles, dans la mesure où l'essence formelle d'un objet en général et les essences régionales semblent jouer de part et d'autre le même rôle. C'est pourquoi on sera tenté de parler, non plus comme jusqu'à présent de régions tout court, mais de régions matérielles et de leur adjoindre la « région formelle ». Si nous adoptons cette façon de parler nous ne devons pas le faire sans quelque précaution. D'un côté nous trouvons les essences *matérielles* ; ce sont elles, en un certain sens, les *essences « authentiques »*. De l'autre côté nous avons bien encore quelque chose de caractère éidétique, mais pourtant de nature foncièrement différente : à savoir une *pure forme éidétique*, une essence certes, mais complètement « *vide* », une essence qui convient à *la façon d'une forme vide à toutes les essences possibles*, qui, grâce à son universalité formelle, tient sous sa dépendance jusqu'aux universels matériels de plus haut degré et leur prescrit des *lois* en vertu des vérités formelles qu'elle implique. Ce qu'on appelle « *région formelle* » n'est donc pas quelque chose qui est coordonné aux régions matérielles (aux régions pures et simples) ; *ce n'est pas à proprement parler une région, mais la forme vide de région en général* ; toutes les régions, ainsi que les particularisations éidétiques d'ordre matériel qu'elles enveloppent, ne sont point à côté d'elles, mais *sous* elle (en un sens purement formel toutefois). » (Husserl, 1950, p. 40-22)

La logique formelle est pour Husserl analytique : la pensée pure ne peut poser à elle seule les lois du réel. Néanmoins, à l'instar de la logique formelle kantienne, elle reste le *canon* des lois du réel à défaut d'en être l'*organon*. Les lois du réel devant être nécessairement pensées pour exister en tant que loi, elles doivent se conformer aux exigences de la pensée et donc respecter le canon de la logique formelle dont l'essence est d'exprimer ces exigences.

3 Essences différentielles

La phénoménologie husserlienne et sa philosophie de la logique nous ont permis de constater que les ontologies du domaine que nous recherchons pour résoudre le problème que nous avons noté dans le chapitre précédent sont les ontologies matérielles ou ontologies régionales que Husserl a posé pour rendre compte des connaissances synthétiques *a priori* structurant l'expérience possible dans un domaine.

Il s'agit bien d'une réponse à notre problème car la notion d'expérience possible permet de rendre compte qu'il s'agit de tel ou tel domaine, et non d'un autre. En structurant des connaissances qui ne sont pas valables pour tous les domaines mais seulement pour un particulier, il résulte que ces connaissances ne peuvent être analytiques, c'est-à-dire relever de tautologies valables dans tous les domaines ou de contradictions fausses dans tous les domaines, ni être simplement empiriques car précédant l'expérience au lieu d'en résulter.

Les essences différentielles sont des invariants éidétiques obtenus par variation, et sont les catégories structurant l'expérience dans le domaine. Cependant, un point fondamental doit être encore souligné avec de passer au chapitre suivant où sont exploités les résultats de ce chapitre. C'est que, à strictement parler, un domaine est un ensemble de pratiques, unifiées autour d'un objectif commun: par exemple les pratiques médicales autour d'un objectif thérapeutique. En d'autres termes, cela signifie que l'expérience possible constituant le cadre au sein duquel les concets ont nécessairement affaire les uns avec les autres ne coïncide pas avec le domaine, mais avec une pratique ou tâche dans le domaine. C'est ainsi que dans le domaine médical, l'unité de l'expérience possible sera celle du diagnostic, qui ne sera pas la même que celle de la codification médicale des soins. Cela provient du fait qu'un domaine est un système de normes sociales, techniques et culturelles, et que par conséquent il constitue une unité humaine où toutes ces normes cohabitent. Mais, au cas par cas, ces normes ne sont pas forcément compatibles, comme le pouve à l'envi le problème de concilier les normes administratives et comptables avec les normes thérapeutiques.

En conclusion, l'unité de l'expérience possible est pour nous l'unité de la tâche ou pratique dans le domaine, et non l'unité *du* domaine. Ce n'est pas revenir à un point de vue proche de celui de KADS dans la mesure où la tâche n'est pas définie indépendamment du domaine, mais dans le domaine. Au contraire, cela revient à affirmer que l'éboration d'une ontologie ne peut être locale qu'à un domaine donné, et à une tâche donnée dans le domaine. Cela implique que l'ontologie modélisée n'est pas d'emblée réutilisable si l'on change de tâche ou de domaine : un concept ou essence différentielle n'est une connaissance du domaine que dans la mesure où elle renvoie à l'unité d'une expérience possible, à l'unité d'une pratique, unité au sens où tous les concepts participent à la réalisation d'une même activité.

Chapitre IV

Ontologie et artéfacture

Nous avons vu dans le chapitre 2 les problèmes que rencontrent l'IA dans sa figure constructiviste en examinant le problème des ontologies. En cherchant à se constituer comme un ingénierie des connaissances, l'IA dégage des principes méthodologiques pour constituer deux types de composants réutilisables de connaissances : les méthodes génériques de résolution et les ontologies du domaine.

Les méthodes de résolution sont indépendantes du domaine dans la mesure où elles décrivent des manières génériques de raisonner pour résoudre un type donné de problème. On suppose ainsi que le raisonnement peut se définir indépendamment du domaine, et que sa mise en œuvre d'un domaine à l'autre ne change pas sa nature. Nous avons vu que cela n'était pas vérifié dans la mesure où les méthodes n'étaient pas décrites avec suffisamment de précision de manière générique pour que l'on puisse affirmer qu'il s'agisse de la même méthode quand elles sont mises en œuvre dans des domaines distincts. Les méthodes génériques sont par conséquent des abstractions qui permettent d'*interpréter* les connaissances du domaine en vue de construire une méthode de résolution dans le domaine, comme le confirme d'une part l'usage pratique de KADS, et d'autre part le notion de modèle d'*interprétation*.

Par ailleurs, les ontologies du domaine consistent d'une part en ontologies primitives, décrivant quels sont les objets primitifs existant d'un domaine, et d'autre part en ontologies pragmatiques, abstractions d'ontologie primitive, formulées dans les termes adéquats aux raisonnements et méthodes de résolution qui seront appliqués au domaine. Les ontologies pragmatiques fournissent les termes du domaine, le vocabulaire, permettant d'instancier les méthodes génériques. Ce que signifie un terme du domaine consiste dans ce cadre dans les relations formelles stipulées par les méthodes qui l'utilisent. Par lui-même, il n'a pour fonction que de renvoyer à des objets existants ou à des abstractions de ces objets.

Or, nous avons pu constater qu'aucun des travaux cités n'élaborait explicitement des ontologies primitives mais qu'au contraire, ils reposaient tous sur l'hypothèse qu'elles existent toujours dans un domaine donné et que l'on pouvait les supposer acquises. En effet, les travaux procédant de l'ontologie formelle, que l'on pouvait penser voir résoudre ce problème, consistent à poser les axiomes exprimant le contenu des catégories formelles permettant de penser le domaine, comme « espèce » ou « propriété ». Ces axiomes s'interprètent dans l'ontologie primitive et permettent alors d'affirmer si tel ou tel terme (*e.g.* « rouge ») est une espèce ou une propriété. Par conséquent, eux aussi, tout comme les ontologies pragmatiques mobilisées par les méthodes génériques de résolution, reposent sur l'hypothèse de l'existence des ontologies primitives.

Le diagnostic que nous avions formulé est que, mobilisant des outils formels, ces travaux ne peuvent en effet considérer le domaine lui-même, mais seulement « formellement ». Cela signifie que le domaine n'est caractérisé qu'en tant qu'il remplit des catégories formelles, que ce soit celles du raisonnement (ontologies pragmatiques déterminées pour les méthodes de résolution) ou celles de l'ontologie formelle. En d'autres termes, les travaux cités élaborent des contraintes sémantiques qui s'appliquent à tous les domaines, en tant qu'ils sont des domaines, mais non les contraintes propres à tel ou tel domaine particulier. Restant au niveau formel, ces contraintes sont analytiques et *a priori*. Comme la notion husserlienne d'ontologie formelle permet de le comprendre, l'ontologie formelle correspond à la forme de région en général et ne caractérise aucune région en particulier.

Or, l'ontologie primitive d'un domaine n'est pas une donnée du domaine. Autrement dit, un domaine pratique comme la médecine ne possède pas des objets que l'on peut considérer comme des existants primitifs à partir desquels tous les autres objets se dérivent. Les travaux formels supposent l'existence des ontologies primitives alors qu'elles n'existent pas comme données, mais doivent reposer sur une modélisation explcite.

Il est donc nécessaire d'établir une étape méthodologique spécifique pour élaborer à partir de la donnée d'un domaine une ontologie primitive. Pour cela, il est nécessaire de pouvoir caractériser le domaine en lui-même, tel qu'il est lui-même et pas un autre. Comme l'ontologie primitive doit être motivée par les principes du domaine, il faut qu'elle repose sur des principes *a priori* du domaine. Mais ces principes ne peuvent être analytiques, puisque dans ce cas on retrouve comme analytique *a priori* les lois formelles qui ne caractérisent pas le domaine mais la pensée. Il faut donc que ces principes soient synthétiques et *a priori*.

L'objectif de ce chapitre est de montrer comment dégager les principes synthétiques *a priori* du domaine à partir de la notion d'essence différentielle posée dans le chapitre précédent pour être capable de constituer l'ontologie régionale, ontologie matérielle qui à ce titre contient les principes synthétiques *a priori* du domaine permettant de construire l'ontologie primitive supposée par les travaux formalistes. Il est alors possible de les utiliser et de retrouver leurs acquis méthodologiques.

En effet, nous avons pu remarquer dans le chapitre précédent que la modélisation que nous recherchons pour représenter le domaine coïncide avec les ontologies matérielles ou régionales de Husserl, et que les essences qui les composent sont des essences différentielles, et nous avons conclu qu'une ontologie est un système d'essences différentielles (contenu conceptuel) déterminées dans le contexte d'une tâche donnée pour laquelle elles constituent des actions possibles (contenu gnoséologique), c'est-à-dire des connaissances.

Nous avions noté également que le caractère « différentiel » des essences, conféré par la variation éidétique effectuée dans l'imagination, est analogue à une caractérisation différentielle des signifiés telle qu'elle est mise en œuvre en sémantique différentielle. Cela signifierait que les essences différentielles sont différentielles dans la mesure où elles sont extraites des données linguistiques. Il ne faut pas oublier en effet que l'ontologie du domaine s'obtient à partir de l'expression linguistique des connaissances. Par conséquent, une ontologie est un système de connaissances exprimées en langue naturelle. Elle ressortit ainsi au linguistique et hérite de son caractère différentiel. La principale problématique de ce chapitre va résider dans le passage d'une description linguistique des connaissances, depuis la sémantique différentielle, à une description ontologique des connaissances du domaine.

Comme nous l'avons suggéré dans notre introduction, le principe de la modélisation en IA repose sur le fait d'adopter dans les systèmes formels automatiques des signifiants linguistiques en guise de symboles de prédicats, de fonctions, ou de constantes. Cet emprunt à la langue permet de charger de sens le symbole et de voir en lui le représentant de la connaissance qu'il exprime quand il est utilisé dans l'expression linguistique et non dans la formalisation logique. Le problème est donc de contrôler le sens linguistique dont le symbole s'investit de manière à ce que son enrôlement dans le formalisme logique soit conforme au contenu prescriptif de la connaissance dont il doit être le représentant. En d'autres termes, il faut que le signifiant linguistique renvoie à une essence différentielle : pour cela il faut contrôler la manière dont il signifie. Le problème dont s'occupe ce chapitre est de savoir comment expliciter la manière de contraindre l'interprétation linguistique des signifiants symboliques du formalisme ou « labels » des concepts pour qu'ils s'investissent d'un contenu prescriptif et renvoient à une essence différentielle. De cette manière, l'ontologie ou ensemble des termes constituant la signature du formalisme est une ontologie matérielle du domaine, c'est-à-dire notre objectif.

Mais ce travail de contrôle de la signification linguistique pour obtenir des connaissances revient à substituer à l'imagination un corpus linguistique et son exploitation informatique pour la constitution des essences matérielles. Autrement dit, nous n'obtenons pas les essences matérielles de la même manière que Husserl : si cependant le résultat est bien le même, s'il s'agit bien des mêmes essences et des mêmes connaissances synthétiques *a priori*, il conviendra de s'interroger sur le statut du synthétique *a priori* dans notre ontologie régionale et refléchir sur le fait qu'il est obtenu par une médiation technique, technique linguistique d'une part, technique informatique d'autre part. En d'autres termes, le statut « transcendantal » de ces connaissances n'est plus très clair et une nécessaire confrontation entre le transcendantal et le technique doit être entreprise. Ce sera l'objet de la conclusion de ce chapitre d'introduire cette réflexion pour la seconde partie du mémoire.

La description ontologies des connaissances du domaine repose sur l'élaboration de trois systèmes ontologiques, l'ontologie régionale, l'ontologie formelle et l'ontologie computationnelle :

L'ontologie régionale décrit, comme son nom l'indique, l'ontologie d'une *région* du réel. Construite à partir des connaissances du domaine et de leur expression dans la langue naturelle, l'ontologie régionale n'est pas un *catalogue* de noms de classe renvoyant à une extension d'individus pré-donnés mais un *système* d'essences différentielles : *différentielles* car construites à partir du système des différences entre les significations de la langue, *essences* car objectivant les objets du domaine. L'ontologie régionale est une théorie du domaine consistant dans une taxinomie arborescente de concepts et de relations, munis de principes interprétatifs pour les comprendre comme des connaissances du domaine. L'ontologie régionale est construite à partir d'une *normalisation sémantique* de l'expression linguistique des connaissances du domaine pour en extraire leur contenu prescriptif et l'organiser en système d'essences différentielles. C'est la normalisation sémantique qui, constituant les essences différentielles, dégage les principes synthétiques *a priori* du domaine.

L'ontologie formelle consiste dans la *formalisation logique* de l'ontologie régionale en dotant les structures de représentation mises en place par cette dernière d'une sémantique formelle. Elle ne contient plus d'essences différentielles mais des *concepts logiques* qui sont manipulés comme des primitives à partir desquelles construire récursivement une représentation formelle des connaissances du domaine.

L'ontologie computationnelle consiste dans *l'opérationnalisation informatique* de l'ontologie formelle, en la dotant d'une sémantique opérationnelle. Elle ne contient ni des essences différentielles, ni des concepts logiques, mais des structures de programmation permettant d'exploiter les connaissances du domaine dans un programme, en particulier dans un SBC.

Chaque ontologie représente un palier de la modélisation des connaissances d'un domaine nécessaire à l'élaboration d'un SBC. Chaque palier est autonome par rapport aux paliers suivants : l'élaboration d'une ontologie régionale est indépendante de sa formalisation et de son opérationnalisation et constitue une activité théorique autonome. Il existe, autrement dit, plusieurs niveaux d'objectivation théoriques : le niveau régional qui dégage les catégories de l'expérience d'une pratique du domaine, le niveau formel qui dégage une représentation formelle dénotant des objets constitués par l'ontologie régionale, le niveau computationnel objectivant des calculs effectifs.

Il peut paraître singulier de poser comme objectivation théorique un travail de normalisation sémantique qui ne repose ni sur une formalisation (axiomatisation) ni sur une mathématisation. Mais ce niveau renvoie, comme nous le verrons (section 4-9), par la constitution d'ontologie régionale, à une manière de caractériser le programme de recherche proposé à la terminologie, qui en quelque sorte traque les connaissances à partir de leur expression dans la langue.

L'ontologie formelle est indépendante de l'ontologie computationnelle dans la mesure où on la considère dans la perspective d'une sémantique référentielle, c'est-à-dire en termes de structures objectives dénotées dans le domaine. L'ontologie formelle correspond donc au problème de formaliser un domaine pour le soumettre aux lois logiques indépendamment d'une exploitation informatique ultérieure.

Enfin, l'ontologie computationnelle correspond aux structures informatiques effectivement utilisées par un SBC. Elle dépend des choix effectués sur le plan informatique pour la construction du système et s'inscrit donc sur le plan du « comment » plutôt que sur celui du « quoi » de l'exploitation informatique des connaissances pour assister l'automatisation des tâches du domaine. Elle consiste à considérer l'ontologie formelle non comme des structures dénotant le domaine mais comme des structures effectives opérationnelles.

Ces distinctions sont importantes dans la mesure où elles recouvrent des entités non équivalentes et que doivent être étudiées les modifications qu'implique le passage d'un palier ontologique à l'autre.

Le reste du chapitre explore ces trois paliers ontologiques. La nature différentielle des essences provient de leur origine linguistique : elles se dégagent à partir d'une analyse de la langue. La section 4-1 présente les rudiments théoriques de la sémantique différentielle à la base de notre méthode. La nature différentielle de la langue dont sont issues les essences et l'ontologie qu'elles composent détermine la structure globale de l'ontologie qui prend la forme d'une taxinomie arborescente (section 4-2). La structure arborescente permet de formuler des principes « différentiels » qui explicitent et norment la signification des unités linguistiques pour en faire des concepts. Ces principes correspondent à la structure locale de l'ontologie (section 4-3). Les structures globales et locales de l'ontologie permettent de construire l'essence différentielle, c'est-à-dire le contenu prescriptif du concept associé à un label ou signifiant symbolique de l'arbre ontologique (section 4-4). Nous reprenons ensuite (section 4-5) pour la détailler la structure de l'ontologie régionale, puis la structure de l'ontologie formelle (section 4-6) et enfin la structure de l'ontologie computationnelle (section 4-7). Nous soulignons par la suite (section 4-8) le fait que normaliser, formaliser et opérationnaliser sont des opérations qui, rendant les connaissances effectives, gagnant ainsi dans leur détermination calculatoire, les appauvrissent du point de vue de leur signification. Par ailleurs, la normalisation, en dégageant le synthétique *a priori* propre à l'unité d'une pratique dans le domaine, compromet la réutilisabilité de l'ontologie régionale ainsi obtenue (section 4-9). Nous résumons ensuite les étapes de notre méthodologie pour la résolution du problème des ontologies (section 4-10) pour conclure sur les leçons qu'elle nous permet de tirer quant à la nature des connaissances ainsi constituées (section 4-11).

1 Sémantique différentielle

Le fait que les connaissances du domaine ne soient accessibles qu'à travers leur expression linguistique en langue naturelle implique que toute modélisation des connaissances doit commencer par une étude linguistique et sémantique[1] de ces expressions pour construire la théorie du domaine. Mais une étude linguistique ne peut suffire : nous verrons qu'une sémantique linguistique ne permet pas en tant que telle de construire une objectivité, c'est-à-dire de modéliser des connaissances. L'analyse sémantique n'est pas un procédé d'objectivation et ne permet pas d'élaborer en tant que tel un système de connaissance. En effet, le problème est qu'en langue, d'une part, ce qui est considéré comme un objet n'est pas stable, et d'autre part, la manière d'être d'un objet n'est pas stable non plus. La raison est qu'en contexte, là où les expressions linguistiques prennent leur signification et peuvent être interprétées comme véhiculant des connaissances, les conditions pratiques imposées par les normes sociales, techniques, pragmatiques, permettent d'interpréter chaque terme de manière non ambiguë. Mais hors contexte, l'expression linguistique, ne portant pas en elle-même les prescriptions interprétatives permettant de lui associer un sens unique et non ambigu, perd les prescriptions interprétatives du contexte et devient donc équivoque. La pluralité des déterminations en contexte ne se généralise pas en une signification hors contexte univoque. Autrement dit, l'expression linguistique, comme en témoigne sa description sémantique, ne contient pas par elle-même de rapport à une expérience possible, rapport permettant aux unités linguistiques d'exprimer des connaissances.

Il est donc nécessaire, pour retrouver dans l'expression linguistique la connaissance

1. La sémantique, ou théorie de la signification, est une discipline disputée par les philosophes, les linguistes, les sémioticiens, etc. Les travaux sont donc fort nombreux. Dans ce chapitre, nous adoptons le point de vue de François Rastier : (Rastier, 1987) théorise une approche différentielle de la sémantique, (Rastier, 1991) fait le point sur la mobilisation de la sémantique dans les recherches cognitives contemporaines et (Rastier *et al.*, 1994) peut être vu comme un traité systématique de sémantique pour l'analyse automatique de la langue naturelle. Il est d'usage de faire remonter la sémantique au travail fondateur de Saussure avec son *Cours de linguistique générale* (de Saussure, 1985). Par ailleurs, parmi les principaux travaux de langue française, on peut citer (Greimas, 1986; Pottier, 1992; Martin, 1992). Eco (Eco, 1988a; Eco, 1988b; Eco, 1992) est également une référence incontournable. La sémantique est également un sujet activement traité dans les traditions formalistes comme en témoigne (Nef, 1988).

dont elle est l'expression de fixer un contexte d'interprétation permettant de voir en chaque expression une action possible se rapportant à des objets du domaine. En d'autres termes, il va falloir *normaliser* l'analyse linguistique pour qu'à la pluralité du système fonctionnel de la langue et des règles interprétatives socialisées succède l'univocité de l'action mise en œuvre dans la résolution d'un problème donné.

1.1 La sémantique différentielle de F. Rastier

La modélisation des connaissances nécessaires à la résolution d'un problème passe, pour les domaines relevant d'une approche IA, par une analyse des expressions linguistiques des connaissances, puisque dans le cas qui nous intéresse ici, c'est ce type d'expression qui en constitue l'expression canonique et opératoire. Il est donc nécessaire d'effectuer une analyse de la signification de ces expressions pour en extraire le contenu prescriptif correspondant aux connaissances. L'analyse de la signification est une analyse sémantique, et il faut donc disposer d'une théorie sémantique de la langue pour mener à bien l'entreprise de modélisation dont il est question ici.

La sémantique des langues est un sujet largement débattu suscitant fort peu de consensus. Plusieurs théories sémantiques existent et donnent des visions hétérogènes voire contradictoires du sens linguistique. Le paradigme linguistique retenu ici est, comme il l'a été dit plus haut, le paradigme *différentiel*. Avant d'en préciser les principales caractéristiques, il est utile pour le définir de le confronter aux autres paradigmes auxquels il s'oppose.

Le paradigme différentiel s'inscrit dans un cadre théorique où l'on veut étudier le sens linguistique pour lui-même, sans présupposés extralinguistiques : le sens linguistique ne doit pas être compris depuis une réalité extralinguistique prise comme cadre de référence, mais depuis la réalité intralinguistique de la langue. C'est en effet le seul moyen d'obtenir une théorie sémantique autonome.

Par conséquent, une sémantique différentielle refusera de rapporter le sens linguistique à une réalité physique, logique ou psychologique. Elle s'oppose donc à une sémantique *cognitive* qui rapporte le sens aux objets de l'expérience, à une sémantique *formelle* qui rapporte le sens à un concept logique ou prédicat, à une sémantique *psychologique* qui rapporte le sens à des représentations mentales. Cela ne veut pas dire qu'il n'existe pas de rapport entre le sens linguistique et les objets du monde, la cohérence formelle et les représentations mentales. Mais qu'aucune de ces dimensions ne capte le linguistique en tant que tel dans la mesure où elles le déterminent depuis d'instances extralinguistiques : l'objet, le concept, la représentation.

En particulier, la sémantique différentielle va s'opposer aux sémantiques référentielles et inférentielles : la sémantique référentielle définit le sens linguistique comme la référence dénotationnelle à un objet extralinguistique. Elle repose sur l'hypothèse de la compositionnalité du sens dont nous avons vu qu'elle n'était pas vérifiée dans les domaines techniques reposant pour l'expression des connaissances sur la langue naturelle. La sémantique inférentielle définit le sens comme une inférence en général de nature logique entre signes linguistiques ou objets mondains. Si une telle sémantique se montre plus réservée quant à la détermination de la référence, elle reprend à son compte la nature compositionnelle du sens puisqu'elle fonde sur elle ses processus inférentiels.

La sémantique différentielle repose donc sur l'analyse *intralinguistique* du sens et de la signification, et sur le postulat que le sens d'une unité linguistique n'est pas sa référence à un objet, son renvoi inférentiel à un autre signe, mais sa différence avec les autres unités. Le sens linguistique procède dès lors de la position respective des unités linguistiques dans la langue considérée comme système : occupant une position que les autres n'ont pas, elle les nie et en tire ainsi sa signification.

Le sens comme différence définie comme une position dans un système conditionne la nature du *signe* linguistique. Le signe est en linguistique saussurrienne cette entité biface comportant un signifiant, le son, la graphie, etc., et un signifié, ce que veut dire ce son ou cette graphie. Cela signifie que le signe linguistique procède d'une double individuation : d'une part sur le plan des signifiants, il s'individue en un signifiant particulier, d'autre part sur le plan du signifié, il s'individue en un signifié particulier.

Cette opposition entre plan du signifiant et plan du contenu est due à Hjemslev (Hjemslev, 1971) qui a posé qu'un signifiant n'est signifiant que parce qu'il s'oppose à d'autres signifiants possibles, et qu'il en était de même pour le signifié. Ainsi, le /i/ et le /i:/ sont des signifiants identiques en français, alors qu'ils ne le sont pas en anglais. Cela veut dire que l'anglais et le français sont deux langues appliquant des systèmes de signifiants différents, des découpages différents sur la même matière phonique. De même, la matière de tout ce qu'il est possible de dire est découpée différemment selon les cultures. *Holz* ne s'oppose pas à *Wald* de la même manière que *Bois* s'oppose à *Forêt* (Eco, 1988a).

C'est donc la systématicité, le système des différences, qui conditionne la signification d'un signe et sa définition de signe en tant que signe : savoir ce qui dans un texte ou énoncé linguistique constitue un signe ou non n'est pas spécifié d'avance mais résulte d'un parcours interprétatif qui, associant une signification à « Socrate », y voit un signe, alors que « rat » dans « Socrate » n'en constitue pas un. En effet, si l'on substitue à « rat » « souris », on obtient une graphie inintelligible : « rat » dans « Socrate » n'est pas un signe linguistique. Mais il l'est dans « un chat mange le rat » car cela s'oppose à « un chat mange la souris », énoncés tous deux pourvus de sens. Comme le souligne F. Rastier :

> « Les signes linguistiques ne sont que le support de l'interprétation, ils n'en sont pas l'objet. Seuls les signifiants, sons ou caractères, sont transmis : tout le reste est à reconstruire. En d'autres termes, l'interprétation ne s'appuie pas sur des signes déjà données, elle reconstitue les signes en identifiant leurs signifiants et en les associant à des signifiés. L'identification des signes comme tels *résulte* donc de parcours interprétatifs. » (Rastier *et al.*, 1994, p. 12)

Un signe est donc un signe en tant qu'il possède une valeur différentielle, qu'il s'oppose à un signifiant que l'on aurait pu mettre a sa place en modifiant le sens de manière à en susciter un autre. Le *morphème* constitue ainsi le signe linguistique minimal : son signifiant ne contient aucune partie dotée d'un signifié. Par exemple, « rétro » est un morphème, mais « tro » n'en est pas un. Il s'agit de la première articulation du discours, celle des énoncés en morphèmes. Mais il existe une seconde articulation du discours [2], celle des morphèmes en *phonèmes*, qui sont les plus petites unités signifiantes qui, bien que dépourvues de contenu, modifie les morphèmes qui les contiennent dans leur face signifiante. Par exemple, en français, substituer le son /i:/ (le « i » long) au /i/ (« i » court) ne modifie pas le sens des morphèmes qui les contiennent : ces sons constitue un même phonème. En revanche,

2. Voir par exemple (Martinet, 1980) pour des présentations plus détaillés sur le plan linguistique de la double articulation.

en anglais si l'on substitue un /i:/ à /i/ dans *ship* on obtient un *sheep*. Ce sont donc deux phonèmes différents en anglais.

En résumé, la sémantique différentielle considère la langue comme un système de différence où chaque position dans le système définit une valeur linguistique, une réalité linguistique. La définition des unités linguistiques en termes de valeur mobilise par conséquent trois principes :

- la valeur est la véritable réalité des unités linguistiques ;
- elle est déterminée par la position des unités dans le système des signifiés, c'est-à-dire par les différences ;
- rien ne préexiste à la détermination de la valeur par le système. La valeur, relation entre signifiés, exclut toute définition du signe qui le rapporterait à des entités extra-linguistiques : résultat du jeu systématique des signifiés, la valeur procède du global vers le local, interdisant toute compositionnalité au sens vériconditionnel.

Il s'agit dès lors de savoir caractériser en quoi consistent les différences constitutives du sens et de définir d'une part les systématicités dont relèvent les différences et d'autre part leur modes d'organisation.

F. Rastier distingue (Rastier, 1987; Rastier *et al.*, 1994) trois types de systématicité gouvernant l'économie des différences constitutives du sens :

le système fonctionnel de la langue prescrit les différences imposées par la langue elle-même ; c'est ce système qui définit l'inventaire des morphèmes et de leurs combinaisons ;

les normes sociales déterminent les différences de sens qui valent dans une société donnée ; c'est typiquement de ce niveau de systématicité que relève l'étude linguistique des connaissances linguistiques d'une spécialité ;

les normes idiolectales déterminent le système des différences valables pour un individu ou une société réduite d'individu ; c'est typiquement à ce niveau que se situe le vocabulaire intime, familial ou câlin.

Par ailleurs, pour expliciter les différences à l'œuvre dans la détermination d'un sens linguistique, il convient d'adopter une approche *onomasiologique*. Cette dernière s'oppose à la méthode *sémasiologique* en usage dans les milieux cognitifs :

> « La méthode traditionnelle, dite *sémasiologique*, est employée par la plupart des dictionnaires. Elle code les unités lexicales par leur signifiant graphique. A chaque signifiant, on associe l'ensemble des signifiés que ce signifiant est susceptible de véhiculer, et l'on cherche ensuite à structurer leur inventaire. » (Rastier *et al.*, 1994, p. 45)

Il est alors impossible de décrire les systèmes de différences constituant le sens dans la mesure où les signifiés sont confrontés non pas par leur opposition dans le système de la langue mais par leur association commune à un même signifiant. Ainsi, on s'interrogera sur le problème profond de savoir ce qui peut bien y avoir de commun entre un célibataire et un jeune phoque, sous prétexte qu'ils constituent tous deux des signifiés possibles de *bachelor*. Pourtant, à part dans les études de linguistique d'obédience cognitiviste, il y a fort à parier que l'on ne trouvera jamais dans un même texte les célibataires et les jeunes phoques. Ou, si cela a lieu, c'est pour opposer les jeunes phoques, comme célibataires, aux phoques plus âgés connaissant les rigueurs de la vie maritale. Par conséquent, il paraît plus judicieux de comprendre « célibataire » par opposition à « homme marié » que par rapport à « jeune phoque ». Ou encore, comme le précise Rastier dans (Rastier, 1987, p.48), « on conviendra que pour définir les deux sémèmes manifestés par *chinois*, il demeure plus rationnel et plus économique de distinguer « passoire » des autres ustensiles de cuisine, et « citoyen de la R.P.C. » des autres citoyens, plutôt que d'interdéfinir le citoyen et la passoire... ». C'est précisément ce que permet la méthode onomasiologique :

> « L'autre méthode, dite *onomasiologique*, part des classes de signifiés pour les structurer. Les dictionnaires thématiques ou notionnels en adoptent certains principes. [...] L'essentiel réside dans le caractère différentiel de la méthode : le sens d'un mot se définit non par rapport à ses autres sens, mais par rapport au sens des mots voisins, aussi bien dans l'ordre paradigmatique que dans l'ordre syntagmatique. Dans le premier ordre, un mot se définit d'abord par rapport à ceux qui lui sont le plus proches, c'est-à-dire ceux qui lui sont opposés. Par exemple *frère* se définit par rapport à *sœur*, *naître* par rapport à *mourir*, etc. » (Rastier *et al.*, 1994, p. 45-46)

Il s'agit donc, dans une perspective onomasiologique de déterminer quels sont les traits distinctifs entre signes linguistiques. Ces traits distinctifs sont des traits différentiels, des traits d'opposition, dont on attend qu'ils permettent d'analyser complètement le sens du signe :

> « Dire qu'un signifié linguistique est relatif à une langue définie comme système, c'est dire qu'il s'analyse exhaustivement en relations d'opposition. Ce sont autant de traits relationnels, qui différencient sa classe des autres classes, ou qui le différencient au sein de sa classe. Les traits qui le composent sont dénommés par des paraphrases intralinguistiques. Ce sont des éléments de définition (non des descriptions de l'« objet » dénoté). Ces traits sémantiques relatifs à des langues ne doivent pas être confondus avec d'autres types de traits définis par des sémantiques non différentielles : nommément le traits référentiels, qui dans certaines sémantiques de la dénotation, sont autant de conditions nécessaires et suffisantes pour apparier une expression et un objet ; et les primitives, qui dans nombre de sémantiques structurales ou cognitives sont autant d'atomes conceptuels indépendants des langues. » (Rastier *et al.*, 1994, p. 33)

Ces traits distinctifs s'appellent des *sèmes* : ils permettent d'opposer les signifiés des signes linguistiques. Si on appelle *sémème* le signifié d'un morphème, un sème peut donc se définir comme un « élément d'un sémème, défini comme l'extrémité d'une relation fonctionnelle binaire entre sémèmes ». Par exemple, le sémème /bistouri/ possède le sème « pour les vivants » et constitue l'extrémité d'une relation fonctionnelle d'opposition avec le sémème /scalpel/ qui possède comme autre extrémité de la relation le sème « pour les morts ».

Dans la description des traits distinctifs d'un signifié, il convient de faire les distinctions suivantes. On oppose les *sèmes génériques* aux *sèmes spécifiques*. Les premiers sont hérités des classes hiérarchiquement supérieures et indexent le sémème dans ces classes. Par exemple, le sème /animal/ est générique dans le sémème /chien/. Ils notent les relations d'équivalence entre sémèmes. Les seconds différencient les sémèmes dans le contexte de lexies appartenant à une même classe. Ils notent des relations d'opposition entre sémèmes.

Par ailleurs, ces sèmes génériques et spécifiques peuvent être *inhérents* ou *afférents*. Les sèmes inhérents sont définitoires du type et sont hérités par défaut du type à l'occurrence. Par exemple, /noir/ est un sème inhérent au sémème /merle/ : chaque occurrence du syntagme « merle » en hérite par défaut. Mais un tel héritage n'est jamais obligatoire et peut toujours être inhibé : ainsi Musset écrivit-il « la merlette blanche ». Les sèmes afférents sont quant à eux de deux types :

> « Les premiers [premier type de sèmes afférents] notent des relations applicatives d'une classe minimale de sémèmes (*taxème*) ou de sémies (*taxémie*) dans une autre. [...] Cette sorte d'application rend compte des phénomènes dits de connotation, ainsi que de phénomènes de prototypicalité. Ces relations applicatives dépendent de normes sociales différentes du système de la langue (mais cependant à l'œuvre dans tout texte). D'où sans doute le caractère « périphérique » souvent attribué aux sèmes afférents. On peut dire que ces sèmes afférents, *socialement normés*, sont associés au type sans avoir de caractère définitoire au même titre que les sèmes inhérents. A ce titre, ils ne sont actualisés dans l'occurrence qu'en raison de prescriptions issues du contexte. Ces sèmes sont des valeurs prises dans l'occurrence par des attributs facultatifs du type. Ils ne sont pas hérités par défaut, mais doivent être actualisés par une instruction contextuelle. » (Rastier *et al.*, 1994, 53-54)

Il faut leur distinguer par ailleurs une seconde sorte de sèmes afférents :

> « Une seconde sorte de sèmes afférents ne dépend pas de relations paradigmatiques entre classes, mais résulte uniquement de propagations de sèmes en contexte. On peut les appeler *sèmes afférents contextuels*. Leur mode d'actualisation les distingue radicalement des précédentes sortes de sèmes, car il ne met pas en jeu le rapport entre types et occurrences, mais uniquement les rapports entre occurrences. Traditionnellement, les sèmes afférents contextuels ne sont pas représentés dans le type lexical. Ils sont propagés dans l'occurrence par le contexte, notamment au moyen de déterminations ou de prédications. » (Rastier *et al.*, 1994, 54)

Répertorier les différents sèmes qui peuvent être actualisés dans un sémème occurrence constitue donc la définition sémantique du sémème type. Mais il faut remarquer qu'empiriquement seuls les sémèmes occurrences existent linguistiquement, et que les sémèmes types sont des artefacts du linguiste. Cela signifie en particulier qu'il est impossible de définir *a priori* quels seront les sèmes présents dans un sémème occurrence. En effet, de multiples opérations interprétatives commandées par la cohésion du sens viennent modifier les traits distinctifs de l'occurrence qui dès lors n'est plus la simple réplique du type. Ces phénomènes sémantiques incitent à remettre en cause la relation type / occurrence unissant le système de la langue à ses réalisations. En effet, la compréhension que se forge le linguiste de la langue en l'objectivant en un système de types sémantiques caractérisés par des traits différentiels ou sèmes est un artefact de la modélisation qui, pris pour la réalité ontologique du langage, entraîne la conception selon laquelle toute mise en discours effectif ne fait qu'altérer le type sémantique et l'adapter aux nécessités de la parole. La réalité du langage réside dans ce cas dans le système de la langue et non dans ses réalisations. La sémantique différentielle bouleverse cette vue des choses dans la mesure où la réalité empirique et attestée du langage n'est pas le système mais ses réalisations discursives dans le texte ou la parole. Par conséquent, c'est l'occurrence qui construit et constitue le type, et non le type qui se dévoie en occurrences :

> « En somme, la hiérarchie entre sens et signification doit être inversée. Le sens n'est pas de la signification déformée par le contexte : la signification est du sens appauvri car coupé de son contexte. Le type est une collection d'accidents, un résumé conventionnel des occurrences retenues comme pertinentes pour sa définition. La sémantique différentielle ébranle ainsi la conception classique en inversant le rapport entre signification et sens. La signification ne serait plus un type diversement déformé dans ses occurrences qui constituent les sens. Elle apparaît alors comme un artefact de l'ontologie, appuyée sur la lexicographie, alors même que les sens varient sans limite, confirmant l'intuition que toute occurrence est unique. » (Rastier *et al.*, 1994, p.35) '

Trois opérations fondamentales permettent de déterminer le sens des lexies, transformant ainsi les significations répertoriées par les types en sens contextuel :

l'inhibition interdit l'actualisation de sèmes inhérents ; ils sont dits alors *virtualisés* ;

l'activation permet l'actualisation des sèmes ; elle ne concernent que les sèmes afférents dans la mesure où les sèmes inhérents sont toujours présents dans le sémème occurrence car, à moins d'être virtualisés, ils sont hérités par défaut du sémème type ;

la propagation intéresse les sèmes afférents en contexte ; principalement véhiculée par la qualification et la prédication, la propagation active des sèmes qui n'apparaissent pas dans la sémie type, fût-ce à titre d'afférence, mais qui sont propagés par les occurrences voisines dans le contexte.

Ces opérations fondamentales mettent en œuvre deux lois fondamentales régissant les rapports entre sèmes, l'assimilation et la dissimilation. La dissimilation opère quand les contrastes sémantiques sont faibles, comme c'est notamment le cas pour les tautologies : dans « une femme est une femme », on affecte une acception différente aux deux occurrences de « femme ». En d'autres termes, la dissimilation est la loi sémantique correspondant au fait qu'on ne parle jamais pour ne rien dire, et que toute répétition se comprend sur un fond de différence.

L'assimilation opère à l'inverse de la dissimilation et atténue les contrastes sémantiques trop forts. Elle consistera à virtualiser les sèmes contradictoires pour activer les sèmes présents dans les sémèmes de manière à constituer la cohérence de l'énoncé. L'assimilation est la loi sémantique correspondant au fait qu'on parle toujours pour dire quelque chose, et que l'interprétation est un travail de construction du sens en annihilant les incohérences.

Les opérations interprétatives permettent de construire le sens d'un énoncé en annulant les incohérences et activant les redondances. Le sens d'un énoncé se définit donc par la redondance des sèmes entre les sémies occurrences : c'est la notion d'isotopie. L'isotopie est tout simplement la présence de sèmes identiques dans les différentes sémies occurrences constituant l'énoncé analysé. Cette notion (voulant dire littéralement « le même sujet ») signifie que toute compréhension se construit sur la base d'une *présomption d'isotopie*. Cette présomption se construit en tenant compte d'une part des indications de la syntaxe qui spécifient quelles sémies doivent participer à la construction d'un sens, d'autre part des problèmes rencontrés lors de la considération des sémies ainsi prescrites, et enfin des relations sémiques prescrites par la langue, les sociolectes, idiolectes d'une part, et le contexte d'autre part permettant de résoudre ces problèmes en faisant jouer les opérations interprétatives vues ci-dessus [3].

Or, la construction d'une isotopie en suivant les prescriptions du lexique quant aux sémies types et du contexte quant aux sémies occurrences font de la compréhension une opération éminemment contextuelle, descriptible *a posteriori* et non *prescriptive a priori*. En effet, c'est le genre textuel qui va prescrire des opérations interprétatives qui se répercutent sur le texte lui-même, possédant ses propres contraintes en tant que texte qui vont s'exercer sur périodes et syntagmes qu'il contient, pour terminer sur les sémies occurrences. Il y a donc une cascade de mises en contexte qui, si on peut les répertorier et classifier, ne peuvent donner lieu à une théorie prescriptive et constructive du sens.

C'est la raison pour laquelle une analyse sémantique ne permet pas de dégager les connaissances linguistiques à partir du corpus linguistiques qui les expriment : car si la théorie sémantique permet bien de décrire le contenu, elle n'en dégage pas l'idéalité en tant qu'action *possible* et donc elle permet de décrire le contenu *exprimé* ou *signifié*, mais non le contenu *prescriptif* des connaissances.

L'analyse sémantique, étant résolument intralinguistique, adopte pour méthodologie un nominalisme radical. Se privant, pour constituer son objet, de toute référence extralinguistique, elle se prive par là-même de la possibilité de comprendre le rapport au monde exprimé par le linguistique. En d'autres termes, le linguistique, en se constituant comme tel, perd toute valeur gnoséologique. Il faudra donc, pour dégager la valeur prescriptive du donné linguistique introduire une perspective gnoséologique, c'est-à-dire un rapport au monde qui donnera le point de vue nécessaire pour dominer l'imprévisibilité de la construction de la cohésion sémantique.

En voulant caractériser le savoir véhiculé par les connaissances exprimées en langue, la construction ontologique part du contenu linguistique de leur expression linguistique, mais ne peut s'en tenir là : elle doit le déterminer en un contenu conceptuel qui n'est un contenu gnoséologique, c'est-à-dire une véritable connaissance comme rapport au monde, que si cette détermination s'inscrit dans l'horizon d'une expérience possible. En effet, le problème de la connaissance n'est pas celui d'une linguistique descriptive comme l'est la sémantique différentielle :

> « Leur valeur de connaissance dépend exclusivement de la pratique dont procède leur production et leur lectuer, et la linguistique n'a pas à en traiter. Le problème linguistique du sens est alors soigneusement distingué du problème philosophique de la connaissance. » (Rastier *et al.*, 1994, p. 14)

Les exemples qui suivent vont illustrer la nécessité d'introduire le point de vue de l'expérience possible ou de la tâche dans l'analyse sémantique pour que, normalisant le contenu signifié depuis le point de vue de la tâche, on puisse en dégager la valeur de connaissance pour la tâche.

3. Ces trois points sont respectivement appelés par Rastier *la condition d'accueil*, *le problème interprétatif* et l'*interprétant*.

1.2 Sens linguistique et sens prescriptif

La signification linguistique ne peut constituer une théorie du domaine : un signifié n'est pas un concept dans la mesure où il n'en possède pas les caractéristiques essentielles. Un signifié est une entité *descriptive* alors que le concept est de nature *prescriptive*. Le signifié rend compte de la signification d'un morphème, décrit ses rapports sémantiques aux autres morphèmes du discours, mais il ne détermine pas selon des règles quels rapports sont obligatoires et lesquels sont interdits. En revanche, un concept prescrit les sens des termes en déterminant le contenu sémantique qu'ils doivent véhiculer et les relations conceptuelles qu'ils entretiennent avec les autres termes du discours.

Par exemple, soit l'énoncé « boire comme un trou ». Cet énoncé possède un sens décrit linguistiquement : /boire/ est une sémie possédant comme sème inhérent le trait /ergatif/. Par propagation sémique et présomption d'isotopie, ce trait s'actualise dans la sémie /trou/, pour lequel il est un sème afférent, ainsi que le trait /animé/, qui n'est pas afférent à /trou/. Ce dernier implique qu'il y a virtualisation du sème /inanimé/ qui lui est inhérent.

Considéré comme exprimant une connaissance, cet énoncé viole les contenus prescriptifs attachés aux concepts qu'il contient. En effet, seul un être *animé* peut boire.

Par conséquent, un énoncé dont le sens est décrit linguistiquement peut ne pas avoir de sens prescriptif, de contenu prescriptif. Cela signifie en d'autres termes qu'il ne véhicule pas de connaissances en tant que telles, mais seulement un sens linguistique. La différence provient du fait qu'un concept est prescriptif dans la mesure où il renvoie à un contenu réel, un contenu objectif, *i.e.* les objets du domaines : il possède un contenu conceptuel et un contenu gnoséologique. Respectant les contraintes de cohérence pour avoir un contenu conceptuel, et la réalité du domaine pour avoir un contenu gnoséologique, le concept doit refléter les contraintes objectives réelles du domaine. En d'autres termes, un énoncé possédant un contenu prescriptif, c'est-à-dire exprimant une connaissance, peut commander l'action : l'énoncé « la sténose est serrée » prescrit, dans le contexte des coronarographies, par son contenu gnoséologique qu'il faut dilater l'artère pour restaurer la circulation. Le contenu linguistique ne prescrit aucune action à entreprendre : décrivant le sens, il ne prétend pas renvoyer au monde objectif, et donc d'y prescrire une action à tenir. Le sens linguistique ne décrit pas le concept comme une action possible dans le domaine, et c'est pourquoi il est incapable d'en rendre compte en tant que connaissance[4].

Par exemple, un énoncé linguistique peut être « la sténose du patient ». Cet énoncé possède un sens linguistique, mais ne véhicule pas de connaissance – ne possède pas de contenu prescriptif – en tant que tel. En effet, une sténose est l'état d'un segment d'artère possédant un diamètre inférieur à la normale et ne peut donc être rencontrée que sur une artère. La connaissance que l'on peut néanmoins associer à cet énoncé est qu'il existe une artère du patient sténosée sur l'un de ses segments. L'énoncé linguistique possédant, outre un sens linguistique, mais aussi un contenu prescriptif est donc l'énoncé suivant : « sténose d'un segment d'une artère du patient ».

Par ailleurs, un énoncé semblant être pourvu d'un contenu conceptuel peut ne pas véhiculer la même connaissance selon les locuteurs. Par exemple, l'énoncé « le sectral a

4. Par sens linguistique, nous comprenons le sens tel qu'il est rendu dans une description linguistique du type de la sémantique différentielle.

changé de principe actif mais pas de classe thérapeutique » semble respecter les prescriptions conceptuelles des termes utilisés. Cependant, cela signifie pour un pharmacologue que le médicament « sectral » a désormais une nouvelle nature : il y a deux objets différents du domaine. En revanche, pour un clinicien par exemple, il y a toujours un même et seul objet dont une propriété a varié, le principe actif. Par conséquent, l'énoncé linguistique est ambigu dans la mesure où il ne possède pas suffisamment de prescriptions interprétatives permettant de déterminer le contenu conceptuel de l'énoncé. On rencontre le même problème à propos de terme pouvant évoquer des concepts de différente nature quoique apparentés : par exemple, le terme « maladie » peut évoquer le contenu conceptuel de « état morbide » ou de « processus pathologique ». Bien que fortement corrélées, ces deux notions sont cependant distinctes, et choisir l'une plutôt que l'autre modifie le type des inférences et raisonnements que l'on peut mener à leur endroit. Cela correspond donc à deux problèmes, le premier étant qu'un objet du domaine n'est pas stable dans son existence, il peut disparaître en tant que notion (cf. sectral), et le second étant qu'un objet n'est pas stable dans son mode d'être (cf. maladie).

Pour déterminer quelles connaissances sont exprimées par des énoncés linguistiques, il est donc nécessaire de leur associer des prescriptions interprétatives permettant de déterminer le contenu conceptuel qu'ils véhiculent, c'est-à-dire les connaissances qu'ils expriment. Il faut donc normaliser la signification des expressions pour qu'il soit possible de leur associer un contenu conceptuel même hors contexte. La normalisation repose sur le constat que le signifié linguistique n'est pas conceptuel et qu'il faut comprendre le concept comme un *signifié normé* (Rastier *et al.*, 1994).

1.3 Normalisation et contenu conceptuel

Ce que montre la sémantique différentielle, c'est qu'utiliser des signifiants linguistiques pour qu'un formalisme les mobilisant comme sa signature représente par son comportement inférentiel les connaissances que ces signifiants véhiculent linguistiquement ne suffit pas. En effet, la sémantique différentielle est descriptive dans la mesure où, tant qu'on s'en tient au sens linguistique, on ne peut que décrire le sens constitué par les signifiants sans être capable de prescrire leur signification.

Or, ce que l'on veut, c'est obtenir une ontologie constituée de signifiants linguistiques dont la signification soit prescrite et contrainte pour qu'ils renvoient à des connaissances du domaine. Il faut donc introduire une dimension supplémentaire, qui par nature n'est pas du ressort de la description linguistique, pour passer du descriptif au prescriptif. Cette dimension supplémentaire est l'expérience possible ou tâche qui sera confiée au système. La normalisation doit s'effectuer en déterminant l'ambiguïté linguistique en une univocité gnoséologique.

De manière générale, la normalisation doit refléter un point de vue pratique et un point de vue pragmatique. Le point de vue pratique correspond au point de vue de la tâche à accomplir dans le domaine. Par exemple, ce sera une tâche de diagnostic ou de codification médicale : cela signifie que c'est dans ce contexte de diagnostic ou de codification que les signifiants prennent leur valeur de connaissances, c'est-à-dire qu'ils s'interprètent comme des connaissances et non seulement comme des significations possibles dans le domaine. Etant donné qu'un signifiant linguistique ne recevra pas la même valeur de connaissance dans le contexte d'une tâche diagnostic ou de codification, la normalisation doit fixer le contexte de la tâche.

Le point de vue pragmatique consiste dans le choix de conditions d'expression, déterminant leur contexte, permettant de fixer les significations. Par exemple, le choix d'un type de discours, le compte rendu médical, d'une spécialité, la coronarographie, et d'une fonction du discours, le compte rendu de sortie, fournit des instructions interprétatives pour la compréhension des expressions. Cela signifie qu'un signifiant ne reçoit pas toujours la même valeur de connaissance selon la manière dont il est utilisé : la normalisation doit trancher également à ce niveau.

La prise en compte des conditions pratiques et pragmatiques permet de considérer qu'une expression linguistique soit l'expression d'une connaissance et reconnue comme telle. Elles permettent de stabiliser l'objet dans son existence et dans sa manière d'être en fixant d'une part son contenu *noématique* et d'autre part son contenu *relationnel* : de déterminer de quel type est un objet (contenu noématique) et quelles relations il entretient avec les autres notions (contenu relationnel).

On peut se poser la question de savoir si une telle normalisation sémantique, qui détermine l'ambiguïté linguistique en fonction du contexte de la tâche pour reconnaître dans l'expression linguistique l'expression d'une connaissance, définie bien les essences matérielles ou différentielles rencontrées chez Husserl. Tout d'abord, on peut se poser la question de savoir si elles possèdent le même statut épistémologique, à savoir si elles sont synthétiques *a priori*. La réponse est oui dans la mesure où elles structurent l'expérience du domaine : *a priori* car *conditionnant* l'expérience, synthétique car *valant pour* l'expérience. D'autre part, on peut se demander si les essences différentielles résultant d'une normalisation sémantique ont le même contenu que les essences matérielles husserliennes. La question est plus délicate mais il nous semble que la réponse doit être également positive. Nous avions souligné dans le chapitre précédent comment les essences husserliennes pouvaient être considérées comme différentielles. Mais, les essences husserliennes ne sont pas linguistiques : reposant sur des variations imaginatives, ce sont des contenus pensés et non des contenus linguistiques. Mais, dans le cadre sémantique qui est le nôtre, les variations linguistiques sont sur le plan du contenu. Le seul obstacle, irréductible celui-là, réside en conclusion sur le fait que Husserl s'appuie sur l'esprit et l'imagination pour dégager l'essence et en avoir l'intuition, alors que nous nous appuyons sur la médiation du corpus et de la technique linguistique pour obtenir un semblable résultat. Par conséquent, si l'essence différentielle semble bien avoir le même contenu et le même statut épistémologique que l'essence husserlienne, il n'en demeure pas moins qu'elles diffèrent sur ce point essentiel. Nous y revenons à la fin du chapitre.

Le principe sur lequel s'appuie la normalisation consiste dans le fait que la description linguistique sémantique délivre une représentation des expressions en langue naturelle des connaissances sous la forme d'un réseau d'identités et de différences entre les unités linguistiques, comme on l'a vu. Il s'agit alors, face au contexte de la tâche visée (la condition pratique) et aux pratiques interprétatives du domaine (condition pragmatique) d'expliciter le sens que *doivent* prendre les unités linguistiques pour être des connaissances. Il faut fixer quel jeu d'identités et de différences permet d'y reconnaître des connaissances, c'est-à-dire expliciter et fixer ces identités et différences en prescriptions interprétatives qui donnent aux unités linguistiques leur contenu prescriptif, leur associant ainsi des signifiés normés ou concepts.

On a vu que la langue consiste en un système linguistique d'identités et différences. Il faut obtenir un système conceptuel d'identités et de différences : le système linguistique n'a pas la systématicité requise pour fonctionner comme un système conceptuel ou système de connaissances. Dans le système conceptuel, le contenu prescriptif d'un concept est déterminée par la position qu'occupe son signifiant, c'est-à-dire le nœud qui le représente.

Mais, en normalisant le linguistique en fixant un jeu particulier d'identités et de différences pour lui faire jouer un rôle de prescriptions interprétatives et plus seulement un rôle de descriptions sémantiques, on reste dans le cadre différentiel : un concept prend son contenu prescriptif par les traits qui le différencient de ses voisins. Ce cadre différentiel possède de nombreuses conséquences quant à la structure du système conceptuel et prescriptif qu'est l'ontologie. Nous distinguons deux niveaux structurels pour expliciter le caractère différentiel de l'ontologie :

un niveau global où le caractère différentiel impose une structure taxinomique et arborescente à l'ontologie ;

un niveau local où le caractère différentiel détermine comment expliciter localement le contenu prescriptif d'un nœud pour qu'il puisse fonctionner comme un concept ;

L'explicitation locale permet de donner le contenu local du concept, c'est-à-dire les traits différentiels vis-à-vis des plus proches voisins. La structure globale arborescente permet alors de passer du local au global, c'est-à-dire d'expliciter le contenu global du concept, c'est-à-dire la signification que possède un nœud compte tenu non seulement de ses plus proches voisins mais aussi de l'ensemble des autres nœuds. En effet, la structure arborescente permet en quelque sorte de « sommer » les contenus locaux pour retrouver les contenus globaux.

Dans ce qui suit, nous explicitons la structure globale (section 4-2) et la structure locale (section 4-3) de l'ontologie régionale.

2 Structure globale de l'ontologie

Nous voyons dans cette section la struture d'une ontologie régionale étant donné qu'elle contient des essences différentielles. Nous voyons d'une part qu'il s'agit d'une structure taxinomique (section 4-2.1) et d'autre part qu'elle est arborescente (section 4-2.2).

2.1 Ontologie différentielle et subsomption

La structure d'une ontologie régionale est en effet une structure taxinomique, c'est-à-dire une structure classificatoire des essences selon la relation générique « est-un ».

En tant que système d'*essences*, l'ontologie détermine ce qu'*est* une chose. Mais, en tant que *système* d'essences, l'ontologie détermine ce qu'est une chose par tout ce qu'elle n'est pas. En effet, par système, il faut comprendre un réseau dans lequel la position de l'essence définit sa signification par le jeu des différences qui l'opposent aux autres nœuds du réseau. Dans un réseau, un nœud ne possède pas de signification par lui-même, mais par le fait qu'il occupe une position et pas les autres : la signification découle de la situation topologique. Le nœud ne possède pas d'individualité, ne se distingue pas des autres nœuds, sinon par sa position ; il n'est pas possible d'importer d'autres considérations que sa position pour déterminer sa signification.

Il est donc nécessaire que la position détermine ce qu'*est* une chose par ce qu'elle *n'est pas*. Or, les différences ne se définissent pas dans l'absolu, mais toujours sur un fond identitaire dont elles relèvent les saillances différentielles. C'est ainsi que l'essence x sera définie comme étant différente de l'essence y en disant quelque chose du genre : x, c'est comme y sauf telle ou telle différence notable. Par conséquent, le système des différences constituant l'ontologie sera déterminé sur la base d'un système d'identités à partir desquelles démarquer les essences en faisant ressortir leurs différences. Cela signifie donc que le système de différences se structure le long de relations taxinomiques de type « est-un ». Ce sont les relations « est-un » établies entre les essences qui permettent de déterminer le fond identitaire à partir duquel expliciter leurs différences et donc leur contenu.

Il ne suffit pas en effet d'affirmer que le sens est différentiel pour qu'il soit dès lors possible d'assigner un sens aux essences. Il faut en outre se donner les moyens de déterminer quelles sont les différences constituant le contenu conceptuel et gnoséologique des essences. En affirmant que le système des essences est une hiérarchie taxinomique, on affirme que les essences sont toutes reliées les unes aux autres par l'intermédiaire de liens taxinomiques. Par exemple, sur l'arbre de la figure 4.1, les essences `physical_object` et `intentional_object` ont leur contenu mutuellement déterminé par le fait qu'ils sont tous deux des `entity` à la différence près qu'un objet physique se définit par sa position dans l'espace et non dans le temps, au contraire de l'objet intentionnel qui, selon les principes de l'ontologie de MENELAS, se détermine dans le temps et non dans l'espace. C'est grâce au lien taxinomique « est-un » qu'il est possible de savoir que la différence pertinente dans l'ontologie entre ces deux essences, indépendamment des associations sémantiques qu'on peut effectuer par ailleurs, consiste dans les traits différentiels que l'on a établis pour distinguer les objets intentionnels des objets physiques. Le lien taxinomique est véritablement le fond sur lequel se détachent les différences.

Si l'on a su poser les nœuds fils d'un père et frère des fils déjà existants en explicitant les différences avec le père et ces fils, alors il est possible, grâce au lien taxinomique, de déterminer *quelles sont les différences entre des essences* et de leur donner ainsi un contenu. Le lien taxinomique est par conséquent la condition nécessaire mais non suffisante (il existe des taxinomies non différentielles) à l'établissement d'un système de différences. C'est donc le fait d'être différentielles qui imposent que les essences soient organisées en réseau de type taxinomique. Chaque nœud se retrouve donc déterminé par un double jeu d'identités et de différences :

- le nœud est affirmé être identique à tous les nœuds auxquels il est relié par le lien « est-un » ; étant affirmé identique, il en hérite les propriétés ;
- le nœud est affirmé être différent de tous les nœuds occupant une autre position que la sienne dans le réseau, *i.e.* tous les autres nœuds.

C'est en particulier le ou les chemins reliant, par l'intermédiaire des liens taxinomiques, deux essences, qui permettent de déterminer leurs différences et par conséquent leur contenu. Ces liens permettent d'expliciter que ces deux essences sont identiques à la

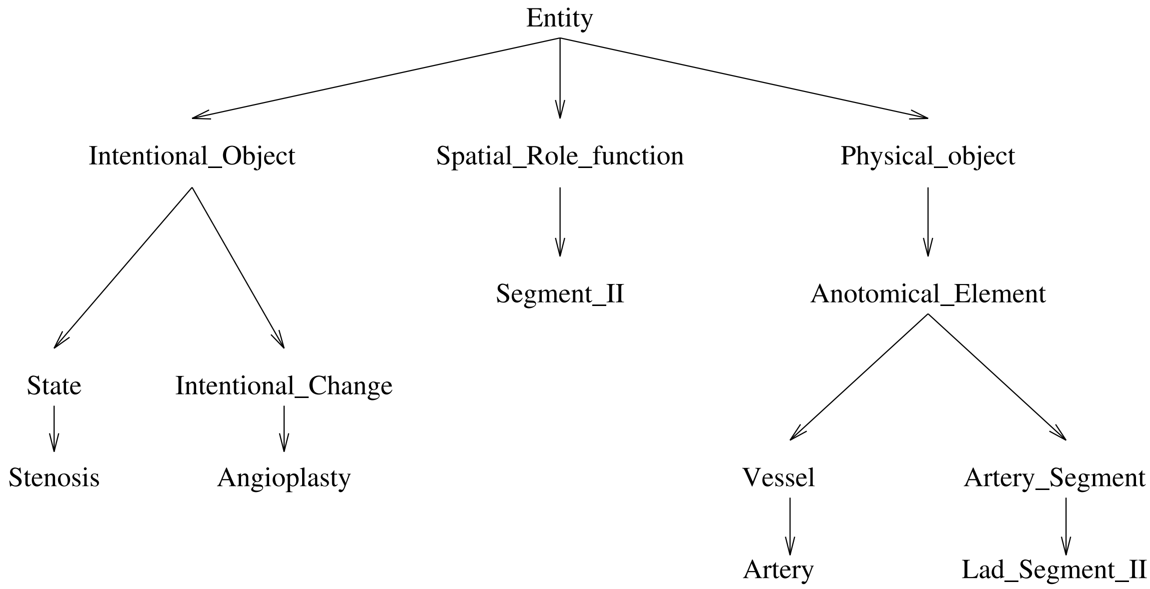


FIG. 4.1 – *Un extrait de l'ontologie du domaine construite dans* MENELAS

différence près des distinctions effectuées à chaque fois que l'on a établi un nœud différent sur le chemin qui les relie. On en tire la propriété suivante :

> le contenu d'une essence se détermine comme l'ensemble des différences que l'on obtient en parcourant tous les chemins taxinomiques qui mènent de cette essence à toutes les autres.

Précisant dans les termes d'une taxinomie hiérarchique le contenu d'une essence différentielle, cette propriété conditionne la nature et les structures de l'ontologie et de sa représentation taxinomique. En particulier, elle prescrit les principes méthodologiques permettant de la constituer (voir section 4-3.2) et les structures permettant de l'expliciter (voir section 4-5).

2.2 ONTOLOGIE DIFFÉRENTIELLE ET ARBORESCENCE

Au vu d'une telle caractérisation de l'ontologie comme taxinomie d'essences différentielles, il apparaît que le contenu d'une essence dépend directement des liens taxinomiques qui l'unit aux autres essences, puisque ce sont eux qui permettent de dégager les différences pertinentes : la structure de la hiérarchie taxinomique conditionne par conséquent le contenu des essences puisque la structure détermine quels liens et chemins sont possibles ou non entre des essences.

Par sa nature taxinomique, l'ontologie doit posséder une structure de treillis. Un treillis se définit comme un ensemble muni d'une structure d'ordre telle que pour toute paire d'éléments, cette paire possède une borne inférieure et une borne supérieure dans l'ensemble. La relation d'ordre est ici la relation « est-un ». Etant donné qu'une ontologie n'est un système différentiel que sur le fond d'un système identitaire, il en découle que toute paire d'éléments, pour être systématiquement reliés dans l'ontologie, c'est-à-dire être dits appartenir à une même ontologie, doivent posséder un fond identitaire à partir duquel expliciter leurs différences. Ce fond identitaire appartient en outre à l'ontologie elle-même. Par conséquent, il existe toujours, pour toute paire, un fond identitaire constituant leur borne supérieure pour la relation « est-un ».

En d'autres termes, cela signifie qu'il existe un nœud suprême T, ou *summum genus*, qui représente le fond identitaire ultime auquel participent tous les nœuds. Un tel nœud ultime existe nécessairement dans la mesure où, constituant un *même système*, ils sont tous commensurables ou comparables, c'est-à-dire que leurs différences sont explicitables à partir d'un fond identitaire. Ce nœud représente le fait que toutes les essences *sont*, peu importe leur manière d'être. A partir du moment où elles *sont*, elles partagent toutes l'existence et participent à la positivité du réel, son effectivité, sa présence. Ce nœud est la pure positivité ou effectivité du réel, considéré indépendamment de toute détermination ou particularisation. Ce nœud est le fond identitaire de tous les nœuds puisque de tous il affirme qu'ils *sont* indépendamment de *ce* qu'ils sont.

Symétriquement, il faut poser un nœud $\perp$ qui représente la pure absence, le néant, la négation de toute réalité. Autrement dit, ce nœud ne représente rien d'effectif ou de positif : à proprement parler, il n'est rien. Ce néant qui est pure absence, il faut le comprendre en suivant la notion aristotélicienne de *privation*. La privation, c'est le fait d'être privé d'une détermination ou d'une propriété. Pure négation sans positivité, la notion de privation rappelle qu'être privé de quelque chose, ce n'est pas encore être soi-même quelque chose.

Dans une ontologie, ce nœud souligne le fait que chaque nœud est la négation des autres nœuds, dans la mesure où, n'étant pas eux, il est privé de ce qui fait leur être ou essence, c'est-à-dire de leurs propriétés. Ce nœud constitue donc une borne inférieure pour tous les nœuds de la taxinomie dans la mesure où ils partagent tous le fait de se nier réciproquement.

Mais, étant acquis qu'une taxinomie ontologie est un treillis, demeure la question de savoir si le treillis doit être contraint pour être un arbre ou non. En d'autres termes, deux essences différentielles, peuvent-elles partager autre chose qu'un fond identitaire et une pure négation des uns des autres ? Ou peuvent-elles par exemple posséder des essences plus spécifiques communes ? Des essences peuvent en effet posséder une essence plus générique comme fond identitaire commun sans que ce dernier doive être nécessairement le *summum genus*. La raison est que des essences peuvent partager, pour y relever leurs différences déterminantes, un fond identitaire constitué non seulement du fait qu'elles *sont* mais aussi de *ce* qu'elles sont. En revanche, il est moins clair si des essences peuvent partager comme négation commune des déterminations positives, c'est-à-dire des nœuds fils communs distincts de $\perp$. Pour cela, il faut préciser ce qu'il faut comprendre par la négation dans une taxinomie ontologique.

Comme nous allons le voir, la structure en forme de treillis et le type de chemins taxinomiques qu'elle contient implique un certain type de signification ou de contenu pour les essences, qui est différent de celui prescrit par une structure arborescente ; la négation ne possède pas en effet la même signification pour construire le système de différences. Dans le cas arborescent, la négation est une négation intensionnelle ; dans le cas du treillis, la négation est une négation extensionnelle.

La raison en est simple : la négation intensionnelle est une négation de propriété. Dire que des essences se nient réciproquement, c'est affirmer qu'elles s'opposent sur des propriétés qui s'excluent. Supposons alors qu'un nœud possède deux nœuds pères. Par définition de la taxinomie comme système de différences, les deux nœuds pères s'opposent l'un à l'autre dans la mesure où affirmer l'un, c'est nier l'autre. Il est donc contradictoire d'affirmer qu'ils puissent posséder un nœud fils commun puisque ce nœud, héritant par le lien identitaire « est-un » des caractéristiques déterminant ses pères, se retrouverait caractérisé par des propriétés héritées d'un père qui excluraient celles héritées de l'autre. Ce nœud aurait pour propriétés déterminantes à la fois P et $\neg P$, ce qui est impossible. Si tout nœud ne peut avoir qu'un seul père, cela signifie que la hiérarchie taxinomique, construite dans un domaine non formel où les termes se définissent différentiellement, c'est-à-dire par opposition et différence, est nécessairement arborescente.

Dans le cas d'une négation extensionnelle, les propriétés ne se définissent pas par le fait qu'elles s'excluent, mais par la classe d'individus qu'elles dénotent. Il est donc tout-à-fait possible que des nœuds différents possèdent des fils communs dans la mesure où, se définissant de manière positive par ce qu'ils dénotent et non de manière négative par ce qu'ils nient, ils peuvent partager leur contenu dénotationnel. La négation extensionnelle ne contraint donc pas la structure du treillis en arbre.

L'opposition entre arbre et treillis pour la structuration des taxinomies ontologiques renvoie par conséquent à l'opposition entre les connaissances linguistiques et les connaissances formalisées. En effet, dans la mesure où il n'existe pas quelque chose comme un univers du discours composé d'individus singuliers et indépendants dans les connaissances linguistiques, la négation extensionnelle ne peut s'y appliquer : seule la négation intensionnelle vaut, qui implique pour le treillis taxinomique une structure d'arbre.

2.3 Taxinomie formelle et graphes conceptuels

Si cette conception des choses bien conforme aux conceptions classiques, les taxinomies fournies par la tradition en zoologie et botanique adoptant le plus souvent une structure arborescente de type aristotélicien, les travaux modernes infirment cette vue dans la mesure où ils adoptent une conception extensionnelle et formaliste des connaissances linguistiques et les systématisent en treillis non arborescents, comme on peut s'en apercevoir en considérant les ontologies explicitées dans le formalisme des graphes conceptuels.

Les graphes conceptuels de Sowa (Sowa, 1984) sont un formalisme de représentation des connaissances qui a servi de cadre informatique pour la mise au point de notre méthodologie. Un graphe conceptuel est un graphe bipartite contenant d'une part des nœuds appelés « concepts » et représentés par des boîtes, et d'autre part des nœuds appelés « relations » représentés par des cercles. Les concepts sont reliés entre eux par des relations.

Un nœud concept est une entité composée de deux éléments, son type et son référent. Le type du concept est une entité intensionnelle correspondant au sens (le *Sinn* frégéen) du concept. Ce peut être par exemple la notion de « chien », ou de « nombre entier ». Le référent est un objet puisé dans la base extensionnelle dans laquelle consiste le monde réel. Un concept est donc la désignation d'un objet réel (extensionnel) à travers son type (intensionnel). Il se représente habituellement à travers la la notation `[type : référent]`.

Des exemples de concept sont ainsi `[chien : Fido]` ou bien `[nombre entier : 2]`. Il est également possible de parler d'un référent sans avoir à le désigner explicitement : c'est le « référent générique » noté `*`. Le concept dont le référent est le référent générique signifie alors qu'il existe un objet indéterminé dont le type est celui du concept. Par exemple, `[chien : *x]` signifie qu'il existe un chien. Un nœud relation est une entité composée d'une seul élément son type. Il est noté `(type)`.

Les types des concepts sont regroupés en un treillis taxinomique, le « treillis des types ». Ce treillis précise les liens de nature intensionnelle existant entre les types, en particulier les liens de nature ontologique. Par exemple, on consignera dans le treillis le fait que tout chien est un animal en établissant le type « chien » comme un sous-type du type « animal ». Par là, on s'attend également à ce que toutes les propriétés vraies des animaux se trouvent également vérifiées par les chiens, dans la mesure où, étant un animal, un chien en possède les propriétés inhérentes.

Les graphes conceptuels ne peuvent être construits qu'à partir du treillis des types. Ce dernier constitue donc la base des descripteurs du domaine, le vocabulaire permettant de nommer les choses.

Les graphes conceptuels ont été conçus pour être un formalisme de représentation des connaissances en IA et plus particulièrement pour constituer un outil de traitement de la langue naturelle. Une propriété des graphes régulièrement revendiquée par Sowa est en effet d'avoir un pouvoir d'expression comparable à celui de la langue, et de même nature.

Cela signifie que les graphes conceptuels sont destinés à appréhender l'objectivité qualitative rencontrée à travers la langue naturelle. Les graphes conceptuels reprennent donc à leur compte les objectifs fondamentaux de l'IA à savoir d'appréhender les connaissances linguistiques, c'est-à-dire les connaissances exprimées dans la langue naturelle. Il est donc naturel que les graphes conceptuels puissent à la fois revendiquer représenter les connaissances et traiter l'analyse de la langue naturelle.

Mais, alors qu'ils partagent l'objectif poursuivi dans notre travail de systématiser et de rationaliser les connaissances linguistiques et de leurs objectivités corrélées, les graphes conceptuels construisent des *treillis* de type de concept, et non des *arbres* comme nous l'avons postulé plus haut. C'est donc qu'il y a un changement de point de vue sur la logique des connaissances qui entraîne une modification formelle des structures fondamentales.

Le changement consiste dans le fait que la logique sous jacente à l'entreprise des graphes conceptuels n'est pas une sémantique différentielle, mais une logique formelle extensionnelle.

Les graphes conceptuels reposent en effet sur une conception résolument extensionnaliste du monde. Le monde est constitué de tous les référents possibles, objets individuels correspondant à la base extensionnelle à partir de laquelle construire les dénotations des relations (relations n-adiques) et des concepts (relations monadiques ou propriétés).

Chaque objet se présente donc au concepteur comme les habitants de l'Eden à Adam pour recevoir un nom, c'est-à-dire un type. Le type est intensionnel dans la mesure où il est possible d'imaginer un monde possible dans lequel il possède une extension différente de celle qu'il a dans le monde réel. Mais sa signification est toujours extensionnelle.

Par ailleurs, pour confirmer cette affirmation, les graphes conceptuels peuvent être traduits en logique du premier ordre, extensionnelle comme on le sait, au moyen d'une fonction Φ dont les principes sont les suivants : un concept dont le référent est un objet particulier, `[P : a]` se traduit par la formule logique $P(a)$; le concept `[P : *x]` se traduit par la formule $\exists\ x\ P(x)$. Enfin, les relations n-adiques se traduisent par les prédicats n-adiques portant sur les référents des concepts qu'elles relient : `[P : a]--(R)-->[Q : *x]` se traduit donc par la formule $\exists\ x\ P(a) \wedge Q(x) \wedge R(a,x)$.

La différence entre la structure arborescente et la structure en treillis provient donc d'une différence entre les types de sémantique permettant d'interpréter et de comprendre les connaissances exprimées. Dans le cas des graphes conceptuels, on utilise une sémantique formelle référentielle où signifier n'est pas différer de ses voisins dans le système de la langue ou des connaissances, mais référer à des objets du monde. Dans ce cas, comme nous l'avons noté, il est tout à fait possible d'avoir des nœuds possédant plusieurs pères puisque les significations se composent de manière non différentielle. En revanche, dans le cas où le système des connaissances est extrait d'expressions linguistiques et qu'il repose en son principe sur la sémantique de la langue, la sémantique adoptée est une sémantique linguistique différentielle. Et comme on l'a vu, la structure taxinomique implique que l'on hérite des différences de ses pères et donc deux pères qui s'opposent ne peuvent avoir un fils commun.

Autrement dit, dans une taxinomie exprimant non des connaissances formalisées mais des connaissances normalisées, il n'existe pas de compositionnalité des labels des nœuds. Dans une structure en forme de treillis, on autorise qu'un nœud soit le fils de plusieurs pères, c'est-à-dire qu'il hérite de leurs déterminations de sens, c'est-à-dire que l'on compose ces déterminations pour composer un nouveau nœud. Cela signifie que le label de ce dernier, nom donné à ses déterminations de sens, est le composé des labels des nœuds à partir desquels il est construit. Par conséquent, la structure en treillis entraîne la compositionnalité des labels. En revanche, la structure d'arbre, interdisant qu'un nœud ait plusieurs pères, implique qu'il est impossible de composer des labels, mais seulement de les spécialiser en explicitant de nouvelles propriétés : alors que la structure en forme de treillis autorise la spécialisation par composition des labels déjà posés, la structure arborescente n'autorise que la spécialisation par explicitation de nouvelles propriétés.

En d'autres termes, cela signifie que le sur-type d'un type dans une taxinomie différentielle constitue son contexte d'interprétation. Ce sur-type doit donc être nécessairement unique pour qu'il n'y ait pas d'ambiguïté. Or, l'unicité du sur-type est précisément ce qui caractérise un arbre par rapport à un treillis.

Par conséquent, la structure globale de l'ontologie régionale est nécessairement un taxinomie arborescente.

3 Structure locale de l'ontologie

Il s'agit à présent de se tourner vers la détermination locale de l'ontologie régionale, c'est-à-dire de considérer comment un concept peut recevoir une explicitation normée permettant de fixer sa signifiation en vue d'une expérience possible, la tâche du système, à partir des différences avec ses voisins.

3.1 Essences : définition

Classiquement l'essence d'un objet consiste dans les conditions nécessaires et suffisantes qu'il doit vérifier pour être ce qu'il est. La notion même d'essence évoque les substances qui se définissent par ce qu'elles ne peuvent manquer d'être, par tout ce qu'elles ne peuvent pas ne pas être. Cela signifie que l'on détermine le contenu essentiel d'une notion indépendamment des autres, par ce qu'elle est, et non par ce qu'elle n'est pas. Cela implique que l'on est capable d'avoir accès directement aux objets du domaine et d'en déterminer les propriétés. Cela voudrait donc dire que l'on revient à une ontologie primitive donnée au préalable par rapport à laquelle le système conceptuel s'ajuste. Il ne s'agirait alors que de la description intensionnelle d'une collection d'individus prédonnée, à l'instar d'une sémantique *référentielle* où le sens des concepts se définit de manière extensionnelle en termes de classes d'individus ou référents.

Mais il ne peut s'agir de cela dans la mesure où, dans les domaines où nous nous situons, ce que sont les objets, leur nature, n'est pas directement accessible. C'est pourquoi nous sommes passés par l'expression linguistique des connaissances qui les visent. Une approche référentielle nous est interdite.

En fait, l'identité d'un objet, son essence, est de nature différentielle et non référentielle : ne disposant pas d'une ontologie méta-linguistique prédonnée, il n'est pas possible de fixer les essences en leur communiquant les propriétés des référents. Il faut au contraire déterminer les essences et viser le référent à partir des propriétés linguistiques qui énoncent ses manières d'être. Or, les propriétés linguistiques individuent les objets du discours les uns par rapport aux autres : un objet du discours, une essence, n'est pas défini par ce qu'il est et ne peut pas ne pas être, mais au contraire par tout ce qu'il n'est pas, et ne peut pas être sous peine de perdre son identité ou essence. C'est donc une approche onomasiologique qu'il faut adopter dans une perspective différentielle, et non une approche sémasiologique dans une perspective référentielle (Rastier, 1987).

Cela veut donc dire qu'un concept, et le terme technique qui le vise, n'est pas objet au sens de la sémantique formelle, mais une essence, une notion. Cette notion se définit comme une position, une valeur dans le système, non de la langue, mais des connaissances. Il s'agit donc d'essence différentielle et non d'essence substantielle dans la mesure où une substance est inanalysable en propriétés. La normalisation correspond donc au processus permettant de passer du système de la langue à un système non régi par le système fonctionnel de la langue, mais par le domaine.

3.2 Essences : détermination

Dans la mesure où l'ontologie régionale est une taxinomie arborescente, les plus proches voisins d'un nœud sont constitués par son père et ses frères. Du fait de la structure taxinomique impliquant une relation d'héritage, un nœud n'est pas déterminé par ses fils, alors que les fils le sont par leur père. Par conséquent, le voisinage proche d'un nœud participant à la détermination de son contenu ne comprend que le père et les frères, sans les fils, pour ne retenir que les plus proches voisins.

C'est dans le cadre de cette structure locale qu'il faut être capable d'expliciter les traits différentiels qui constitueront les prescriptions interprétatives conférant au nœud sa valeur prescriptive. Les traits différentiels doivent en effet fonctionner non comme des descriptions sémantiques du contenu du nœud mais comme des prescriptions : c'est ce que doit vouloir signifier ce nœud. La détermination des traits différentiels s'effectuent en retenant, parmi les traits proposés par la description sémantique, ceux qui correspondent à la tâche visée en vue de laquelle les nœuds de l'ontologie doivent valoir comme connaissances.

Nous proposons quatre critères permettant de déterminer l'essence d'un concept en fonction de son voisinage dans la hiérarchie. Quand un concept est placé dans la hiérarchie, il convient de préciser les quatre points suivants (Zweigenbaum *et al.*, 1995a) :

Principe de communauté avec le père – P1 : si un concept est placé sous un autre, du fait qu'il s'agit d'une hiérarchie de subsomption, c'est qu'il possède dans sa définition les propriétés caractérisant le père, il en hérite. Il convient donc de préciser en quoi le nouveau concept possède les déterminations du concept père, en quoi choisir le fils, c'est aussi affirmer le père ;

Principe de différence avec le père – P2 : si néanmoins un concept est placé sous un autre concept, c'est qu'il en est différent ; sinon, en effet, on en resterait au concept père. Il faut donc préciser en quoi le concept fils est différent du père, préciser en quoi choisir le fils, c'est nier le père ;

Principe de différence avec les frères – P3 : si un concept possède une fratrie, c'est que l'on a jugé bon de poser un nouveau concept dans la mesure où les concepts frères déjà posés ne permettaient d'exprimer la notion en vue ; il faut donc préciser en quoi choisir le nouveau concept, c'est nier les concepts frères ;

Principe de communauté avec les frères – P4 : si un concept est posé avec ses frères, c'est qu'il possède une dimension sémantique commune avec eux sur laquelle chacun ils s'opposent. De cette manière, on est assuré que choisir une valeur sur cette dimension sémantique, c'est exclure les autres. Par conséquent, ce principe permet de respecter le précédent.

Le dernier principe est en effet nécessaire : supposons que l'on ait choisi pour opposer deux concepts frères le fait que l'un se rapporte aux traitements chirurgicaux, et l'autre aux actes infirmiers. Le fait d'avoir un traitement chirurgical n'exclut pas que ce soit également un acte infirmier. Il peut très bien s'avérer qu'un même acte relève des deux catégories. Par conséquent, l'opposition entre les frères est mal décrite. Il faut en effet que l'opposition entre les frères soient posées intensionnellement, c'est-à-dire que l'on puisse déduire automatiquement du fait que l'on a retenu un concept que ses frères soient écartés. Par exemple, si les frères s'opposent sur la dimension sémantique unique de « qui fait l'acte », on peut définir les actes infirmiers faits par les infirmiers et les actes médicaux, faits par les médecins. On est ainsi assuré que si un acte est fait par un infirmier, il est exclu qu'il soit fait par un médecin. De la même manière, sur la dimension sémantique de la « fraternité », les valeurs sémantiques « frère » et « sœur » sont mutuellement exclusives [5].

Ces principes reçoivent l'appellation de *principes différentiels*. Ils confèrent à l'ontologie régionale des propriétés qui s'avèrent très utiles dans le cadre d'une modélisation effective

5. Cela correspond au fait que les différences entre concepts frères ne peuvent être défini que dans le cadre d'une classe relevant du système fonctionnel de la langue ou des connaissances. Ainsi, « frère » et « sœur » appartiennent à un même taxème (Rastier, 1987).

de connaissances d'un domaine. Ainsi, selon (Gruber, 1995), une ontologie doit posséder les propriétés suivantes pour être propre à un usage effectif :

Clarté : une ontologie doit communiquer la signification visée des termes définis. Les définitions doivent être autant que possibles indépendantes de considérations informatiques d'implémentations ou de considérations sociologiques sur l'usage des termes. Gruber note que le formalisme et des définitions axiomatiques permettent de remplir ces conditions, même si toutes les définitions doivent être documentées en langue naturelle. Dans notre ontologie, ces objectifs de « clarté » sont assumés au niveau de la langue sans recourir au formalisme ;

Cohérence : l'ontologie doit pouvoir établir quelles inférences sont conformes aux définitions qu'elle contient, que ce soit de manière formelle ou informelle.

Extensibilité : il doit être possible de définir de nouveaux termes pour des usages spécifiques à partir des termes existants, de telle manière qu'il ne soit pas nécessaire de revoir les définitions existantes.

Un biais minimal d'encodage : (*minimal encoding bias*) la conceptualisation du domaine que propose l'ontologie ne doit pas dépendre d'une implémentation particulière ; elle doit s'effectuer au niveau des connaissances et non au niveau des programmes ;

Engagement ontologique minimal : une ontologie doit assumer un engagement ontologique minimum pour permettre qu'une même ontologie puisse être spécialisée dans différentes directions.

Dans notre ontologie, chaque nœud est explicité par les principes différentiels qui spécifient dans quel but ce nœud a été posé à cet endroit dans l'arbre. Autrement dit, les principes différentiels forcent le modélisateur à expliciter en quoi la notion qu'il vise à travers le nœud qu'il pose renvoie à une connaissance que la résolution de la tâche mobilise. Le nœud ne doit pas être « gratuit », mais il doit être motivé en fonction de la tâche visée, pour laquelle il véhicule une connaissance. A ce titre, notre ontologie remplit les objectifs de clarté : chaque nœud est motivé et justifié eu égard à la tâche.

Les principes différentiels contraignent une cohérence locale entre un nœud et son père et ses frères. Comme nous le verrons plus bas, la structure arborescente permet de passer au niveau global et fournit ainsi des critères pour vérifier la cohérence non plus locale mais globale.

Les principes différentiels permettent par ailleurs de compléter ou d'étendre l'ontologie sans modifier les structures existantes. Mais, ils permettent en outre de spécifier *comment* étendre et *où* ajouter de nouveaux concepts. En effet, on peut formuler (Bouaud *et al.*, 1994) en pseudo-code, un protocole récursif, PlaceOrSearchType, qui utilise ces principes pour placer un nouveau concept. Supposons que Type est le concept qu'il faut rechercher ou ajouter s'il n'est pas retrouvé dans l'ontologie. Root est le nœud racine d'une structure locale, c'est-à-dire c'est le père d'une fratrie. Le protocole est initialisé avec le concept à rechercher ou ajouter et le nœud racine de l'ontologie, T. Selon l'usage, le pseudo-code est en pseudo-anglais.

PlaceOrSearchType(Type,Root)

0. **If** Type et Root sont identiques
   Then 1. Return(Type)
   Else 2. vérifier si Type partage les propriétés similaires à la racine Root [P1]
   3. **If** un axe sémantique n'a déjà été déterminé [P4]
      Then 3.a. Déterminer un axe sémantique pertinent
   4. **Let** Dif la projection de Type selon cet axe
   5. **Let** Sub défini par Root puis Dif [P2]
   6. **If** Sub n'est pas un fils ou sous-type de Root [P3]
      Then 6.a. ajouter le nouveau sous-type Sub à Root
   7. PlaceOrSearchType(Type,Sub)

Il faut remarquer que cette procédure, construire sur la structure locale, opère globalement du fait de la structure arborescente de l'ontologie.

Enfin, l'ontologie est par définition indépendante au codage. En revanche, il n'est pas clair que l'engagement ontologique soit minimal : l'objectif n'est pas ici d'être générique ou réutilisable, mais de représenter adéquatement les concepts du domaine pour être capable de caractériser quels sont les objets du domaine.

4 Du local au global

Mais la principale propriété des principes différentiels consiste dans le fait qu'ils confèrent aux nœuds une signification univoque et indépendante du contexte, permettant de voir en eux des primitives de modélisation.

En effet, les prescriptions interprétatives que sont les principes différentiels utilisent le fait que le nœud consiste en un label exprimé à travers des unités linguistiques pour contraindre l'interprétation que tout locuteur de langue et spécialiste du domaine associe à cette unité. Autrement dit, les principes contrôlent l'interprétation linguistique des lables des nœuds pour qu'ils renvoient à des connaissances du domaine.

Mais les principes différentiels ne donnent que le sens local d'un nœud. Or, l'ontologie est un système conceptuel dans lequel un nœud prend sa valeur prescriptive en fonction de l'ensemble du système et pas seulement par rapport à ses plus proches voisins. Mais la structure arborescente permet de passer du local au global, comme on l'a vu avec la procédure d'ajout ou de recherche d'un concept. En effet, les déterminations locales se somment pour donner la signification globale du nœud : en parcourant le chemin reliant un nœud à la racine de l'arbre, au « quelque chose en général » du domaine, il est possible de sommer les traits différentiels qui ont été à chaque fois posé pour différencier un ancêtre situé sur ce chemin de son voisinage immédiat. Un nœud signifie toutes les différences qui ont été posée sur le chemin le reliant à la racine pour différencier un nœud du chemin de son père et de ses frères. Autrement dit, les principes différentiels et la structure arborescente permettent de retrouver la signification d'un nœud à partir de sa position dans l'ontologie.

Par ailleurs, dans la mesure où les principes différentiels explicitent l'interprétation qu'il faut donner aux nœuds de l'ontologie, ils assurent une interprétation non contextuelle des signifiants linguistiques adoptés pour exprimer les labels des nœuds. En effet, les principes fixent un contexte particulier, celui de la tâche, si bien que les significations élaborées dans ce contexte deviennent insensibles au contexte de l'interprétation imposé par un utilisateur de l'ontologie : les principes différentiels imposent le contexte d'interprétation et l'utilisateur s'y conforme, retrouvant ainsi la connaissance représentée dans l'ontologie, au lieu de s'en remettre à la variabilité du contexte de lecture.

Si bien que les principes différentiels permettent aux essences différentielles, c'est-à-dire aux concepts signifiés par les nœuds de fonctionner comme des primitives, dans la mesure où, non contextuelles et univoques, elles peuvent être composées pour construire de nouvelles significations.

La normalisation sémantique que permettent les principes différentiels effectue en effet un travail préliminaire de définition des labels qui les détermine de manière à en autoriser la composition. En effectuant ce travail, on fait passer les labels du statut sémantique de *composant* à celui logique de *primitive*. Il faut donc distinguer deux étapes, l'une élaborant un *arbre de composants sémantiques normalisés*, correspondant à une *systématisation* des connaissances telles qu'elles sont, l'autre élaborant un *treillis de primitives conceptuelles*, correspondant à une *formalisation* des connaissances. Or, il faut bien voir qu'il est impossible de construire un treillis de primitives conceptuelles sans être passer au préalable par un arbre de composants. Cette étape est en effet indispensable pour donner un sens au treillis.

Reconsidérons, pour fixer les idées, le cas des graphes conceptuels. Les types des concepts sont des primitives de représentations à partir desquelles construire les graphes. Les primitives reçoivent une signification relationnelle ou combinatoire au sein du formalisme en fonction de leur place dans le treillis et des relations qui peuvent les relier. En d'autres termes, les primitives ne possèdent aucune signification par elles-mêmes, mais uniquement par l'utilisation qui en est faite dans l'écriture des graphes, c'est-à-dire en fonction de leur position dans le treillis, puisque l'utilisation dépend de cette position. C'est la signification qu'elles reçoivent par la syntaxe.

Mais les graphes doivent posséder, outre une signification relationnelle interne, syntaxique, une signification externe ou sémantique dans la mesure où ils doivent représenter des connaissances, c'est-à-dire renvoyer au monde. Il faut donc interpréter les graphes conceptuels, leur donner une signification. Mais, alors que les relations entre les concepts et les types sont imposées par le formalisme et les règles de déduction, la signification des primitives elles-mêmes est sous-déterminée par le formalisme : leur interprétation est entièrement libre, pour peu qu'elle soit compatible avec leur signification relationnelle interne. Les constructions conceptuelles s'obtiennent par des règles de *composition* syntaxique appliquées aux primitives, types de concept ou types de relation. Par définition de la sémantique formelle, la signification sémantique de la construction est une composition de la signification sémantique des primitives : il est indispensable de posséder d'abord l'interprétation des primitives pour avoir celle de la construction. Mais, comme son nom l'indique, la primitive est « première » dans le système : condition de la construction, elle ne peut être conditionnée par celle-ci ; la primitive ne peut tirer du système de construction son interprétation.

Où prendre alors l'interprétation des primitives ? Puisqu'il s'agit de représenter des connaissances, l'interprétation des primitives découlent des connaissances à représenter. Mais ces connaissances ne sont accessibles qu'à travers leur expression linguistique. Par conséquent, c'est dans le système différentiel de la langue et des connaissances qu'elle permet d'exprimer qu'il faut puiser l'interprétation des primitives.

Mais, dans le système de la langue, les significations ne sont pas compositionnelles, comme on l'a vu plus haut, contrairement à ce qu'exige la représentation logique des connaissances. Par conséquent, il ne suffit pas de faire simplement correspondre aux primitives l'expression linguistique d'une connaissance : ressortissant à une sémantique non compositionnelle, elle ne peut être utilisée sans autre forme de procès dans la taxinomie formelle, sous peine d'obtenir, comme nous l'avons vu plus haut avec le logotron, des incohérences. Il faut par conséquent *normaliser* la signification linguistique des connaissances pour que la primitive possède une signification univoque et que sa sémantique puisse être intégrée à la construction logique des graphes.

Comment fixer leur signification ? Il est impossible de les définir en termes extensionnels puisque l'extension n'est pas accessible en tant que telle. C'est-à-dire qu'il est impossible d'affirmer que la notion de « thrombus » correspond à tel ensemble d'individus du monde puisqu'il est impossible de le montrer ou d'y accéder indépendamment de la signification du terme qu'il est destiné à définir... Il faut donc s'en remettre non pas une signification déterminée en termes de référence, mais à une signification en termes de signifié linguistique, c'est-à-dire de différences.

Fixer l'interprétation des primitives, c'est donc fixer le système de différences qui les déterminent : c'est par conséquent construire un système où la position différentielle d'une composante détermine de manière univoque sa signification et la détermine comme une primitive. Ce système est donc *arborescent* et *normalisé* : arborescent puisque, comme on l'a vu plus haut, un système de différence ne peut être qu'un arbre ; normalisé puisqu'il s'agit que cet arbre *fixe* la signification des nœuds de l'arbre pour qu'ils puissent tenir lieu de primitives conceptuelles dans le treillis des connaissances formalisées.

Reprenons donc les étapes de cette argumentation :

- dans un formalisme logique de représentation des connaissances, les types conceptuels sont des primitives de représentation qui, en tant que telles, ne sont pas déterminées par le formalisme lui-même ;
- le but est d'utiliser le formalisme pour représenter des connaissances sur une réalité externe, il est nécessaire de préciser l'interprétation des primitives : cette interprétation est nécessairement extrinsèque au formalisme puisque le formalisme repose sur les primitives et leur donne une sens combinatoire interne au formalisme lui-même, mais on serait bien en peine de leur fournir une correspondance avec un univers d'interprétation ;
- interpréter les primitives, c'est donc sortir du formalisme pour leur donner une sémantique en termes de connaissances d'un domaine ; or, de telles connaissances sont exprimées en langue naturelle et possèdent une sémantique de type linguistique ; par conséquent, l'interprétation des primitives revient à leur donner une sémantique dans la langue du domaine ;
- les primitives sont des signifiés qui s'interdéfinissent par opposition et différences ; elles ne peuvent donc être structurées en un treillis, mais nécessairement en une hiérarchie arborescente ; c'est au moyen de cette hiérarchie qu'elles se retrouvent pourvues d'une signification univoque et déterminée ;
- l'arborescence retenue permet de fixer la sémantique des labels en la normalisant ;

ainsi, ces labels acquièrent l'univocité leur permettant de fonctionner comme des primitives.

Si l'affirmation que la taxinomie des types primitifs d'un formalisme de représentation est nécessairement arborescente peut paraître paradoxale, c'est qu'il faut distinguer deux manières de considérer cette taxinomie :

la taxinomie régionale qui sert à déterminer le sens des primitives ; dans cette taxinomie, les types ne sont pas des primitives, mais seulement des unités linguistiques et conceptuelles dont il faut fixer la signification ; la taxinomie régionale est un arbre de composantes sémantiques normalisées ;

la taxinomie formelle qui structure les primitives du formalisme de représentation ; ces primitives possèdent une signification fixe et stable et sont utilisées pour construire les structures composées que sont les graphes conceptuels.

La taxinomie régionale ne contient pas de primitives, mais permet de les définir. La taxinomie formelle contient des primitives, qu'elle ne peut par conséquent définir. La taxinomie régionale contient des nœuds renvoyant à des unités linguistiques et conceptuelles dont le contenu est de nature différentielle. Elle est donc nécessairement arborescente. La taxinomie formelle contient des primitives dont l'interprétation est donnée par la taxinomie régionale, et mises en œuvre dans la syntaxe du formalisme pour construire les graphes. La taxinomie formelle n'est donc pas nécessairement arborescente et peut hériter des propriétés de compositionnalité associées au formalisme.

Il est nécessaire de distinguer les taxinomies régionale et formelle dans la mesure où le formalisme prescrit la manière d'utiliser les primitives pour construire les graphes, mais ne donne aucune indication quant à leur interprétation. La taxinomie régionale donne l'interprétation des primitives sans prescrire leur utilisation dans le formalisme. La taxinomie régionale est donc arborescente pour que les labels conceptuels aient une signification déterminée et puissent, grâce à cette détermination, fonctionner comme les primitives du formalisme de représentation.

Dès lors que les essences différentielles sont utilisées comme des primitives, il devient possible d'adopter non plus une sémantique différentielle mais une sémantique référentielle. En effet, chaque essence possède une signification correspondant à la tâche posée pour laquelle elle définit les objets sur lesquels la tâche opère. A partir du moment où les principes différentiels fixent la signification des concepts, ils en fixent la référence. D'une certaine manière, le travail interprétatif de la normalisation sémantique consiste à fixer la référence, à construire une cohésion de sens gouvernée par la tâche permettant de déterminer la référence. Par exemple, le fait de déterminer la notion de « sténose » comme déformation géométrique d'une artère fixe sa référence comme étant une propriété géométrique d'un certain type d'objet, et non comme une classe d'objets.

Cette fixation de la référence permet alors de donner la matière dont les travaux de l'ontologie formelle ont besoin : en établissant en effet qu'une sténose est une propriété et non un objet physique, il devient possible d'appliquer les axiomes logiques de l'ontologie formelle concernant les catégories comme « propriétés » ou « objets physiques ». Ces axiomes trouvent un domaine et des objets sur lesquels s'interpréter.

En reposant sur le principe que les essences s'analysent en traits sémantiques car renvoyant à un système de différences au lieu d'une ontologie primitive de référents, les principes différentiels contraignent l'analyse de manière à ce que les signifiés soient objectivants, reflétant le point de vue d'une action possible dans le cadre d'une tâche spécifiée dans la pratique d'un domaine particulier. Dès lors que chaque nœud de l'arbre est déterminé par un jeu de relations nécessaires et suffisantes, ce qui était impossible auparavant, du fait de la non-compositionnalité de la langue, le devient, à savoir la description du domaine comme un ensemble d'individus.

Par conséquent, la normalisation et la formalisation dégage chacune un type de prescription que l'on peut caractériser en reprenant la notion d'« engagement ontologique » que nous avions rencontrée dans le chapitre 2. En effet, la normalisation prescrit la signification des labels utilisés en termes de différences intralinguistiques de signification avec les autres labels : utiliser un label, c'est s'engager à ce qu'il nie ou s'oppose à tel ou tel autre label et de telle ou telle manière. Il s'agit d'un *engagement sémantique*. La formalisation, construite à l'aide de la normalisation, prescrit la signification des labels utilisés en termes de références : utiliser un label, c'est s'engager à ce qu'il dénote tels ou tels objets, et de telle ou telle manière. Il s'agit d'un *engagement ontologique*. Notre travail méthodologique a donc consisté à dégager un niveau d'engagement supplémentaire, un niveau où l'on s'engage sur le sens avant de s'engager sur les objets. Cela est nécessaire dans la mesure où pour s'engager sur les objets, il faut qu'ils aient été constitués ou modélisés, ce que permet de faire l'engagement sémantique.

Mais, il ressort de nos considérations qu'il est nécessaire que l'engagement ontologique reflètent l'engagement sémantique, et que la sémantique référentielle prescrivent des significations compatibles avec la sémantique différentielle. Il faudra également que l'opérationnalisation, but final de toute méthodologie d'acquisition et d'exploitation des connaissances, soit également conforme à la sémantique référentielle, et à travers elle, à la sémantique différentielle. En d'autres termes, il faut que l'engagement *computationnel*, selon lequel chaque label est une primitive non de formalisation mais de programmation et commande par conséquent le comportement du système, soit compatible avec les engagements ontologiques et sémantiques. Nous détaillons les relations entre les niveaux ontologiques en reprenant chaque palier de la modélisation : régional, formel, computationnel.

5 Ontologie régionale et normalisation sémantique

D'après les considérations qui précèdent, une ontologie régionale se présente donc comme un arbre d'essences différentielles déterminées dans le cadre d'une tâche d'un domaine, permettant ainsi à la signification des labels de prendre la valeur d'action possible et d'acquérir ainsi le contenu prescriptif faisant d'elle une connaissance du domaine. Chaque essence est déterminée à l'aide des quatre principes différentiels : ils constituent la grille de lecture permettant d'investir les labels des nœuds de la taxinomie d'un contenu prescriptif. Il s'agit à présent de préciser comment se structure plus avant une ontologie régionale. Les exemples sont pris dans l'ontologie régionale que nous avons conçu selon ces principes pour le projet Menelas et on se reportera à la figure 4.3.

5.1 Définition et encyclopédie régionales

Un nœud de la taxinomie n'est pas seulement un nom, mais une notion ou concept du domaine, défini à partir des autres nœuds. C'est pourquoi, outre les quatre principes déjà mentionnés, un nœud se définit également par des assertions définitionnelles déterminant en quoi consiste le contenu prescriptif du label décrit. Ces assertions constituent d'une part ce que nous appelons la *définition régionale* et d'autre part l'*encyclopédie régionale*.

Les définitions assertionnelles précisent en effet d'une part les propriétés toujours vérifiées par le concept, la définition régionale donc, et d'autre part celles qui correspondent à des associations relevant de topoi normalisés du domaine, l'encyclopédie régionale. En d'autres termes, la définition régionale contient les sèmes inhérents caractérisant la notion explicitée, sèmes qui seront hérités par défaut pour chaque notion occurrence. Par exemple, la sténose est l'état d'une artère dont le diamètre est inférieur à la normale.

L'encyclopédie régionale contient les sèmes afférents spécifiés par les normes du domaine qui sont activés lorsque le contexte le requiert. L'encyclopédie régionale explicite les propagations de sens normalisées dans le domaine : par exemple qu'une dilatation puisse entraîner une resténose.

La définition régionale explicite le résultat de l'application des quatre principes comme grille de lecture sur le concept : ces propriétés ne font qu'expliciter la signification du concept dans la taxinomie définie par sa position dans la taxinomie. Ces propriétés explicitent donc les identités et les différences constitutives de l'essence différentielle.

Au niveau de l'ontologie régionale, ces définitions s'explicitent en langue naturelle : elles correspondent à une lecture interprétative intralinguistique de la taxinomie, elle-même formulée à l'aide de termes empruntés à la langue naturelle. En toute rigueur, les labels des nœuds d'une taxinomie peuvent être des symboles totalement quelconques dans la mesure où ils ne sont pas censés intervenir dans le processus de détermination du sens du nœud, ce sens étant complètement déterminé par la seule position : une taxinomie prescrit des rapports d'ordre en des nœuds. Mais, pour que la taxinomie ait un rapport au monde, nous avons montré qu'il est indispensable que ce label soit un signifiant de la langue et qu'il évoque par son signifié la signification du nœud. Ce signifié n'étant que linguistique, ne peut tenir lieu de contenu prescriptif du nœud et l'explicitation des 4 principes différentiels s'appliquent comme contraintes interprétatives pour que le nœud acquiert un contenu notionnel.

5.2 Interprétation sémantique de la taxinomie

La taxinomie est un arbre d'essences dont l'analyse différentielle est conditionnée par les principes. Ces principes imposent l'interprétation sémantique suivante des concepts de la taxinomie.

Chaque nœud de la taxinomie possède un label en langue naturelle ou pseudo-naturelle, c'est-à-dire dans un jargon propre à l'application. Ce label constitue une lexie possédant son signifié, sa sémie. Cette sémie s'analyse en traits sémantiques ou sèmes, exprimés en langue naturelle : c'est le caractère intralinguistique de l'analyse différentielle.

Les sèmes, exprimés en langue, constituent donc des lexies. On impose que ces sémies soient, autant que possibles, des labels de nœuds de la taxinomie. L'explicitation des sémies en sèmes repose donc sur le caractère circulaire de l'ontologie régionale prise comme système. Cela implique que les définitions et encyclopédies régionales qui explicitent l'analyse sémantique du label ne contiennent que les lexies verbalisées dans l'arbre.

Il faut donc bien noter que l'ontologie n'est pas une ontologie de descripteurs élémentaires, ou de primitives sémantiques, mais une ontologie de lexies constituant sémantiquement des objets. Les sémies signifiées par les labels sont des *composantes* sémantiques qui peuvent être utilisées pour expliciter le sens différentiel des autres labels. En d'autres termes, il n'existe pas d'un côté des labels, et de l'autre des traits différentiels. Les labels sont les traits différentiels que les labels utilisent pour s'interdéfinir. Par conséquent, les traits différentiels ne constituent pas un système de propriétés qui fonctionneraient comme des primitives de description des labels. L'ontologie régionale est circulaire dans la mesure où l'explicitation des nœuds est intra-ontologique.

De la même manière, l'ontologie n'est pas une hiérarchie de sème, mais de sémèmes. Il ne s'agit pas d'une hiérarchie de traits différentiels s'interdéfinissant, mais d'une hiérarchie de notions s'interdéfinissant, c'est-à-dire de définitions d'objets du domaine s'interdéfinissant.

En effet, si l'on avait une hiérarchie de traits différentiels ou sèmes, on aurait au sommet les traits caractérisant la dimension sémantique comme /humain/ *versus* /animal/, ou /végétal/ *versus* /animal/. Ces traits sont souvent lexicalisées et donc facile à mettre en évidence : « vénéneux » *versus* « venimeux » par exemple pour la seconde opposition, et « bouche » *versus* « gueule » pour la seconde. Ensuite, il faudrait consigner les traits caractérisant le champ sémantique. Le champ est un ensemble structuré de taxèmes pertinents dans une pratique concrète.

A un niveau de spécificité encore supérieur, il faudrait répertorier les *domaines*, comme la /cuisine/ ou la /navigation/. Un domaine est lié à une pratique sociale déterminée : c'est au sein d'un domaine que l'on détermine les « connaissances d'un domaine ». Un domaine se reconnaît habituellement par le fait qu'il n'y existe pas de polysémie lexicale et qu'entre les unités d'un même domaine, il ne s'établit pas de connexion métaphorique.

Enfin, au niveau le plus spécifique se situeraient les taxèmes : le *taxème* est la classe minimale en langue. C'est en son sein que sont définis les sèmes spécifiques du sémème ainsi que son sème le moins générique, le sème microgénérique ou taxémique qu'il partage avec les autres éléments du taxème. Les taxèmes reflètent des situations de choix : « sœur » et « frère » forment le taxème bien connu de la fratrie. Un taxème n'est pas obligatoirement composé d'antonymes binaires comme « faim » / « soif » ou « vendre » / « acheter ». Par exemple, « veine », « artère », « capillaire » constituent un taxème dont le sème microgénérique commun est le fait d'être un vaisseau sanguin.

> « On peut distinguer deux types de sèmes, correspondant à deux types de relations fonctionnelles : les relations symétriques et.ou réflexives, l'identité et l'incompatibilité notamment, définissent les éléments dits *inhérents*. Les relations antisymétriques et/ou non réflexives, notamment l'implication définissent les éléments dits *afférents* ; dans leur cas, il convient toujours de distinguer explicitement le sémème-source et le sémème-but de la relation fonctionnelle. Rappelons que les relations du premier type sont définies entre sémème au sein du même taxème, et celles du second type, entre sémèmes appartenant à deux taxèmes différents. » (Rastier, 1987, p.54)

Mais la preuve que notre ontologie ne peut consister en hiérarchie de traits sémantiques dont nous venons de voir les paliers de spécificité est que tous ces traits sémantiques ou sèmes dimensionnels, domaniaux, champêtres, et taxémiques, ne constituent pas une hiérarchie arborescente dans la mesure où un niveau de spécificité, par exemple la dimension, peut être transversal dans les oppositions qu'il définit aux autres niveaux de spécificité. Par exemple, le domaine /médical/ est transversal aux dimensions /animé /, /inanimé/ même s'il s'oppose au domaine /vétérinaire/ selon l'opposition /humain/ /animal/ qui relève bien du niveau dimensionnel.

L'ontologie régionale n'est donc pas une hiérarchie de traits sémantiques structurés selon leur spécificité respective et relative. Par ailleurs, puisqu'elle n'est pas non plus une hiérarchie de classes d'objet, structurant leurs mutuelles relations d'inclusion par un arbre ou un treillis, l'ontologie régionale n'est ni une hiérarchie de traits, ni une hiérarchie de classes extensionnelles, et est donc bien une hiérarchie d'essences.

En effet, un nœud placé sous un autre signifie que sa sémie participe des sémies des nœuds placés plus haut sur sa branche, héritant ainsi de leur déterminations positives et négatives, et des sémies des nœuds constituant sa fratrie.

Ainsi, l'essence « objet animé » est une sémie possédant le sème dimensionnel inhérent /animé/. Ce sème, conformément au principe de définition circulaire, constitue la sémie d'un nœud de l'arbre : un nœud placé sur la branche des propriétés, en l'occurrence des propriétés qualifiant les objets physiques. On tient compte ainsi d'une part que les dimensions ne s'organisent pas hiérarchiquement avec les autres niveaux de généralités, les sèmes dimensionnels sont sur une autre branche que les sémèmes les mentionnant, et d'autre part qu'elles interviennent dans des sémies organisées, quant à elles, hiérarchiquement.

6 Ontologie formelle et formalisation logique

Puisqu'il s'agit de construire un SBC dans le but de réaliser une tâche dans le domaine, ou d'en assister la réalisation, l'ontologie régionale ne suffit pas. Il faut en effet déterminer l'ontologie régionale en un format compatible avec une exploitation informatique. L'obtention d'un tel format passe par la détermination d'une *ontologie formelle*. L'ontologie formelle correspond à la formalisation de l'ontologie régionale en associant à la taxinomie non plus une sémantique différentielle, mais une sémantique formelle, se rapportant à une conception extensionnelle.

6.1 Formalisation et normalisation

Cette obtention n'est pas immédiate car la normalisation n'est pas une formalisation. La normalisation produit une ontologie régionale, constituée de concepts régionaux ou linguistiques, ou encore essences différentielles. La formalisation produit une ontologie formelle constituée de concepts formels. Or, il faut distinguer le signifié, le signifié normé ou concept matériel ou linguistique, et le concept formel. Seul ce dernier se manipule de manière combinatoire, compositionnelle, et récursive. Notamment, un concept formel appartenant à l'ontologie formelle est une primitive. C'est une brique élémentaire permettant de construire les autres. En revanche, le concept régional est une composante, il sert à expliciter d'autres concepts tout en étant lui-même soumis à l'explicitation. La normalisation fixe les règles interprétatives sans brider et figer l'interprétation : tout concept matériel peut être soumis à la sémiosis illimitée de l'interprétation tout en respectant la normalisation imposant les stabilités noématiques et relationnelles. Cela signifie en particulier qu'une essence régionale est un ensemble indéfini (ni fini, ni infini) de propriétés nécessaires et suffisantes : il faut et il suffit que les propriétés déterminant l'essence par opposition aux autres essences permettent de déterminer la place de l'essence dans l'ontologie ou système du domaine.

Du fait des principes de la sémantique différentielle, l'ontologie régionale est nécessairement soumise à la sémiose illimitée du sens, c'est-à-dire qu'il est toujours possible de poursuivre l'explicitation des notions représentées. En effet, par principe, tout nœud se définit en opposition aux autres nœuds. Il mobilise donc des traits différentiels pour se définir. Mais, puisque la sémantique différentielle est intralinguistique, les traits différentiels doivent être eux-mêmes des nœuds de l'ontologie. Puisque le nœud à définir se définit par rapport à tous les nœuds présents, il introduit des oppositions par rapport à tous ces nœuds. Il faut donc introduire des traits différentiels distincts de tous les nœuds présents, et donc, puisque les traits sont des nœuds, de nouveaux nœuds. Par conséquent expliciter totalement un nœud, c'est entreprendre une tâche infinie.

L'ontologie régionale, bien que conceptuelle, appartient au domaine du signifié linguistique. En tant que telle, elle bénéficie de la sémiose illimitée du sens, c'est-à-dire en pratique qu'il demeure toujours quelque chose d'implicite qui reste à expliciter. La langue et son système de signification recèle toujours de l'implicite, c'est-à-dire de l'explicitable, on ne pourra jamais tout expliciter. Passer de l'ontologie régionale à l'ontologie formelle, c'est précisément vouloir expliciter l'ontologie régionale dans un formalisme logique. Cette explicitation ne peut donc être que partielle.

Dans ces conditions, l'ontologie formelle ne peut être l'équivalent exact de l'ontologie régionale : nécessairement finie, l'ontologie formelle ne peut rendre compte de cette explicitation à l'infini et doit fixer des limites. Ces limites, elle les fixe en laissant implicite le contenu des nœuds puisqu'elle les utilise comme des primitives. Utilisant les nœuds pour définir de nouvelles notions composées, elle les utilise comme des primitives sans signification. Elle explicite alors les relations entre les nœuds pour rendre compte, au niveau formel, des engagements sémantiques de l'ontologie régionale.

Par conséquent, on retrouvera une relation taxinomique formelle « est-un » traduisant au niveau formelle la relation taxinomique sémantique « est-un » de l'ontologie régionale. A cet égard, on voit qu'au niveau formel, il est nécessaire de retrouver la structure arborescente des relations taxinomiques entre les notions du domaine : l'ontologie formelle se construit autou d'un squelette arborescent reflérant au niveau formel l'arbre des concepts de l'ontologie régionale.

Mais il ne s'agit que d'un squelette dans la mesure où au niveau formel, la sémantique est référentielle et permet à ce titre la composition des significations. Il est alors possible de composer les nœuds de l'ontologie formelle, considérés comme des primitives, pour constituer de nouveaux nœuds, appelés « types définis ». Ces nœuds ne possèdent pas de signification directe dans l'ontologie régionale. Leur signification doit s'effectuer deux étapes :

1. décrire selon les règles de composition sémantique stipulées par la sémantique réfé-

rentielle la signification du type défini en fonction de la signification référentielle des primitives à partir desquelles il est construit ;

2. utiliser l'ontologie régionale pour déterminer le sens des primitives et trouver leur référence objective ;

Ces étapes interprétatives sont valables pour toutes les représentations formelles construites dans le formalisme utilisant l'ontologie formelle comme signature. Autrement dit, pour savoir ce qu'une représentation formelle signifie, il est nécessaire d'une part de la décomposer en significations primitives grâce aux règles du formalisme et d'autre part de considérer l'interprétation linguistique des primitives dans le cadre de l'ontologie régionale pour déterminer leur signification objective. L'ontologie formelle n'a pas de sens par elle-même, mais uniquement à travers l'ontologie régionale qui lui donne son rapport au domaine, puisque c'est au niveau de l'ontologie régionale qu'est établie la détermination du sens en fonction de l'expérience possible dans le domaine.

6.2 Structure de l'ontologie formelle

L'ontologie formelle reprend les structures de l'ontologie régionale en les comprenant de manière formelle. Détaillons le statut de ces structures dans l'ontologie formelle en prenant comme formalisme celui des graphes conceptuels, retenu dans le projet Menelas:

la taxinomie différentielle des essences différentielles devient une *taxinomie formelle* de concepts formels ; la taxinomie formelle est graphiquement la même que la taxinomie différentielle, seule change l'interprétation que l'on en donne ; en particulier, les concepts correspondent à des classes d'individus et la taxinomie formelle s'interprète comme une hiérarchie d'inclusions de classes ;

la définition régionale devient une *définition formelle* ; en pratique, cela revient à traduire en graphes conceptuels la définition régionale et à réaliser complètement le principe de définition circulaire : chaque concept de la taxinomie se définit complètement à l'aide des autres concepts de la taxinomie formelle ;

l'encyclopédie régionale devient une *encyclopédie formelle* ; de même que pour la définition formelle, elle consiste en la traduction de l'encyclopédie régionale en graphes conceptuels écrits à l'aide de la taxinomie formelle.

L'ontologie formelle est donc la traduction dans un formalisme de représentation des connaissances, ici les graphes conceptuels (voir la figure 4.2), de l'ontologie régionale dont les structures sont transcrites en graphes conceptuels écrits à l'aide des labels de la taxinomie formelle, c'est-à-dire de la taxinomie différentielle considérée formellement.

La transcription en graphes conceptuels est une aide précieuse pour l'amélioration de la description en langue des structures régionales. En effet, elle force à expliciter dans les termes de la taxinomie ces structures : bien que cette explicitation ne puisse être complète pour les raisons vues plus haut, elle permet de contrôler la complétude de la taxinomie par rapport aux besoins de la définition des concepts. Mais, il est bien entendu que ces définitions formelles ne peuvent être qu'incomplètes, puisque les nœuds de l'ontologie formelle ne sont plus seulement des composantes, mais des primitives en nombre fini, et qu'elles ne peuvent dès lors prendre à leur charge la sémiose illimitée du sens.

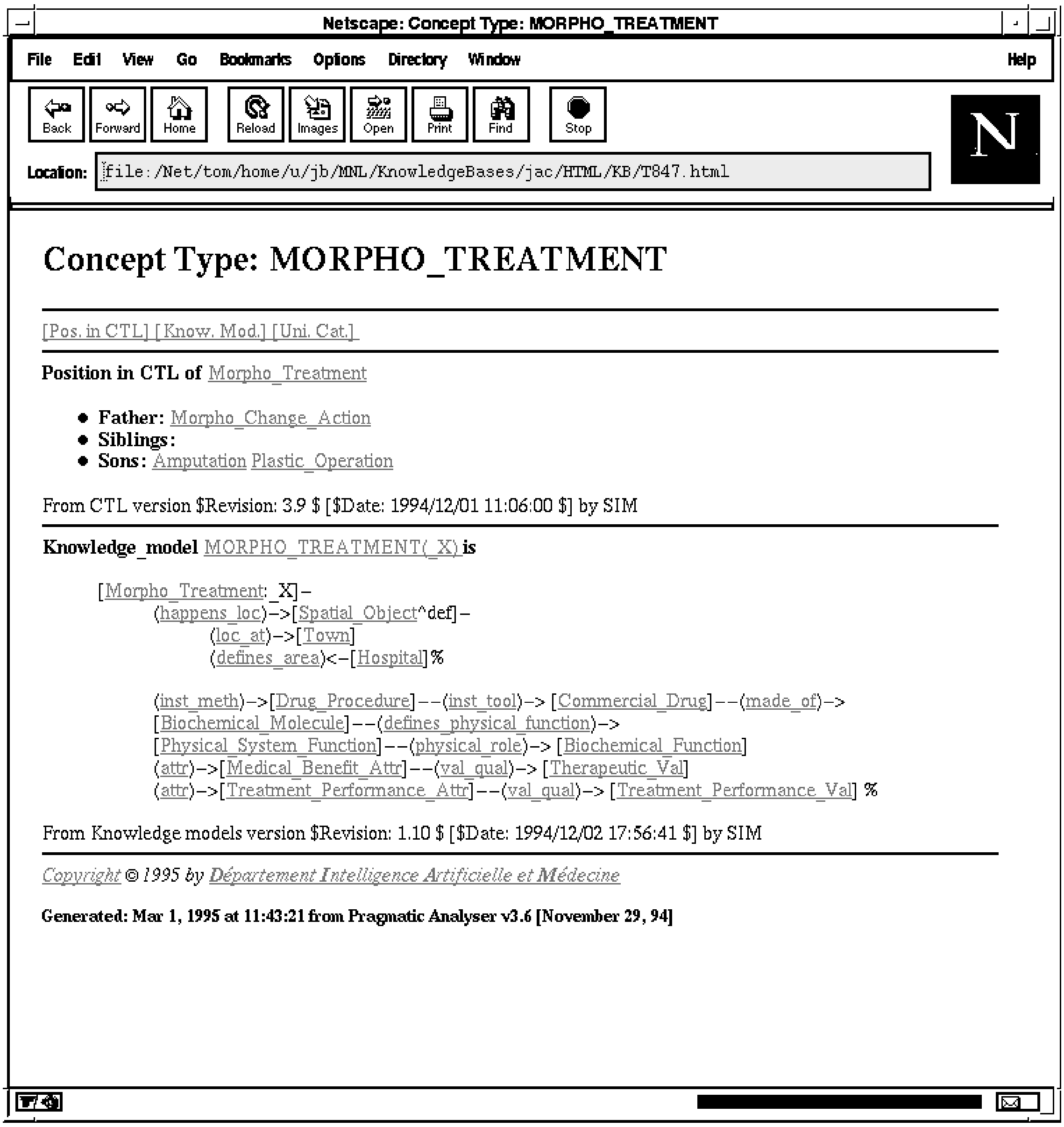


FIG. 4.2 – *Un extrait de l'ontologie formelle de* MENELAS

6.3 La sémantique des graphes conceptuels

L'explicitation en graphes correspond à une formalisation et à l'association d'une sémantique formelle aux nœuds de la taxinomie différentielle ainsi qu'aux autres structures régionales dans la mesure où les graphes possèdent une interprétation extensionnelle. Comme nous l'avons vu plus haut, la fonction Φ permet de traduire les graphes conceptuels en formules de la logique du premier ordre et ainsi de retrouver la sémantique extensionnelle qui les caractérise. Mais, à ce niveau plusieurs remarques doivent être faites.

En premier lieu, il convient de se rappeler que Sowa, le concepteur du formalisme des graphes conceptuels, voit dans les graphes une manière sinon équivalente mais à tout le moins proche de la langue naturelle d'exprimer les connaissances linguistiques. Le caractère graphique des graphes ainsi que la mobilisation du contenu linguistique attaché aux types des concepts font des graphes conceptuels des structures de représentation relativement aisées à lire. Mais, s'ils sont investis d'un contenu linguistique, les graphes n'en sont pas pour autant investis d'un contenu prescriptif, faisant de ces graphes une représentation de *connaissances* et pas seulement d'un contenu linguistique. Il est donc nécessaire d'encadrer l'interprétation des graphes à l'aide des principes différentiels et des structures régionales qu'ils impliquent pour faire des graphes des connaissances et non de simples expressions.

Il convient donc de lire les structures écrites en graphes conceptuels de l'ontologie formelle uniquement dans le cadre interprétatif des structures régionales pour appréhender leur signification. C'est d'ailleurs de cette manière qu'il est possible d'évaluer la complétude, la cohérence et la systématicité des structures régionales : c'est en lisant de manière régionale, c'est-à-dire différentielle les graphes conceptuels de l'ontologie formelle que l'on contrôle l'ontologie régionale.

Mais, si l'interprétation différentielle des graphes conceptuels donne le sens que *véhiculent* les graphes pour être des représentations de connaissance et pas seulement des expressions linguistiques, elle ne donne pas le sens qu'ils posséderont au sein du SBC qui les mettra en œuvre. Les graphes seront utilisés *computationnellement*, et la signification qu'ils posséderont dans ce contexte sera opérationnelle. Une manière de contrôler cette opérationnalisation est d'élaborer une sémantique dénotationnelle des graphes : intermédiaire entre le contenu prescriptif et le contenu opératoire des graphes, le contenu dénotationnel permet de contrôler le passage de l'un à l'autre, et de s'assurer qu'il existe bien un rapport de pertinence entre l'exploitation informatique des représentations symboliques et leur interprétation en termes de connaissances.

Il est donc nécessaire de s'assurer que la fonction Φ, fonction dotant les graphes d'une sémantique extensionnelle chez Sowa, assure correctement sa fonction. Or, rien n'est moins sûr. En effet, l'un des principaux attraits des graphes conceptuels est de permettre la représentation ou l'expression des aspects extensionnels et intensionnels des connaissances, en distinguant le type de la référence dans les concepts. Or, en interprétant les types comme des prédicats s'appliquant sur les référents, l'aspect intensionnel est perdu lors de la traduction logique. Il serait nécessaire d'interpréter les types des concepts non comme des prédicats, mais comme des types.

Même si, *in fine*, on se ramènerait toujours à des formules de la logique du premier ordre, la réforme de la fonction Φ permettrait d'aboutir à un *autre* ensemble de formules que sa version actuelle, qui traduirait plus fidèlement la sémantique différentielle et le contenu prescriptif des graphes lus à travers les structures régionales qu'ils expriment.

7 Ontologie computationnelle et opérationnalisation informatique

Dernière et ultime étape pour obtenir une taxinomie exploitable par un SBC, l'ontologie computationnelle. L'ontologie computationnelle consiste à doter l'ontologie formelle d'une sémantique opérationnelle permettant de considérer les structures taxinomiques comme des structures de programmation devant être exécutées.

Alors que la sémantique formelle décrit en termes extensionnels ce que fait le programme, la sémantique opérationnelle décrit comment il le fait. La sémantique formelle permet une formalisation au niveau des connaissances de l'ontologie régionale, alors que la sémantique opérationnelle en permet une formalisation au niveau des programmes.

L'opérationnalisation de l'ontologie renvoie à la problématique du contrôle en IA (Hayes-Roth, 1985; Bachimont, 1994). Ce dernier est le problème de savoir quand et comment utiliser une représentation structurée pour qu'elle se comporte comme une connaissance permettant la résolution d'un problème. Or le problème est bien à présent de savoir représenter et utiliser l'ontologie régionale pour que les structures de programmation obtenues se comportent en reflétant dans le programme le sens que possèdent dans le domaine les concepts contenus dans la taxinomie ontologique. La traduction des structures de l'ontologie formelle en structures de programmation dans l'ontologie computationnelle s'effectue ainsi :

la taxinomie formelle devient une *taxinomie computationnelle*, c'est-à-dire un arbre de généralisation et de subsomption permettant de comparer les graphes entre eux ; notamment il permet les opérations de jointure ;

la définition formelle devient une *définition computationnelle* permettant les opérations d'expansion et contraction de type ; un graphe correspondant à une définition est remplacé par le concept dont c'est la définition ;

l'encyclopédie formelle devient une *encyclopédie computationnelle* consistant en règle `Si Alors` permettant d'effectuer les inférences correspondant aux parcours interprétatifs correspondant à l'actualisation des afférences au niveau de l'ontologie régionale.

C'est donc au niveau de l'ontologie computationnelle qu'est abordé le problème d'une modélisation informatique des opérations interprétatives à la base du parcours interprétatif permettant de dégager la teneur conceptuelle et gnoséologique des expressions linguistiques. Prenons par exemple l'expression « sténose du patient » et essayons d'en dégager le contenu prescriptif.

Les lexies « sténose » et « patient » sont considérées comme n'étant pas seulement des unités linguistiques possédant la double face signifiant/signifié, mais comme des unités conceptuelles. Mais, pour que cela soit possible il faut être capable de traiter des problèmes qui se résolvent avantageusement au niveau purement linguistique.

Considérés comme des concepts, les lexies composant l'énoncé « la sténose du patient » permettent de construire le graphe conceptuel suivant :

```
[sténose]--(du)-->[patient]
```

Le problème est alors de construire la relation existant entre `[sténose]` et `[patient]`. Il faut en effet respecter les liens conceptuels unissant ces deux notions et sélectionner celui qui convient.

Si l'on considère la définition computationnelle de « sténose » et de « patient », on obtient pour chacun un graphe conceptuel spécifiant les propriétés correspondant aux sèmes inhérents de la notion. Il est inhérent à la notion de « sténose » d'être un état pathologique affectant la morphologie d'une artère. Ces traits sémantiques inhérents se retrouvent pris en compte dans l'ontologie computationnelle d'une part par la taxinomie computationnelle où « sténose » se trouve placée sous le concept « état morphologique » et d'autre part par la définition computationnelle contenant le graphe suivant :

```
 [stenose]- (state_of)-->[patient]
(consists_of)-->[artery_segment]--(part_of)-->[artery]
--(part_of)-->[patient]
```

On voit que cette définition contient le concept « patient » et donc qu'il existe une relation entre les concepts « patient » et « sténose ». Par conséquent, la relation qui doit être trouvée entre ces deux notions est

```
(consists_of)-->[artery_segment]--(part_of)-->[artery]--(part_of)-->
```

On constate que ce mécanisme d'exploration des définitions conceptuelles permet de trouver la relation conceptuelle existant entre un concept et un autre concept contenu dans sa définition. Or, la définition conceptuelle formalise et opérationnalise les traits sémantiques inhérents d'une notion. Par conséquent, ce mécanisme permet de trouver la relation existant entre d'une part un concept et d'autre part la conceptualisation (mise sous forme conceptuelle) de l'un des ses traits sémantiques. Ce mécanisme permet donc de traiter des *métonymies* (Bouaud *et al.*, 1996a; Bouaud *et al.*, 1996b).

Les métonymies se définissent en effet d'un point de vue sémantique comme la « relation entre un sémie et la lexicalisation d'un ou plusieurs de ses traits. » (Rastier *et al.*, 1994, p. 129). Ici, les relations ne s'établissent pas entre les sémies et leur lexicalisation, mais entre un concept et la conceptualisation de son contenu prescriptif.

Ce mécanisme d'exploration des définitions computationnelles peut être élargi aux encyclopédies computationnelles, où la relation trouvée exprime une métonymie entre un concept et la conceptualisation de ses traits afférents, alors que l'exploration des définitions computationnelles ne concernait que la conceptualisation des traits inhérents. Cela ne signifie pas que d'un point de vue linguistique, encore moins d'un point de vue cognitif, la nature des processus sous-jacents aux métonymies est de nature computationnelle. Cela signifie simplement que la modélisation effectuée à travers l'ontologie computationnelle permet d'appréhender un parcours interprétatif tout en respectant les prescriptions conceptuelles attachées aux notions définies non pas intralinguistiquement, c'est-à-dire d'un point de vue sémantique, mais définies prescriptivement, c'est-à-dire d'un point de vue gnoséologique.

Le mécanisme proposé ici est donc un mécanisme général d'*actualisation* des traits sémantiques. Il repose sur les phénomènes d'assimilation dans lesquels, en vue de la cohésion du sens que suppose la présomption d'isotopie, des sèmes sont actualisés pour créer la redondance et la récurrence nécessaire à la constitution du sens.

En revanche, dans les domaines spécialisés, il y a peu de phénomènes de virtualisation sémantique. En effet, ces derniers proviennent essentiellement du fait qu'il faut inhiber certains sèmes inhérents à un domaine dans la mesure où la lexie est utilisée dans un autre domaine. Par exemple, dans « Rocard monte au créneau », le sème /guerrier/ est virtualisé, le domaine étant politique et non militaire. Dans les domaines spécialisés où le recours imagé à d'autres domaines est limité, les sèmes inhérents n'ont pas à être virtualisés. Ils sont pleinement valables et doivent être actualisés pour la compréhension de la notion.

En particulier, tout phénomène de virtualisation sémantique peut être modélisé par le mécanisme ci-dessus d'exploration. En effet, ces virtualisations sémantiques sont suscitées par la préservation de la cohérence qui peut être rétablie en adoptant un lien conceptuel rendant compatibles les notions mises en cause. L'exploration des structures computationnelles permet de supprimer le problème dont la virtualisation sémantique a été la solution.

8 Normalisation, formalisation et appauvrissement

Tant la normalisation que la formalisation et l'opérationnalisation correspondent à un appauvrissement du domaine. Il ne faut pas confondre l'essence modélisée par l'ontologie régionale comme étant le reflet de la vérité suprême du domaine, qui serait vraie en toutes circonstances. L'essence différentielle reflète un point de vue normalisateur, qui en est la condition de possibilité. L'essence n'existe pas en soi dans le domaine : elle n'existe que dans le contexte interprétatif de la normalisation. Cela veut donc dire que l'essence est au plus un signifié normé. Le point de vue selon lequel le signifié, tel qu'il est constitué par l'usage en contexte et en discours, consiste dans l'essence agrémentée de variations accidentelles, contingentes, en un mot inessentielles, est faux. S'il en était autrement, il serait possible d'avoir une essence des termes hors contexte valable pour tous les contextes. Or, normaliser un corpus donné selon un point de vue et donc constituer des essences, peut être incompatible avec un autre point de vue normalisateur du même corpus. C'est donc que les différences sont essentielles, et non contingentes, et que ces différences essentielles permettent de construire des ontologies d'essences irréductibles et incommensurables.

Il n'y a donc pas une *signification* unique d'un terme ou d'une notion, valant pour tout contexte et par là-même universelle, recevant de sa contextualisation différente des *sens* différents rendant compte des accidents de la signification en contexte. Cela signifie que le conceptuel n'est pas la vérité du linguistique, mais son abstraction : les essences différentielles sont conceptuelles dans la mesure où elles figent le sémantique. Ainsi, « On peut appeler *conceptuelles* ces formations sémantiques : peu importe, si l'on convient que le conceptuel est du *sémantique allégé*. » (Rastier *et al.*, 1994, p.75)

Par ailleurs, on a vu comment l'ontologie formelle ne pouvait rendre compte de la richesse sémantique de l'ontologie régionale : la complète explicitation qu'elle implique entraîne la clôture du sens et sa stérilisation.

Enfin, l'ontologie computationnelle ne rend pas complètement compte de l'ontologie formelle dans le mesure où la sémantique extensionnelle n'est pas nécessairement constructive et ne donne pas toujours lieu à des opérationnalisations effectives. Les ontologies ne révèlent pas donc la vérité du domaine, mais sont incluses dans la pratique qui le constitue : elles sont un aspect du domaine, un procédé d'objectivation venant s'ajouter aux procédés déjà existants.

Une telle perspective implique qu'il faut reconsidérer le statut de l'ontologie formelle. En effet, dans les travaux de l'IA et selon la phénoménologie husserlienne, nous avons vu que l'ontologie formelle exprime les axiomes régissant l'usage par la pensée des catégories dites « formelles ». La pensée se caractérise par le fait de penser quelque chose, et les manières de le penser obéissent à des lois qui ne tiennent pas à ce qui est pensé, mais au fait qu'il est pensé. Ces lois sont donc nécessairement *a priori* dans la mesure où, elles précèdent et conditionnent toute pensée. Elles sont analytiques puisque, s'appliquant à tout ce qui est pensé, elles ne nous apprennent rien sur ce qui est pensé. L'ontologie formelle possède donc un statut d'universalité que l'on retrouve dans le fait qu'elle est la forme (logique ou pure) de région pour Husserl : elle est la forme logique de tout domaine possible.

Or, nous venons de voir dans ce chapitre que le formel n'est qu'une manière d'exprimer les connaissances pour se rapprocher d'un format exploitable informatiquement : la méthodologie que nous avons proposé ici consiste dans le fait qu'il faut à chaque étape changer de type de représentation, de format de représentation pour exprimer les connaissances. Or, constater que normaliser fige les connaissances dans un contexte particulier celui de la tâche, que normaliser n'est pas formaliser et qu'il faut pour cela laisser de l'implicite dans l'inexprimable, que formaliser ne donne pas obligatoirement de l'opérationnel, implique que changer de format de représentation n'est pas neutre quant à ce qui est représenté. Et, en particulier, qu'aucun format ne possède de primat dans la mesure où ses lois s'imposeraient nécessairement aux autres formats qui ne pourraient alors que les spécialiser ou les instancier.

Les travaux en IA traitant d'ontologie formelle suppose, en recourant à des sémantiques formelles extensionnelles, que le domaine est un ensemble d'objet. L'universalité putative de l'ontologie formelle repose donc sur cette hypothèse que nous avons vu ne pas tenir dans tous les cas. Cela signifie que, en modélisant le domaine comme nous l'avons fait pour obtenir une caractérisation du domaine en termes d'objets dans la mesure où la modélisation régionale permet de fixer la référence, on se ramène aux hypothèses du formalisme : autrement dit, on *traduit* les connaissances du domaine en représentations formelles. La formalisation n'est donc pas un travail consistant à dégager le savoir de sa gangue linguistique (cf. l'introduction), mais un travail de traduction. En termes naïfs, ce que permet la traduction, c'est de parler le langage des ordinateurs pour obtenir des modèles effectifs ; mais en parlant comme les ordinateurs, on ne parle plus comme le domaine.

Les parcours interprétatifs permettant la compréhension linguistique d'un énoncé sont entièrement conditionnés par la nature sémantique des termes de cet énoncé : la détermination de ces parcours relève du domaine et de ses normes pratiques. Cela implique qu'il n'existe pas de généralisation de ces parcours interprétatifs à un niveau formel, faisant abstraction des contingences matérielles. Le matériel est essentiel du point de vue du domaine, et le formel est essentiel du point de vue des outils utilisés pour opérationnaliser les connaissances du domaine.

Comme nous tenterons de le montrer dans la partie suivante, le formel est lui-même du matériel, c'est-à-dire qu'il renvoie à une région particulière possédant des principes synthétiques *a priori*. Le formel appartient à la région étudiée par l'informatique qui est une science expérimentale. Par conséquent, notre méthodologie consiste simplement à traduire la région du domaine où est posé le problème que doit résoudre le SBC à la région du computationnel où est posé la solution opérationnelle.

9 Généricité et universalité des ontologies

Ne véhiculant aucune vérité essentielle au sens de substantielle sur le domaine mais seulement des vérités pratiques vues depuis le point de vue d'une tâche en particulier, les ontologies régionales ne peuvent avoir nulle prétention à une quelconque universalité et réutilisabilité. Néanmoins, il convient de faire les remarques suivantes.

La normalisation sémantique permet d'obtenir un système conceptuel à partir de l'analyse linguistique du corpus. Ce système conceptuel est l'ontologie du domaine analysé et à ce titre mérite la dénomination d'*ontologie régionale*, en tant qu'ontologie d'une *région* partielle du réel. La définition de l'ontologie régionale du domaine peut être une fin en soi et indépendante de la volonté d'élaborer un SBC. La normalisation est donc une activité autonome qui pourrait constituer une discipline à part entière. Il nous semble que ce pourrait être précisément la tâche de la *terminologie* (Lerat, 1995; Felber, 1987), qui entend dégager les termes d'un domaine en leur associant un système notionnel. En effet, un terme se définit comme un couple (mot – concept), le mot étant un label normalisé en langue d'une notion conceptuelle. Mais, assimiler ainsi terminologie et construction d'ontologie régionale entraîne certaines conséquences.

La première conséquence est qu'une terminologie ne peut être valable en général pour un domaine, mais, à l'instar de l'ontologie, reflète un point de vue normalisateur défini en fonction d'une tâche à effectuer. L'appauvrissement lié à la normalisation est inévitable car il en est la condition de possibilité. C'est à la seule condition de se donner un point de vue de la tâche qu'il est possible d'élaborer un système conceptuel.

Si la terminologie veut refléter une vue des domaines indépendante des tâches à effectuer, il lui faudra pour cela rogner sur les critères de systématicité et de conceptualité adoptés pour définir le système notionnel sur lequel pointent les termes du langage de spécialité. En d'autres termes, la terminologie en restera à un niveau très proche de l'expression linguistique, sans pouvoir dominer la souplesse et la richesse d'expression de la langue par des régularités conceptuelles.

L'enjeu est donc de savoir dans quelle mesure la terminologie doit consister en une activité lexicographique, élaborant des *dictionnaires* plutôt que des théories conceptuelles structurant les objets du domaine. Tout dépend donc de ce pour quoi les terminologies sont faites. Deux points de vue semblent se dégager. Si les terminologies sont des notions purement descriptives, fonctionnant comme des observatoires des langues de spécialité, il n'est sans doute pas nécessaire d'adopter une perspective normalisatrice. Si en revanche la terminologie doit avoir une fonction normalisatrice et prescriptive, les critères de normalisation doivent refléter l'utilisation pratique et effective de ces termes, dans le domaine de spécialité: en d'autres termes ils doivent être déterminés en fonction des contextes d'utilisation, en fonction de la tâche.

Il ne faut sans doute pas exagérer la limitation imposée par cette contrainte de la tâche. En effet, les domaines concernés sont des domaines pratiques et non théoriques, qui sont donc structurés autour de tâches précises. La médecine se structure autour des tâches diagnostiques, de suivi, de prescription, de codification, etc. Par conséquent, si un domaine existe en tant que tel, c'est qu'il existe un ensemble déterminé de tâches qui le structure et les connaissances n'existent pas en dehors de ces tâches pratiques qui les mettent en œuvre. Une connaissance n'y est même qu'une possibilité d'agir dans l'accomplissement d'une tâche, qui a été systématisée et mémorisée en une expression sémiotique.

Pour finir sur l'indépendance douteuse des terminologies vis-à-vis de la tâche, il faut distinguer la tâche pratique définie dans le domaine de la tâche confiée au système informatique. La tâche du système est toujours définie à partir d'une tâche du domaine dans la mesure où le système, devant avoir une utilité et utilisation effectives, doit s'y intégrer.

La normalisation sémantique s'effectue à partir de la tâche du domaine, et non à partir de la tâche du système. Car c'est le cadre pratique de la tâche du domaine qui permet de conférer un contenu conceptuel aux expressions des connaissances, en les renvoyant à une expérience possible du domaine, à une dimension pratique effective. Il existe donc un niveau de généralité dominant la réalisation des systèmes informatiques dans la mesure où une terminologie ou une ontologie régionale sont valables pour toutes les tâches informatiques relevant d'une même tâche du domaine.

Mais cet optimisme doit être tempéré : la conception d'un SBC modifie la structure pratique d'un domaine car il modifie la manière de travailler des acteurs du domaine, des spécialistes. Les connaissances étant des possibilités d'agir, toute modification dans la manière de travailler se répercute sur le contenu des connaissances du domaine. Il faut donc envisager une normalisation des connaissances à partir, non pas de la tâche du système telle qu'elle est définie *a priori* par les concepteurs, mais à partir du système *construit*, *utilisé effectivement*, pour que les connaissances présidant à sa mise en œuvre puissent être déterminées.

Il y a donc un cycle de vie du système qui, conçu en fonction des connaissances du domaine, *i.e.* à partir d'une tâche du domaine (puisque toute connaissance est action au sein d'une tâche), constitue dès lors par son fonctionnement effectif de nouvelles tâches du domaine (et non du système). Par conséquent, le système devient lui-même la source de nouvelles connaissances pour les acteurs du domaine. Le cycle s'effectue donc entre les termes suivants : tâches du domaine, tâche d'un système définie à partir des tâches du domaine, nouvelles tâches du domaine définies à partir d'un nouvel acteur dans le domaine, le système.

Cela signifie qu'il est impossible d'établir une fois pour toute une terminologie dans la mesure où elle devra se faire constamment l'écho des nouvelles pratiques du domaine, qu'elles soient suscitées par l'introduction d'un SBC ou non. Cependant, à un instant donné, le niveau de généralité nécessaire à la conception d'une ontologie régionale ou terminologie est celui des tâches du domaine. Dans ces conditions, c'est une activité indépendante du fait qu'il est prévu d'élaborer un SBC ou non à partir de ces ontologies.

10 ETAPES MÉTHODOLOGIQUES POUR LA CONSTITUTION D'ONTOLOGIES

Il est utile à présent de rassembler les considérations précédentes en récapitulant les étapes méthodologiques nécessaires à l'élaboration d'une ontologie pour un SBC. Ces étapes sont les suivantes :

La constitution d'un corpus constitue la première étape. Il s'agit de réunir l'ensemble des traces linguistiques correspondant à des expressions possibles des connaissances linguistiques nécessaires à la réalisation d'une tâche donnée dans un domaine. Cette étape est fort délicate car elle exige de déterminer clairement la fonction du SBC à concevoir en fonction des tâches définies par la pratique du domaine. On peut dire que cette étape est proche de l'établissement du modèle d'organisation de KADS.

La modélisation linguistique du corpus constitue la seconde étape. Elle consiste dans l'analyse sémantique selon le paradigme différentiel des données linguistiques collectées dans le corpus. Il s'agit de repérer les notions mises en jeu et de définir les taxèmes du domaine permettant de recenser leurs sèmes inhérents et afférents.

La modélisation régionale constitue la troisième étape et résulte dans l'ensemble des structures constituant l'ontologie régionale. Les essences différentielles correspondent à la conceptualisation des caractérisations sémiques des notions élaborées à l'étape précédente. Alors que la modélisation linguistique élabore des contenus linguistiques ou signifiés, la modélisation régionale élabore des concepts régionaux ou essences différentielles pourvues d'un contenu prescriptif. Elle formule les engagements sémantiques liés au fait d'user une notion en explicitant ce que cet usage implique quant à l'usage des autres notions : l'engagement sémantique est intralinguistique et intranotionnel.

La modélisation formelle constitue la quatrième étape et résulte dans l'ensemble des structures constituant l'ontologie formelle. Elle consiste en théorie à fournir à l'ontologie régionale une sémantique extensionnelle et en pratique à traduire l'ontologie régionale en un formalisme de représentation des connaissances doté d'une telle sémantique. Dans notre cas, la traduction s'effectue en graphes conceptuels. Le principe de la traduction est que les structures sémantiques et conceptuelles soient entièrement explicitées en termes de concepts formels dont est désormais composée l'ontologie formelle, en lieu et place des essences différentielles. La modélisation formelle formule des engagements ontologiques explicitant en termes référentiels ce qu'implique l'usage d'un nœud de l'ontologie compris à ce niveau comme une primitive de formalisation.

La modélisation computationnelle constitue la cinquième étape et résulte dans l'ensemble des structures composant l'ontologie computationnelle. Elle consiste en théorie à fournir à l'ontologie formelle une sémantique opérationnelle, telle que ses structures logiques soient des structures de programmation, et en pratique elle consiste à doter d'un contrôle les composantes de l'ontologie formelle. La modélisation computationnelle formule des engagements computationnels explicitant en termes opérationnels ce qu'implique l'usage d'un nœud de l'ontologie compris comme un primitive de programmation.

L'ontologie formelle est l'équivalent du modèle du domaine de KADS. L'ontologie computationnelle correspond au modèle de conception, prêt à l'implémentation. Pour cette méthodologie, l'ontologie régionale telle que définie ici est inconnue du fait qu'elle suppose acquise l'existence des ontologies primitives : données avec le domaine, elles ne reposent pas sur une modélisation explicite.

Ces étapes méthodologiques sont assez proches des étapes dégagées par (Gaines *et al.*, 1993). Il dégage en effet trois types de structures de connaissances intervenant dans le processus d'acquisition des connaissances :

- une base informelle de connaissances (*informal knowledge base*), constituée par des procédures d'« élicitation » en utilisant des techniques conceptuelles d'induction, de raisonnement analogique, à base de cas, etc. ; cette base serait l'analogue de notre ontologie régionale dans la mesure où elle est informelle et conceptuelle ;
- une base formelle de connaissances (*formal knwoledge base*), constituée par des procédures de formalisation en utilisant des techniques de modélisation des connaissances ;

cette base serait l'analogue de notre ontologie formelle ;

– une base computationnelle de connaissances (*computational knowledge base*), constituée par des procédures d'opérationnalisation ; cette base serait l'analogue de notre ontologie computationnelle.

Néanmoins des différences doivent être soulignées dans la mesure où (Gaines *et al.*, 1993) considère que la base informelle renvoie à un modèle mental : notre ontologie régionale n'a en effet rien de psychologique ou de mental ; la base formelle serait également le modèle conceptuel du domaine : l'ontologie formelle n'est pour nous qu'une formalisation d'un modèle du domaine qui est précisément l'ontologie régionale informelle ; enfin la base computationnelle renvoie à un modèle « objectif » du domaine. D'une certains manière, l'ontologie computationnelle, du fait de son effectivité, produit de l'objectivité, c'est-à-dire qu'étant opérationnelle, c'est cette ontologie qui produit des connaissances dans le domaine.

11 Artéfacture et intelligence artificielle

La méthodologie proposée dans ce chapitre permet de construire dans le cadre d'une tâche donnée la théorie d'un domaine en exploitant l'expression linguistique des connaissances du domaine.

La description linguistique des expressions en langue des connaissances se traduit par un réseau d'identités et de différences entre les unités linguistiques du corpus. La normalisation sémantique permet d'expliciter selon le point de vue de la tâche un *système* d'identités et de différences particuliers et de le fixer. Pour cela, on met en œuvre des principes différentiels dont le respect oblige l'explicitation intralinguistique du contenu de sens visé par l'adoption de tel ou tel signifiant linguistique pour exprimer une connaissance ou un concept du domaine.

C'est donc cette étape de normalisation sémantique qui permet de dégager et de caractériser les connaissances du domaine qu'il faut mettre en œuvre dans le SBC. Cette étape, comme nous l'avons souligné, est constructive : le système obtenu, l'ontologie régionale ou modèle régionale du domaine, est construit ; il n'est pas immanent au domaine mais résulte de l'activité de modélisation nécessaire à la conception d'un SBC.

La normalisation sémantique ne met pas en œuvre de principes formels. Elle dégage ce qui caractérise le domaine en tant que tel dans le cadre d'une tâche fixée dans ce domaine. En d'autres termes, elle dégage les principes synthétiques *a priori* autour desquels vont se constituer les connaissances nécessaires à la réalisation du système. Ces principes sont synthétiques dans la mesure où ils reposent sur la tâche comme expérience possible, comme rapport au monde déterminant la valeur de connaissance des expressions linguistiques. Ils sont *a priori* dans la mesure où ils conditionnent la structure du SBC et ses actions. Ces principes explicitent les conditions de l'expérience possible dans le cadre du système.

Enfin, la normalisation sémantique permet de dégager des *primitives* de modélisation : en fixant un système d'identités et de différences, les signifiants linguistiques correspondant aux labels des nœuds du système prennent une valeur univoque et non contextuelle. Il est alors possible d'associer à ces signifiants des objets du domaine dans la mesure où le système prescrit comment les obtenir en mobilisant le rapport au monde proposé par la langue. En effet, c'est en considérant le sens associé dans le système de la langue au signifiant linguistique et en le soumettant aux prescriptions interprétatives du système que l'on est capable de déterminer les référents visés. En d'autres termes, le système est un système de fixation de la référence.

Par conséquent, la normalisation sémantique donne l'ontologie primitive qui, comme nous l'avons rappelé au début de ce chapitre, est supposée acquise par les travaux formels que nous avons consultés. Nous avons donc rempli le programme que nous nous étions fixé dans cette partie, à savoir d'ajouter l'étape méthodologique permettant d'appliquer les résultats méthodologiques concernant la constitution de méthodes génériques de résolution et d'ontologies du domaine.

Il s'agit à présent d'inscrire ce résultat dans le cadre du programme que nous avions fixé dans notre introduction. Il s'agit en effet de rendre compte du fait que les modèles de l'IA ne sont pas des modèles dénotationnels tout en étant des modèles du domaine. La méthodologie exposée ici rend pleinement compte de ce fait. En effet, on comprend facilement à présent que les modèles de l'IA formalisent un modèle régional et que ce n'est qu'à ce titre qu'ils modélisent le domaine. Autrement dit, le modèle formel dénote une caractérisation du domaine comme univers de référents construite grâce à la normalisation sémantique, c'est-à-dire grâce au modèle régional. Par conséquent, nous avons bien construit une méthodologie exploitant et théorisant le fait que les modèles de l'IA sont des modèles formels (car exprimés dans des langages formels de représentation) non dénotationnels (car modélisant le domaine non par ce qu'ils le dénotent, mais parce qu'ils formalisent un modèle régional et dénotent une image construite grâce à ce modèle).

Nous avons donc réussi à contrôler l'investissement de sens des signifiants linguistiques constituant la signature formel des langages de représentation des connaissances. Les principes différentiels sont les principes explicitant et contrôlant la signifiance linguistique.

Notre méthodologie rend par conséquent compte du premier aspect du constat positif que nous avions dressé dans notre introduction, à savoir que les sytèmes formels automatiques de l'IA sont spontanément interprétables par l'utilisateur du domaine dans la mesure où il reconnaît dans la signature fonctionnelle et relationnelle du système formel les signifiants linguistiques utilisés dans le domaine pour exprimer les connaissances. Notre méthodologie explicite et rationalise l'interprétation linguistique pour que le système formel automatique puisse fonctionner comme un modèle du domaine.

Mais, cette méthodologie ne rend pas compte de l'autre caractéristique fondamentale du constat positif que nous avions notée quant aux enseignements de la pratique effective des systèmes IA. C'est le fait que le système informatique, du fait de sa dynamicité, permet de constituer les connaissances du domaine. En d'autres termes, nous avons utilisé l'aspect automatique du système formel pour méthodologiser son interprétabilité, mais nous n'avons pas méthodologisé cet aspect dynamique.

C'est pourtant nécessaire dans la mesure où l'utilisation de l'outil informatique dans la constitution des ontologies régionales est fondamentale et doit être considérée pour comprendre le véritable statut épistémologique de ces ontologies.

Comment en effet constituer les essences différentielles ? On a vu que pour Husserl, ces structures exprimant les connaissances synthétiques *a priori* du domaine sont des structures mises en évidence par une opération de l'imagination qui effectue la variation éidétique. Mais, ici, ce travail de variation n'est pas effectué par l'imagination, mais repose sur une utilisation intensive de l'outil informatique, et ce à deux niveaux :

- d'une part, l'outil informatique permet d'exploiter le corpus qui constitue le matériau empirique à partir duquel dégager les connaissances du domaine ; il s'agit dans un premier temps de proposer des « candidats termes » qui sont en fait les signifiants linguistiques proposés pour être le représentant linguistique de l'essence ; il existe de nombreux travaux[6] autour de ce problème appelé habituellement « extraction de terminologie » ;

  Dans un second temps, il s'agit de donner les occurrences en contexte de ces candidats termes pour expliciter ou paraphraser intralinguistiquement le terme ; il est ainsi possible d'expliciter les traits sémantiques différentiels constituant la signification en corpus du terme ; c'est la problématique habituellement dénommée « extraction de collocations » et mobilisent également de nombreux travaux[7] ; cette étape complète l'extraction du candidat terme de la première étape par l'extraction de « candidats différentiels » ; ces deux étapes explicitent le donné empirique à partir duquel effectuer la normalisation sémantique ;

- à une second niveau, l'outil informatique intervient pour normaliser et fixer les candidats termes et différentiels en essences et différences ; l'outil permet d'opérationnaliser la logique du système différentiel ; en effectuant la sommation des traits différentiels pour déterminer le contenu d'un nœud de l'arbre ontologique, en réalisant des tests de cohérence (non formelle) permettant de vérifier que les principes différentiels sont explicités de manière pertinente et cohérente entre les différents nœuds, l'outil informatique permet de vérifier que le réseau ontologique fonctionne bien comme un système.

De manière essentielle, il faut remarquer que l'intervention de l'outil informatique permet de suppléer au travail parfois défaillant de l'imagination requis par Husserl pour dégager les structures synthétiques *a priori* du domaine. En d'autres termes, le synthétique *a priori* du domaine n'est pas une structure *a priori* de la raison, mais une construction de la raison reposant sur la médiation technique de l'informatique.

En effet, l'outil permet de constituer des catégories du domaine que l'on serait incapable de dégager sinon. Notre méthodologie repose en effet sur le fait que la détermination du système différentiel s'effectue *localement*, et que l'outil informatique permet de passer au global. En d'autres termes, on laisse la rationalité s'exercer à un niveau où elle peut contrôler le sens des éléments en présence : un nœud confronté à son père et ses frères. Par ailleurs, la structure arborescente permet de passer du local au global en sommant sur les chemins, d'un nœud à la racine, entre deux nœuds, les traits différentiels explicités. Ce passage est nécessaire dans la mesure où c'est la position globale du nœud dans le réseau qui permet de déterminer sa signification et d'expliciter le contenu de l'essence différentielle dont il est le représentant taxinomique.

6. On peut consulter à ce propos (Bourigault, 1994), (Mikheev & Finch, 1995), (Daille, 1995), (Soderland *et al.*, 1995).

7. On peut consulter à ce propos (Grefenstette, 1994), (Habert & Jacquemin, 1993; Habert *et al.*, 1995; Habert & Nazarenko, 1996). *Stricto sensu*, ces travaux se recoupent étroitement avec les travaux concernant l'extraction de terminologie et il est souvent impossible d'entreprendre l'un sans l'autre. La distinction opère davantage sur l'objectif (extraire des termes ou extraire des contextes) que sur les méthodes.

Sans le recours informatique, l'ontologie régionale resterait virtuelle. En effet, pour des raisons de complexité calculatoire évidentes, les opérations effectuées pour définir le contenu d'une notion dépasse les possibilités mentales. Il est nécessaire de recourir à l'outil informatique auquel on délègue la mise en œuvre effective de la logique du système. En pratique, nous avons conçu une station de lecture / écriture pour les ontologies qui assiste la modélisation régionale. Elle permet en effet de constituer l'ontologie régionale en utilisant le support informatique. Ce support permet d'objectiver des inférences formelles dont l'interprétation sémantique en termes de connaissances du domaine aide à les systématiser. Le paradoxe est donc que le travail méthodologique que nous avons exposé dans ce chapitre, tout en voulant se démarquer du formalisme pour élaborer un modèle non formel du domaine, repose en grande partie pour sa possibilité pratique sur l'outil informatique. La réalisation d'une ontologie de 2000 concepts documentés, effectuée à l'aide de cette méthodologie, n'aurait pas été possible sans recourir au support informatique.

En d'autres termes, l'imagination se trouve dépassée pour constituer les essences par variation éidétique à deux niveaux : d'une part la détermination des voisins en corpus pour déterminer les candidats termes et les candidats différentiels, d'autre part le passage du local au global en sommant les résultats de la normalisation locale effectuée sur les candidats trouvés. Puisque l'imagination est finie et bornée dans ses capacités, seule l'informatique permet de dépasser ces limitations et de dégager les essences, c'est-à-dire de *constituer les connaissances synthétiques* a priori *du domaine*.

Le synthétique *a priori* n'est donc pas une caractéristique de la raison pure, ou une structure *a priori* de la raison. Il s'agit plutôt d'une structure *construite* par la raison grâce à une médiation technique. Les essences différentielles ne sont donc pas des structures *a priori* puisqu'elles sont construites. Mais elles ne sont pas non plus empiriques : elles n'appartiennent pas au domaine, elles ne représentent pas des faits empiriques du domaine. Les essences différentielles donnent les catégories de l'expérience dans le domaine, mais ces catégories ne sont dégagées que par la médiation de l'outil technique.

Ni empiriques, ni transcendantales, les essences différentielles fonctionnent comme des catégories structurant l'expérience possible dans le domaine, sans ressortir exclusivement aux structures de l'esprit. Pour dégager ces catégories, l'esprit doit s'appuyer sur la médiation technique : autrement dit, le transcendantal comme condition *a priori* de l'expérience est constitué, et sa constitution est technique. La technique permet de modifier et de constituer nos catégories transcendantales, c'est-à-dire nos catégories de l'expérience.

Les questions qui se posent à présent sont alors les suivantes. En considérant que l'informatique est une technique permettant de constituer des connaissances, on suppose par là même que l'informatique permet d'accéder à du nouveau. En d'autres termes, on suppose que l'informatique est une science expérimentale. Par ailleurs, on considère que cette technique informatique permet de suppléer aux défaillances cognitives de la raison et de lui fournir la possibilité de se constituer un domaine d'expériences possibles. Notre partie suivante entreprend une réflexion sur ces deux problèmes :

- quel type de science expérimentale est l'informatique, et de quel type d'objet est-elle la connaissance ; c'est le propos du chapitre 5 de travailler cette question ;
- que signifie la médiation technique dans le travail de la raison ? quel est le rôle de la technique dans la constitution du sens et des connaissances ? C'est le propos des chapitres 6 et 7 d'explorer ce problème en recourant à la phénoménologie husserlienne.

En d'autres termes, il nous faut, après avoir vu une mise en œuvre partielle de l'artéfacture, de comprendre ce sur quoi elle repose, ce que nous appelons l'herméneutique matérielle.

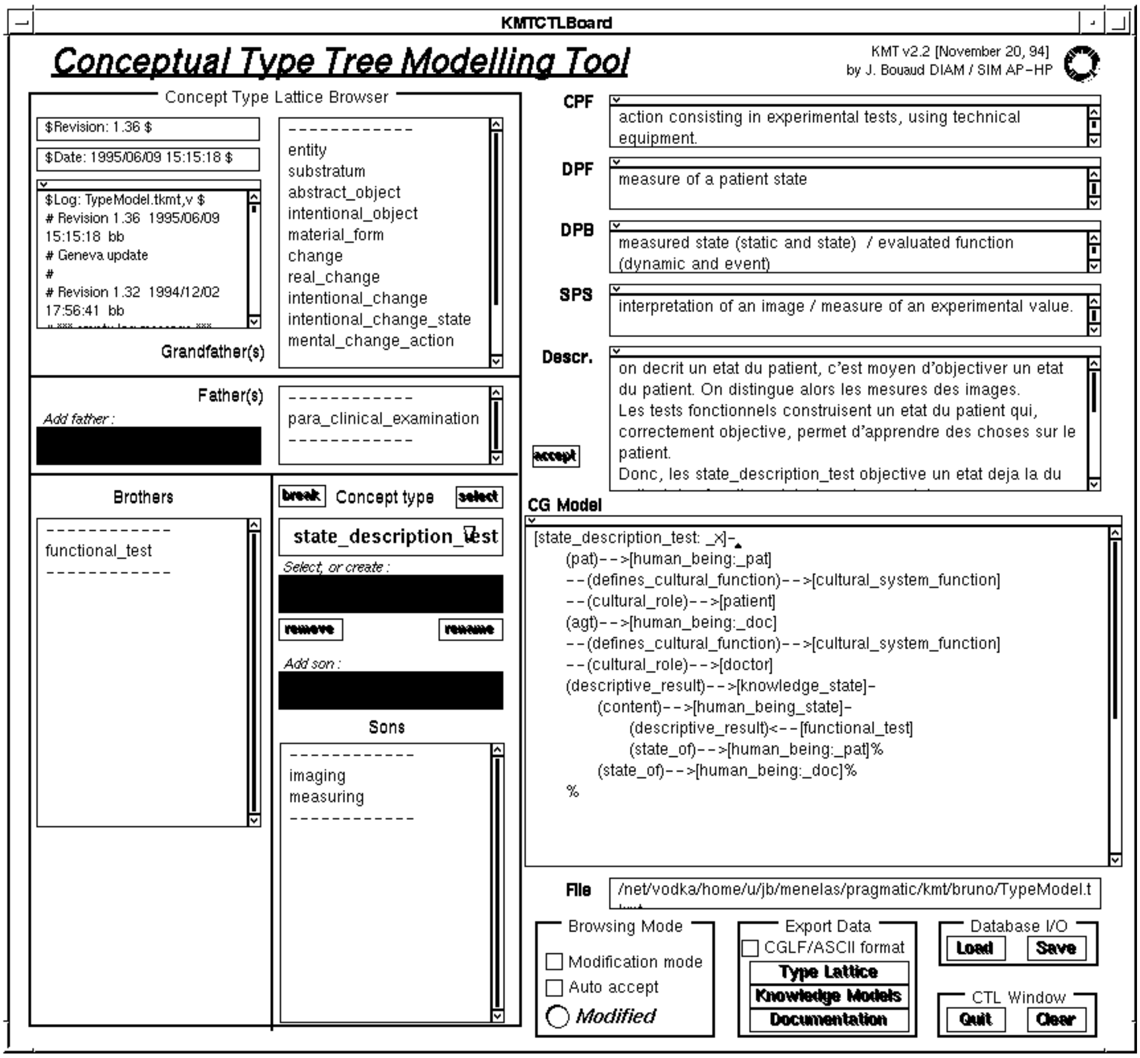


FIG. 4.3 –*Ceci est la copie d'écran de la description du contenu local d'un nœud tel qu'on peut le lire sur la station de lecture / écriture pour ontologie que nous avons développée. En haut à droite, on lit les 4 principes différentiels : CPF renvoie au principe de communauté avec le père, DPF renvoie au principe de différence avec le père, DPB au principe de différence avec les frères, et SPB au principe du communauté des fils entre eux. On voit au milieu à droite un commentaire en langue naturelle de la notion associée au concept défini. En bas à droite, on voit la traduction formelle du contenu local du nœud. En haut à gauche, on voit le chemin unissant le nœud à la racine, (unique du fait de la structure arborescente). En dessous est donné le père du concept défini. Enfin, en bas à gauche on trouve les frères et les fils.*

Deuxième Partie

Artéfacture et herméneutique matérielle

Le problème de la constitution informatique des connaissances

Chapitre V

Sens et épistémologie transcendantale

L'IA, selon la caractérisation que nous avons retenue, construit des modèles pour opérationnaliser les connaissances permettant de résoudre ou traiter un problème posé dans un domaine de spécialité. Nous avions remarqué dans notre introduction que les modèles ainsi conçus, bien qu'exprimés dans des formalismes logiques de représentation des connaissances, n'étaient pas des modèles du domaine dans la mesure où ils le dénotaient. Les modèles de l'IA sont des modèles du domaine dans la mesure où ils mobilisent dans leurs règles inférentielles des signifiants symboliques empruntés à la langue spécialisée du domaine permettant d'exprimer les connaissances caractérisant le problème posé. Les structures symboliques du formalisme sont dès lors interprétables dans les termes du domaine et renvoient ainsi à la réalité objective du domaine. Mais, pour que ces renvois fonctionnent de la manière dont un modèle signifie ce qu'il modélise, il convient de contrôler l'investissement de sens linguistique des signifiants symboliques. Nous avons montré dans la première partie comment effectuer un tel contrôle en l'illustrant sur la problématique des ontologies.

La méthodologie que nous avons présentée dans la partie précédente nous a permis de voir que contrôler l'investissement sémantique, c'est considérer l'interprétation linguistique dans son rapport à une expérience possible, à une action ou pratique dans le domaine où le signifiant symbolique non seulement signifie un sens du domaine, mais une connaissance dans la mesure où il prescrit une action possible.

Cet investissement de sens en fonction de la pratique dans le domaine repose sur la manière dont les primitives symboliques du formalisme sont elles-mêmes pourvues de sens. C'est la signification de ces primitives qui détermine ce que signifient en termes de connaissances du domaine les représentations formellement construites par leur moyen. Autrement dit, la signification des primitives détermine *a priori* la valeur gnoséologique des représentations formelles : elle constitue leur cadre *a priori* au sein duquel ces représentations renvoient à une expérience possible dans le domaine. Du fait que les représentations formelles doivent constituer un modèle du domaine, c'est-à-dire rendre compte de notre connaissance du domaine, il résulte que les primitives du formalisme doivent renvoyer aux connaissances nécessairement vraies de manière *a priori* dans le domaine pour que toute représentation possède par principe de construction une valeur de connaissance dans le domaine. Les primitives doivent ainsi constituer les principes synthétiques *a priori* du domaine.

Puisque la détermination de telles primitives est la tâche que s'assigne la problématique des ontologies en IA, il ressort que la détermination des primitives comme connaissances synthétiques *a priori* du domaine renoue avec le projet husserlien des ontologies matérielles ou régionales. Notre méthodologie permet ainsi de construire l'ontologie régionale d'un domaine, faisant ainsi suite aux travaux permettant de rapporter le domaine aux structures analytiques *a priori* de l'ontologie formelle.

Nous avons alors constaté que l'analyse transcendantale du domaine, puisque c'est de cela dont il s'agit (cf. chapitre 3), c'est-à-dire la caractérisation de la logique transcendantale du domaine, repose sur une médiation technique, celle de l'informatique. La question qui s'est alors posée consistait à s'interroger sur le statut de cette médiation technique : contingente ou nécessaire. En effet, la médiation informatique n'est-elle qu'un aspect contingent de l'analyse transcendantale, c'est-à-dire que l'on aurait pu l'effectuer sans elle ? Ou bien la médiation informatique est-elle nécessaire dans la mesure où seule son intervention permet de structurer les catégories de l'expérience possible dans le domaine ? Autrement dit, la médiation informatique conditionne-t-elle notre expérience du domaine ou bien est-elle un outil en lui-même neutre de notre expérience ?

L'objectif de cette partie est de traiter cette question et d'aborder son extension au problème de la technique. Le traitement que nous proposons s'effectue selon les étapes suivantes :

- dans un premier temps, nous tentons d'éclaircir le rôle de l'informatique en IA en entreprenant une analyse transcendentale de cette discipline ; nous arrivons à la conclusion que l'informatique est une science expérimentale, donc à ce titre une science de la nature, et qu'elle a pour objet le calcul, c'est-à-dire la détermination temporelle du devenir en algorithme ; cela nous permet d'une part de comprendre que l'informatique et les constructions qu'elle autorise ne possède aucune capacité particulière à représenter le domaine, si bien que si les modèles conçus en IA modélisent un domaine, ce n'est pas parce qu'ils sont informatiques : cela permet de mieux comprendre notre constat négatif ; par ailleurs, cela permet également de mieux appréhender le rôle médiateur de l'informatique dans l'élaboration des connaissances dans la mesure où, science du calcul, on peut lui déléguer l'effectivité de la représentation symbolique de nos connaissances pour expliciter des contenus qui sinon n'auraient jamais pu l'être ;

- dans un second temps, nous entreprenons d'approfondir ce rapport entre l'informatique et la connaissance en mobilisant la phénoménologie husserlienne, qui fournit une théorie descriptive des connaissances ; cette dernière permet en effet de se poser la question si le rôle médiateur de l'informatique dans l'élaboration des connaissances ne provient pas tout simplement du fait que la pensée est structurée formellement ou computationnellement, et donc la pensée peut se reconnaître et s'assister dans l'ordinateur dans la mesure où elle est d'une certaine manière un ordinateur ; mais, en analysant cette question, nous aboutissons à la conclusion que la pensée n'est pas structurée computationnellement mais qu'elle repose en sa possibilité même sur la médiation externe du signe ; autrement dit, l'élaboration des connaissances doit en passer par la médiation de support de connaissances externes ; on comprend alors mieux le rôle de l'informatique qui, comme physique des signes, propose le calcul comme support de connaissances et participe de manière constitutive et pas seulement contingente à l'élaboration des connaissances ;

- dans un troisième temps, nous reprenons ce constat fait quant à l'informatique pour l'élargir à la notion de support externe de connaissance en la considérant à travers le concept d'écriture.

En d'autres termes, en constatant que la médiation informatique est constitutive des connaissances du domaine et des principes synthétiques *a priori* qui le structurent, nous avançons l'hypothèse que la médiation technique constitue l'expérience possible dans le domaine. Cela renvoie au fait bien connu que, par exemple, l'introduction de nouvelles techniques modifient considérablement les méthodes de travail du domaine jusqu'à modifier également la nature même du domaine. Nous y revenons dans le dernier chapitre.

Ce chapitre est donc placé sous le signe de l'épistémologie transcendantale. L'épistémologie transcendantale est un courant important de l'épistémologie poursuivie aujourd'hui par Jean Petitot (1986; 1992a) qui l'utilise, à la suite de Cavaillès et de Lautman (ce dernier étant davantage platonicien), pour ébaucher une épistémologie de la physique, ce qui est habituel en épistémologie, mais surtout, une épistémologie des mathématiques et des sciences cognitives. La filiation transcendantale possède également une dimension herméneutique avec Jean-Michel Salanskis (1991) qui exploite le legs esthétique kantien à l'aune des problématiques herméneutiques de Heidegger et surtout de Gadamer. Il aboutit à la conclusion que les sciences mathématiques ressortissent au questionnement herméneutique tout autant que les sciences humaines, alors que ce type d'interrogation leur semblait jusque là réservé.

En adoptant une perspective transcendantale, nous nous inscrivons dans cette continuité dans la mesure où nous pensons que la philosophie transcendantale donne les instruments nécessaires à la compréhension du rapport à l'expérience qui est mobilisé dans le cadre des sciences de la nature. Le plan adopté pour la suite chapitre est le suivant :

- nous exposons les éléments de la philosophie critique kantienne qui sont utiles à notre propos (section 5-1) ;
- de cette présentation, nous tirons deux questions qui nous intéressent plus particulièrement : d'une part les conditions de possibilité d'une psychologie scientifique (section 5-2), et d'autre part le sens transcendantal de la connaissance logique et informatique (section 5-3);
- la considération des mathématiques et de la logique nous amène à élaborer une épistémologie transcendantale de la logique et de l'informatique, à l'instar des rapports entre mathématiques et physique mathématique dans l'édifice kantien ; en particulier, de la même manière que Jean Petitot a montré que les mathématiques constituaient un discours d'objectivation, nous le montrons pour la logique ; et de la même manière que les mathématiques sont la condition d'objectivation des objets physiques, nous montrons que la logique est la condition d'objectivation des objets informatiques;
- cependant, notre conception se heurte à une objection fondamentale concernant la possibilité de considérer les discours logiques et mathématiques comme des discours d'objectivation ; cette objection consiste à remarquer que toute objectivation ne peut se construire qu'à partir d'une réceptivité passive ou esthétique ; or, tant les mathématiques que la logique reposent sur la pure spontanéité du sujet, n'étant soumises à rien d'empirique (section 5-4);

- pour répondre à cette objection, nous reprenons l'approche transcendantale pour critiquer la notion de jugement synthétique *a priori*. Nous nous inspirons pour cela de la philosophie des mathématiques de C.S. Peirce, et concluons sur le fait que, tout comme dans l'expérience physique empirique qui repose sur la passivité de la perception, la logique et les mathématiques reposent sur l'expérience de la pensée, dans le sens où le sujet constate passivement qu'il y a de la pensée en lui. Nous sortons alors du cadre kantien qui fut jusqu'alors le nôtre pour élaborer une théorie de la connaissance reposant sur la médiation constitutive du signe externe (section 5-5);
- pour conclure, nous réévoquons les raisons pour lesquelles nous avons prôné une passivité de la pensée et de la raison et pas seulement de la perception. Ces arguments nous entraînent inéluctablement à remettre en cause l'opposition de l'espace et du temps, et que toute notion de passivité et spontanéité repose sur une dichotomie artificielle bien que traditionnelle de l'espace et du temps (section 5-6).

1 La théorie kantienne de la connaissance

Il n'est pas question pour nous de nous faire l'exégète de la pensée kantienne [1], notre propos n'étant pas celui d'un historien de la philosophie. Comme nous l'avons déjà expliqué dans notre introduction générale, notre réflexion veut puiser dans le fonds de la tradition ce dont elle a besoin, mais elle ne vise pas à cerner l'essence du kantisme, ni celle de la phénoménologie husserlienne dans le chapitre suivant. Notre démarche est guidée par notre problématique scientifique et philosophique qui, pour entrer en résonnance avec l'héritage philosophique, ne saurait s'y réduire. C'est la démarche d'une épistémologie expérimentale où nous recherchons une construction théorique nous permettant de comprendre et de traiter notre problématique, et d'évaluer cette construction à l'aune du succès pratique [2] qu'elle implique.

La présentation partielle effectuée ici est principalement orientée sur les conditions de possibilité de la psychologie scientifique d'une part et sur la caractérisation transcendantale de la logique et des mathématiques d'autre part. Comme cela a été souligné plus haut, la discussion centrale pour une épistémologie de l'IA porte en effet sur la pertinence de la modélisation logico-computationnelle en IA. Ces préliminaires étant entendus, nous pouvons revenir à notre exposition partielle et orientée du contenu de la critique kantienne.

La philosophie kantienne est une réaction au « sommeil dogmatique » où était plongé

2. « Pratique » plutôt qu'« expérimental », dans la mesure où l'on ne peut affirmer qu'utiliser une base de connaissances, ou une ontologie dans le sens de la première partie, corresponde à ce protocole normé qu'on appelle expérience scientifique. Cette utilisation est en effet beaucoup plus floue et empirique, d'où notre appellation de « pratique ».

2. Sur un auteur aussi important que Kant, la bibliographie ne peut immanquablement qu'être immense. Quelques études remarquables se détachent cependant comme (Philonenko, 1968) et (Rivelaygue, 1992). (Pascal, 1957; Crampe-Casnabet, 1989) sont plus élémentaires. (Deleuze, 1963) aborde la doctrine des facultés et propose une vue synthétique du kantisme, articulé en ses moments critiques. Pour une étude phénoménologique et ontologique de la *Critique de la raison pure*, les trois œuvres d'Heidegger (1982; 1953; 1971) sur Kant s'imposent. Sur les relations de la science à la philosophie, (Vuillemin, 1955) propose une étude complète et détaillée.

le monde philosophique, assimilant l'héritage leibnizien[3]. Mais, le chemin est long du dogmatisme au criticisme. A la suite de Descartes et Leibniz, Kant propose en effet dans sa *Dissertation* de 1770 une vision encore toute gouvernée par la métaphysique (donc dogmatique) bien que déjà grosse de nombreuses thèses de la critique. Kant y affirmait en effet qu'il existait une science des noumènes, *i.e.* des choses telles qu'elles sont, et une science des phénomènes, *i.e.* des choses telles qu'elles apparaissent. Les noumènes sont les objets intelligibles donnés à et pensés par l'entendement, qui les atteint dans leur vérité ontologique. Les phénomènes sont les objets tels que les sens nous les donnent, *i.e.* à travers une perception qui nous les situent d'office dans un espace et un temps.

La science des noumènes est une métaphysique qui connaît les choses sans l'expérience physique grâce aux idées claires et distinctes. On passe du concept à l'objet indépendamment de l'expérience concrète. Elle étudie les noumènes, les choses intelligibles, en utilisant les principes de l'entendement, à savoir les principes d'identité et de non contradiction. Sa matière d'étude est donc le noumène, et sa forme consiste dans ces principes (les noumènes se révèlent à travers les principes de l'entendement qui manipulent leurs idées intelligibles). La science des noumènes est la science associée à l'entendement. Au contraire, la science des phénomènes est la science associée à la sensibilité. C'est une physique qui connaît les choses à travers l'expérience sensible concrète. Sa matière est la sensation, et sa forme est l'espace et le temps (les phénomènes se révèlent à travers un temps et un espace).

La sensibilité est la source de jugements synthétiques[4] issus de l'expérience : ils sont empiriques ou *a posteriori*. L'entendement, conceptuel et spontané, est la source de jugements analytiques qui peuvent correspondre à un usage réel ou logique. L'entendement possède un usage réel à travers lequel il connaît les noumènes : l'usage de ses principes est réel dans la mesure où les conclusions tirées indépendamment de l'expérience reflètent *réellement* la nature des choses pensées. Cet usage est analytique et *a priori* : indépendant de l'expérience sensible, il explicite le contenu pensé des concepts selon les principes logiques de l'entendement. Les seules forces de l'entendement lui permettent de donner un objet, un contenu, aux concepts pensés. Par ailleurs, l'usage logique de l'entendement l'articule à la sensibilité : les principes de l'entendement permettent de classer, ordonner, catégoriser logiquement les phénomènes. L'entendement n'ajoute rien à l'expérience phénoménale, mais il l'organise selon ses critères logiques. L'usage logique est donc un usage analytique (il explicite sans rien ajouter à la connaissance).

Mais l'usage réel est métaphysique et ne résiste pas à la critique de Hume : la nécessité logique de l'ordre des pensées ne possède aucun titre pour être valable pour l'ordre des choses. Il est impossible de passer du concept d'un objet à l'objet lui-même : la pensée de *Pégase* ne suffit pas à poser son existence. La pensée de l'objet, son concept, ne suffit pas à m'en donner une connaissance. Il faut que cette pensée s'articule à l'expérience possible de l'objet pour qu'il y ait connaissance de l'objet à travers son concept.

3. Kant avoue dans les prolégomènes : « J'en conviens franchement : l'avertissement de Hume fut précisément ce qui, voilà bien des années, vint interrompre mon sommeil dogmatique, et donna une tout autre orientation à mes recherches dans le domaine de la philosophie spéculative. » *in* (Kant, 1965, p. 18).

4. « Les jugements analytiques ne nous apprennent proprement rien de plus sur l'objet que ce qu'en contient déjà le concept que nous en avons, parce qu'ils n'étendent pas la connaissance au-delà du concept du sujet, mais qu'ils ne font qu'éclaircir ce concept. » *in* (Kant, 1984b, p. 506). Les jugements synthétiques « ajoutent au concept du sujet un prédicat qui n'avait pas été pensé en lui et qu'on n'aurait pu en tirer par aucun démembrement. » *in* (Kant, 1984b, p. 37). Un jugement est *a posteriori* s'il a sa source dans l'expérience, *a priori* sinon.

La science des noumènes est donc impossible et il ne peut y avoir de science que de science des phénomènes. Ainsi, il ne peut plus y avoir comme jugements portant sur la nature des choses que des jugements *a posteriori*, puisque seuls les phénomènes peuvent donner des enseignements sur la nature. L'usage réel de l'entendement est écarté et seul son usage logique subsiste, il ne peut y avoir comme jugements *a priori* que des jugements analytiques classant et catégorisant les apparences sensibles.

Mais cela entraîne qu'une science est impossible : d'une part il ne saurait y avoir de la nécessité posée expérimentalement, *i.e. a posteriori*: « L'expérience nous apprend bien ce qui est, mais non ce qui ne puisse pas être autrement. Aussi les arguments empiriques ne peuvent-ils fournir aucune preuve apodictique. » *in* (Kant, 1984b, p. 505) ; d'autre part, la pensée et la logique, nécessaires quant à elles, ne donnent aucune connaissance. Par conséquent, s'il est impossible d'augmenter ses connaissances avec nécessité, il est impossible d'avoir une science.

Cette impossibilité repose la question des jugements synthétiques *a priori*: en effet, il est nécessaire de sortir du dilemme de la nécessité logique sans portée empirique et de la connaissance empirique sans nécessité logique. Or, puisque la nécessité ne peut provenir de l'expérience, elle doit venir du sujet et donc être *a priori*. Par ailleurs, le jugement, pour être informant, doit procéder de l'expérience, sous peine de n'expliciter que ce qui est pensé sous un concept, sans qu'il y ait véritablement d'*objet* pensé. Le jugement doit donc être synthétique. Si bien que la possibilité de la science repose sur la possibilité des jugements synthétiques *a priori*.

Or, la science existe, à témoin la science newtonnienne, grande réussite du temps que s'efforce de penser philosophiquement Kant. Force est de constater qu'il existe des jugements valables *a priori*, *i.e.* antérieurement à toute expérience, et synthétiques, *i.e.* qui ajoutent au sujet du jugement quelque chose qui n'était pas compris en lui. Par exemple, les jugements mathématiques sont synthétiques *a priori* : que la somme des angles d'un triangle soit égale à deux droits est un jugement qui ajoute quelque chose à ma connaissance du triangle, et qui est effectuable sans avoir besoin de mener des expériences dans le monde. Par ailleurs, certains jugements physiques sont synthétiques *a priori*: que tout événement a une cause n'est pas un enseignement de la physique, mais une condition nécessaire à la constitution d'une physique comme telle. Il faut donc comprendre philosophiquement comment de tels jugements sont possibles, puisqu'ils existent, et donc comprendre comment la science est possible :

> « Puisque ces sciences sont réellement données, il est convenable de se demander COMMENT elles sont possibles ; qu'elles doivent être possibles, c'est démontré par leur réalité. » *in* (Kant, 1984b, p. 44)

C'est la question inaugurale de la *Critique de la Raison Pure* (Kant, 1984b). La *Critique de la Raison Pure* est le *MeisterWerk* de Kant, et peut-être même celui de toute la tradition rationaliste. La problématique kantienne s'insère dans le cadre théorique ouvert par Descartes de la connaissance comme représentation. L'entreprise kantienne peut être comprise comme une tentative d'étudier les différents types représentations et de définir à quelles conditions elles correspondent à de véritables connaissances scientifiques.

1.1 Sensibilité et entendement

La raison comprend plusieurs sources de ses représentations : l'intuition qui lui délivre des représentations singulières, l'entendement qui établit des représentations générales [5], la raison qui élabore des raisonnements. L'intuition est une réceptivité : les représentations nous y sont données et se rapportent immédiatement à l'objet. L'entendement est une spontanéité : les représentations singulières et générales sont spontanément liées pour former de nouvelles connaissances. Le fait de lier des représentations (singulières ou générales) pour fonder une nouvelle connaissance est une *synthèse*.

L'intuition correspond à la perception [6]. Or, il est facile de constater que l'on ne perçoit pas n'importe comment, mais dans des cadres définis *a priori*. Ces cadres sont l'espace et le temps, les formes pures (*i.e.* pures de tout contenu sensitif, résultant de l'expérience) de l'intuition ou de la perception. L'espace est la forme pure de la perception externe, le temps de la perception interne.

La sensation délivre un divers spatio-temporel et non un objet. Cela signifie que la sensation donne un contenu inscrit dans un continuum temporel et dans un continuum spatial, mais ce contenu n'est pas objectivé en un objet. C'est en cela que l'intuition délivre des représentations singulières : la sensation n'est pas encore subsumée par le concept d'un objet. Par ailleurs, le temps et l'espace sont des structures transcendantales [7] du sujet : elles accompagnent toutes les perceptions (elles en sont une condition de possibilité) et ne sont jamais rencontrées pour elles-mêmes en dehors des perceptions. Ce sont des formes de la perception, des abstractions formelles (et non conceptuelles) des sensations.

Pour qu'il y ait objet perçu, il faut l'intervention de l'entendement, dont la tâche sera de synthétiser, d'intégrer le divers senti en un objet scientifique selon une règle donnée par un concept. Par exemple, les diverses sensations d'un chien devant soi donnent des représentations singulières qui doivent être rassemblées, intégrées, synthétisées, en une représentation générale du chien.

Précisons. L'imagination empirique est une faculté chez Kant consistant à ajouter librement des sensations (*i.e.* intégrer le divers à d'autres sensations, au hasard, de manière contingente et arbitraire, sans règle précisant comment effectuer ces synthèses de divers. Cette absence de règle implique qu'il n'existe aucune assurance d'effectuer la même synthèse en face de mêmes sensations à des moments différents, ou dans des sujets différents. Synthétiser des représentations particulières (divers) en une représentation est un jugement : l'imagination empirique permet donc d'effectuer des jugements dits de perception, car ils sont contingents et subjectifs. Ce n'est pas sur eux qu'il est possible de construire une science. Ces jugements correspondent à nos connaissances linguistiques et leurs objets, fruits de synthèses contingentes, ressortissent à l'objectivité linguistique.

Cela signifie en particulier qu'il n'y a pas véritablement d'objet donné par la sensation

5. « Toutes les connaissances, c'est-à-dire toutes les représentations rapportées consciemment à un objet sont ou bien des *intuitions*, ou bien des *concepts*. – L'intuition est une représentation *singulière*, le concept est une représentation *générale* ou *réfléchie*. » (Kant, 1982, p.99)

6. La *perception* se rapporte à la sensation, ce qui est donné par les sens. L'*intuition* est plus générale et désigne toute faculté réceptive. La perception est donc une intuition sensible. Chez Kant, intuition et perception se confondent. Elles se distingueront chez Husserl où l'intuition pourra être non sensible : intuition des essences et intuition catégoriale.

7. Est transcendantal ce qui est *a priori*, *i.e.* antérieur à l'expérience, et valant pour l'expérience.

et l'association empirique de l'imagination. En effet, comment parler d'objet donné par l'expérience si ce dernier peut ne pas exister pour un autre sujet, s'il peut être différent l'instant suivant, dû à d'autres associations de l'imagination. Ne peut être considéré comme objet donné par l'expérience que ce qui est universel et nécessaire. Il faut donc, pour avoir des objets donnés dans l'expérience, des principes guidant les synthèses.

> « Cette loi de la reproduction suppose que les phénomènes eux-mêmes soient soumis réellement à une règle et que ce qu'il y a de divers dans leurs représentations forme une suite ou une série suivant certaines règles ; ... Si le cinabre était tantôt rouge, tantôt noir, tantôt léger, tantôt lourd, si un homme se transformait tantôt en un animal, tantôt en un autre, si dans un long jour la terre était couverte de fruits, tantôt de glace et de neige, mon imagination empirique ne pourrait jamais trouver l'occasion de recevoir dans la pensée le lourd cinabre avec la représentation de la couleur rouge... il faut donc qu'il y ait quelque chose qui rende possible cette reproduction des phénomènes, en servant de principe a priori à une unité nécessaire et synthétique des phénomènes. » in (Kant, 1984b, p.113).

Ces principes seront donc les conditions nécessaires pour construire de véritables objets, *i.e.* des objets scientifiques. Les synthèses correspondantes seront alors des jugements scientifiques ou connaissances scientifiques. Ces synthèses nécessaires sous l'égide de principes permettant de construire des objets scientifiques sont des *synthèses objectivantes*. Les principes sont les catégories, où formes pures[8] de la pensée. Ce sont les concepts ou règles indiquant comment effectuer des synthèses valant universellement. Ce sont les principes d'intégration du divers. Les catégories sont les structures transcendantales du jugement d'expérience qui permettent de poser les objets scientifiques : Kant appelle dans les *Prolégomènes* « jugements d'expérience » les jugements scientifiques. Ce sont les jugements de perception subsumés par les catégories : comme les concepts sont des méthodes d'intégration, cela signifie que la synthèse de l'imagination s'effectue sous le contrôle de la méthode prescrite par la catégorie. L'imagination n'est alors plus empirique mais transcendantale. En conclusion, chez Kant, l'objet n'est pas perçu mais construit à partir des données de la perception.

Les catégories sont les concepts purs de l'entendement. Ces concepts sont *a priori* : règles pour effectuer les synthèses, ils ne peuvent que les précéder. Ils sont donc transcendantaux : ce n'est que par leur entremise qu'un objet peut être donné, et ils ne sont valables que pour les synthèses de l'expérience.

Si ces principes sont transcendantaux, c'est-à-dire non psychologiques, ils ne peuvent être trouvés par introspection ou par l'évidence de la réflexion car il ne leur correspond aucune représentation mentale. Il faut donc les *déduire*, et montrer que les principes déduits sont bien ceux rendant possible l'expérience scientifique, c'est-à-dire les synthèses objectivantes.

La déduction métaphysique permet de trouver quelles sont les catégories, et d'en dresser la liste. Kant appelle déduction transcendantale le raisonnement qui lui permet de montrer d'une part que les catégories valent pour l'expérience et que pour elle, et d'autre part qu'elles sont la possibilité de l'expérience : sans catégories, impossible d'unifier un

8. Forme signifie chez Kant abstraction, où l'on extrait d'un contenu un aspect appelé forme. Forme et abstraction formelle sont donc synonymes. Pur qualifie quelque chose obtenu par abstraction du contenu empirique. Ce qui est pur n'est jamais rencontré tout seul, mais toujours avec dont il est forme, *i.e.* un contenu empirique. La pureté est donc une notion épistémologique et non psychologique.

divers en un objet. Enfin, la déduction transcendantale objective montre comment l'expérience est fondée sur les catégories, et la déduction transcendantale subjective comment les catégories opèrent effectivement dans la conscience. Alors que la déduction objective montre la nécessité conceptuelle et épistémologique des catégories, la déduction subjective montre leur nécessité en décrivant comment la connaissance se construit dans l'esprit.

1.1.1 La déduction métaphysique

Le principe de cette déduction est le suivant. La liste des catégories se déduit en prenant en considération l'activité de synthèse en tant que telle. Or, synthétiser, c'est poser ensemble des représentations. Quand il s'agit de représentation générale ou concept, la synthèse consiste en un jugement de type « S est P » où le terme sujet S et le terme prédicat P reliés par la copule sont unis dans une même opération de la pensée. Il est donc possible de distinguer autant de types de jugements que de manières possibles de penser l'unité du sujet et du prédicat. Si l'on dispose d'une typologie des jugements possibles, il est alors possible de déduire les principes d'unité permettant de synthétiser les représentations. Or, la logique, en tant que théorie du jugement, fournit la table des catégories de jugements : la déduction métaphysique doit donc prendre appui sur le travail des logiciens pour déduire les catégories ou principes de synthèse du divers. Pour justifier cette déduction, Kant reprend la caractérisation classique du jugement, telle que la logique classique le théorise :

> « Dans chaque jugement est un concept, qui vaut pour plusieurs concepts, et qui en cette pluralité, comprend aussi une représentation donnée, qui est en effet rapportée immédiatement à l'objet. » A 68 / B 93 (Kant, 1980, p.131)

Dans un jugement « S est P », le prédicat P est un concept qui se rapporte à plusieurs concepts sujets. Dans ce qu'il faudra appeler l'ancienne théorie du jugement après la critique kantienne, celle héritée d'Aristote et transmise par les médiévaux, il faut distinguer trois domaines, la conceptualisation, le jugement, le raisonnement, articulant trois notions clefs, le terme ou concept, la proposition ou jugement jugé, le syllogisme. La conceptualisation abstrait du sensible le concept, le jugement lie des concepts en affirmant ou niant l'inhérence du prédicat au sujet. Le concept et le jugement sont donc de nature différente dans la théorie classique.

Kant analyse pour sa part le jugement en affirmant que le terme sujet est un terme issu de l'intuition, c'est-à-dire qu'il est abstrait du sensible, et que le prédicat en affirme une propriété générale. Cette propriété générale étant partagée par d'autres concepts que le concept sujet, le concept prédicat a pour fonction de lier ensemble en une même représentation conceptuelle tous les sujets dont on peut affirmer ce prédicat :

> « Tous les jugements sont d'après cela des fonctions de l'unité parmi nos représentations, puisque justement, au lieu d'une représentation immédiate [une intuition pour Kant], on emploie pour connaître l'objet une représentation plus *élevée*, qui comprend sous elle celle-ci et plusieurs autres, et qu'ainsi beaucoup de connaissances possibles sont rassemblées en une. » A 69 / B 94 (Kant, 1980, p.132)

La généralité du prédicat permet de penser l'unité des sujets. Kant donne comme exemple le jugement « Tous les corps sont divisibles » :

> « Ainsi par exemple, dans le jugement : *Tous les corps sont divisibles*, le concept de divisible se rapporte-t-il à divers autres concepts ; mais parmi eux il est rapporté ici particulièrement au concept du corps, et celui-ci à certains phénomènes qui se présentent à nous. Donc ces objets sont représentés médiatement par le concept de la divisibilité. » A 68-69 / B 93-94 (Kant, 1980, p.131-132)

Dans cette analyse du jugement, deux faits essentiels méritent d'être relevés. Premièrement, tout jugement est la synthèse en un terme prédicat des termes sujets dont on peut affirmer ou nier ce terme prédicat. Deuxièmement, le terme sujet peut être un terme singulier, c'est-à-dire une intuition, c'est dans ce cas un jugement singulier. Dans le vocabulaire kantien, il faut en effet distinguer les représentations singulières, qui sont des intuitions ou termes singuliers, des concepts qui sont des représentations générales ou termes généraux. Contrairement à la logique aristotélicienne, mais conformément à la logique médiévale, Kant accepte les jugements singuliers ayant pour sujet des termes singuliers, ces jugements revenant à des jugements universels : « Socrate est mortel » se lisant en effet comme « Tous ce qui est Socrate est mortel » ; c'est le *omne quod est* des médiévaux[9] (Blanché, 1970). En appliquant l'analyse kantienne des jugements aux jugements singuliers, on constate que le terme prédicat d'un jugement singulier est la synthèse des intuitions sujets. Ce qui veut donc dire que le concept ou prédicat est la synthèse des intuitions ou sujets, tout comme le concept prédicat est la synthèse des concepts sujets dans les jugements généraux :

> « Penser est la connaissance par concepts. Mais les concepts se rapportent, comme prédicats de jugements possibles, à quelque représentation d'un objet *encore* indéterminé. Ainsi le concept du corps signifie quelque chose, par exemple un métal, qui peut être connu par ce concept. Il est donc un concept du seul fait que d'autres représentations sont contenues sous lui, au moyen desquelles il peut se rapporter à des objets. C'est donc le prédicat d'un jugement possible, par exemple tout métal est un corps. » A 69 / B 94 (Kant, 1980, p.132)

La nouveauté, dans la théorie kantienne du jugement, consiste donc à annuler la distinction entre le concept et le jugement. Les scolastiques opposaient les concepts abstraits du sensible aux jugements affirmant l'appartenance ou la séparation de ces concepts, c'est-à-dire qu'ils distinguaient et opposaient la fonction de conceptualisation ou abstraction du sensible et la fonction de liaison dans le jugement : le jugement n'a qu'un rapport médiat aux objets, *via* les termes qu'il relie, alors que le concept à un rapport immédiat ou direct avec les choses dont il est l'abstraction directe. Kant affirme au contraire que le concept, à l'instar du jugement, n'a qu'un rapport médiat, indirect à l'objet, au lieu d'en être directement abstrait, dans la mesure où il n'est également qu'une liaison de représentations, immédiates et c'est alors un concept empirique, médiates et c'est alors un concept logique.

En effet, les concepts prédicats des jugements, s'ordonnent en une hiérarchie d'abstraction (extension) croissante, le concept abstrait étant lié aux objets qui tombent sous lui par l'intermédiaire des concepts qu'il subsume. Par conséquent, les concepts, en tant que

9. Ce traitement des jugements singuliers est insatisfaisant pour un moderne dans la mesure où il confond l'élément et le singleton.

représentations générales, ne sont que des fonctions de liaison entre des concepts plus spécifiques, des fonctions permettant de les penser en une unité plus abstraite. En particulier, un concept ou représentation générale est une synthèse de représentations singulières ou intuitions : dans les jugements singuliers, le sujet est un terme singulier et le prédicat un terme général. La liaison faite dans ce jugement est donc la liaison opérée par le concept parmi tous les sujets possibles du prédicat. Il s'agit du problème du schématisme, où il s'agit de comprendre comment des représentations singulières, les intuitions, peuvent être synthétisées, liées, intégrées, en un concept [10]. C'est donc pourquoi Kant peut poser que les catégories ou concepts purs, sont les principes de liaison des représentations et coïncident de ce fait avec les fonctions du jugement, qui ne sont que les principes d'unité sous lesquels penser et synthétiser les termes sujets possibles.

> « La même fonction qui donne l'unité aux représentations diverses *dans un jugement*, donne aussi à la simple synthèse de représentations diverses *dans une intuition* l'unité, qui, exprimée généralement, s'appelle le concept pur de l'entendement. Le même entendement donc, et bien par les mêmes actes par lesquels il établissait dans les concepts, au moyen de l'unité analytique, la forme logique d'un jugement, apporte aussi dans ses représentations, au moyen de l'unité synthétique du divers dans l'intuition en général, un contenu transcendantal, en raison de quoi ces représentations s'appellent de purs concepts de l'entendement, qui se rapportent *a priori* aux objets, ce que ne peut faire la logique générale. » A 79 / B 104-105 (Kant, 1980, p.140)

Les concepts sont donc des principes de liaison des intuitions, du divers de l'intuition, si bien qu'ils sont un contenu transcendantal dans la mesure où ils s'appliquent *a priori* à l'expérience. Par conséquent, les catégories « ne sont autre chose que les simples formes des jugements en tant qu'appliqués à des intuitions » [11]. Le problème est à présent de montrer comment la liaison subjective permet de construire un objet, puisse avoir une valeur objective dans l'expérience. C'est le problème de la déduction transcendantale.

1.1.2 La déduction transcendantale

Mais, ce qui est valable pour la synthèse conceptuelle de représentations générales en jugement logique ne vaut pas obligatoirement pour la synthèse de représentations singulières en concept. Il faut donc justifier dans quelle mesure les catégories du jugement sont aussi les catégories de la synthèse du divers.

Toute synthèse comprend deux moments logiques : en tant que « poser ensemble », une synthèse doit proposer un cadre, une instance où les termes à synthétiser sont juxtaposés ; en tant que production d'une représentation, une synthèse est une unité et elle consiste alors en une modalité permettant de penser l'unité. Par conséquent, il faut, pour caractériser une synthèse, préciser d'une part quelle est l'instance du « poser ensemble », et d'autre part quel est le mode d'unité. Par définition, le jugement est un cadre permettant de « poser ensemble » sujet et prédicat. Son mode d'unité est la catégorie du jugement comme cela a été dit plus haut. Mais, en ce qui concerne les synthèses des représenta-

10. A la remarque près que le schématisme traite davantage de l'interprétation du concept dans le temps, de sa détermination temporelle par le moyen de l'imagination. Ici il s'agit en quelque sorte du problème inverse où il s'agit de « remonter » de l'intuition au concept.

11. *in* (Kant, 1990, p.17 note)

tions singulières, il est plus difficile de penser un cadre permettant de poser ensemble des représentations qui seront alors unifiées selon un mode donné d'unité.

Ce cadre commun pour penser les représentations singulières est celui de l'expérience possible : c'est dans le cadre d'une expérience possible, sous la forme d'un divers empirique, que les représentations singulières sont données, et ce sont les modes d'unité de l'expérience possible qui permettront d'unifier ces représentations.

Les modes d'unité du jugement deviennent dans ce cadre les modes d'unité des l'expérience possible. Or, l'unité de l'expérience est la possibilité de l'expérience : c'est parce qu'il est possible d'unifier des représentations qu'il est possible d'avoir l'expérience d'un objet. Par conséquent, les catégories de l'expérience possible sont la condition de possibilité de l'expérience et de la constitution d'un objet de l'expérience. En effet, puisque l'expérience possible est le cadre dans lequel les catégories permettent d'unifier les représentations en objet, il apparaît donc que les catégories sont, outre les conditions de l'expérience possible, les conditions de possibilité de l'objet de l'expérience :

> « Les conditions de la possibilité de l'expérience en général sont aussi des conditions de la possibilité des objets de l'expérience et ont pour ce motif une valeur objective dans un jugement synthétique a priori. » in (Kant, 1984b, p.162)

Dans ces conditions, qu'est-ce qu'un objet de l'expérience ? C'est donc, d'après cette déduction transcendantale, ce qui est déterminé par les catégories lors de la synthèse du divers. C'est donc quelque chose qui, pensé sous la catégorie de la quantité, est déterminé quant au nombre (unité, pluralité, totalité), sous la catégorie de la qualité, est déterminé quant au degré de réalité (réalité, négation, limitation), sous la catégorie de la relation, est déterminé quant au rapport aux autres objets de l'expérience, (inhérence, causalité, communauté), et enfin sous la catégorie de la modalité, est déterminé quant au rapport aux conditions de l'expérience, *i.e.* est déterminé quant au rapport à l'entendement (possibilité, existence, nécessité).

Un objet scientifique est donc un pôle objectif X de détermination, les principes de détermination étant les catégories. Il appert qu'un objet est construit, dans la mesure où c'est la détermination du X qui lui confère une teneur objective, le fait qu'il soit quelque chose et non une illusion subjective.

1.2 Du concept aux phénomènes

Une fois posé que les catégories sont les conditions nécessaires d'une synthèse objectivante nécessaire et universelle du divers, il faut comprendre comment celle-ci peut fonctionner. Les catégories, formes pures de la pensée, ne se donnent jamais directement, mais à travers les jugements d'expérience qu'elles autorisent. En cela, on ne peut les trouver psychologiquement en soi, on ne peut que les déduire. Ce sont des notions épistémologiques et non psychologiques [12].

12. C'est en cela que Kant rend acte à Hume du fait qu'il est impossible d'avoir présent à l'esprit un concept général comme un triangle général. Mais leur nier une dimension psychologique n'est pas nier leur existence épistémologique.

En effet, Kant montre que les catégories sont nécessaires, sinon les objets de la connaissance sont impossibles. C'est la déduction transcendantale objective, où l'on déduit que les catégories sont nécessaires pour l'unité de l'expérience, *i.e.* pour la possibilité d'une expérience scientifique. Mais il ne dit pas comment les catégories opèrent. Il faut donc avoir une déduction subjective, où l'on dit comment cela se passe effectivement. C'est une phénoménologie de la connaissance scientifique (Philonenko, 1968) que Kant doit exhiber ici pour asseoir sa conception transcendantale. Cela n'est pas immédiat, évident, puisque les catégories sont des concepts épistémologiques et non psychologiques. Puisqu'il n'existe pas de représentations mentales des catégories, cette phénoménologie ne peut être une phénoménologie de l'évidence, et doit reposer sur une déduction. Mais, cette déduction, devant aboutir à une caractérisation psychologique des catégories, est subjective, et non pas seulement objective, quoique cette dernière soit seule en droit nécessaire [13].

Néanmoins, il faut bien en passer par la réalisation psychologique des catégories, ne serait-ce que pour savoir reconnaître un jugement scientifique. Il faut donc montrer comment elles se manifestent dans la conscience. Or, une manifestation psychologique est un phénomène du sens interne, donc un phénomène temporel, puisque le temps est la forme du sens interne.

Il faut donc associer à chaque catégorie une règle montrant comment reconnaître sa manifestation dans des phénomènes temporels. Ces règles sont les schèmes et donnent de manière générale la méthode pour appliquer tout concept [14]. Ils correspondent à la méthode suivie par l'imagination pour effectuer la synthèse temporelle.

L'intégration temporelle s'effectue au moyen des schèmes. Le schème prend place dans l'imagination car il intègre des représentations mentales du sens interne. Le schème indique comment la catégorie se concrétise dans la conscience : comme la forme du sens interne (*i.e.* la conscience) est le temps, il indique la relation temporelle correspondant à la catégorie. Ce principe est donc *a priori* comme l'est la catégorie, et donc il caractérise non pas l'imagination effectuant des synthèses au hasard, imagination qualifiée d'empirique, mais l'imagination transcendantale car elle contient les méthodes qui déploient dans la conscience la règle de synthèse qu'est la catégorie. C'est ainsi que la succession temporelle universelle est le schème de la causalité (A cause B si A est toujours suivi de B). Les schèmes sont ainsi les *principes de l'imagination*.

L'intégration spatiale résulte des *principes de l'expérience*. Ils correspondent à l'application des schèmes (intégration temporelle des représentations en tant qu'elles s'inscrivent dans le sens interne) aux représentations spatiales du sens externe. Ces principes sont les axiomes de l'intuition, les anticipations de la perception, les analogies de l'expérience, et les postulats de la pensée empirique.

Les catégories, formes pures de l'entendement s'appliquent donc aux phénomènes empiriques en s'appliquant aux formes pures de la sensibilité : le temps et l'espace. L'application

13. « Car si l'on veut prouver QUE mes catégories dont la raison doit se servir pout toute sa connaissance, ne peuvent être utilisées que par rapport aux objets de l'expérience (parce que c'est dans ce domaine seul qu'elles rendent possible la forme de la pensée), la réponse à la question de savoir *comment* elles la rendent possible, est assurément assez importante pour *achever* si possible cette déduction, mais elle n'est nullement nécessaire, elle est simplement *méritoire* en ce qui concerne la fin principale du système qui est de déterminer les limites de la raison pure. » (Kant, 1990, p. 17, note)

14. Les catégories ou formes transcendantales se réalisent par le schématisme transcendantal ; les concepts mathématiques par le schématisme pur, les concepts empiriques par le schématisme empirique.

au temps produit les principes de l'imagination, les schèmes qui sont ainsi les formes pures de l'imagination, qui se fait pour l'occasion imagination transcendantale. L'application à l'espace produit les principes de l'expérience. Ces principes, résultant de l'application des catégories conceptuelles au temps et à l'espace, sont donc les conditions de possibilité de toute science de la nature, c'est-à-dire de toute science de phénomènes spatio-temporels donnés à l'intuition. Mais, si ce sont là les conditions de possibilité de toute physique, il n'est pas pour autant donné une physique effective de la nature. La raison en est simple : la spontanéité du concept ne saurait remplacer la réceptivité du phénomène. Si le concept est en quelque sorte la condition d'accueil, le moule tout prêt pour recevoir la matière de l'expérience empirique, *i.e.* la sensation, pour y imprimer la forme de sa légalisation, il a besoin d'une matière donnée, reçue, pour s'appliquer.

1.3 Nature formelle et nature matérielle

1.3.1 Le problème

Les principes de l'expérience sont donc les principes de la construction d'un objet à partir d'un divers. Cela signifie que ces principes, pour une sensation donnée, *i.e.* un divers spatio-temporel, peuvent la déterminer en objet. Mais, sans cette sensation, les principes ne définissent que les conditions formelles qu'une science de la nature doit respecter pour être une véritable science. Kant dit que ces conditions formelles définissent *une nature formelle* par opposition à la *nature matérielle* qui est la nature formelle appliquée à des sensations réelles.

Explicitons. Ce que le divers spatio-temporel indique, c'est une existence effective dans le monde. Ce que la construction indique, c'est la nature de ce qui existe, sa réalité ou essence. Or, l'existence ne peut être donnée qu'*a posteriori* et n'est jamais déduite ni posée par un concept, *i.e.* par la pensée ou l'entendement :

> « car des lois, c'est-à-dire des principes de la nécessité de ce qui appartient à l'existence d'une chose, se rapportent à un concept qu'on ne peut construire parce que l'existence ne peut se représenter dans aucune intuition a priori. » in (Kant, 1990)

Les principes ne peuvent donc poser l'existence. En revanche, ils peuvent se fonder sur le concept d'un objet existant, même si ce concept est déterminé seulement *a posteriori*. La notion d'un objet existant, sans que celui-ci soit déterminé, c'est la notion d'expérience possible. Le raisonnement est : pour une expérience possible, *i.e.* pour un objet existant donné (mais que je ne puis préciser davantage, car si j'ai le concept de son existence, je n'ai pas le contenu, je ne sais s'il existe vraiment ni comment), le principe de causalité affirme que rien n'arrive sans cause. Ce principe est au niveau de l'entendement et donc est discursif et non intuitif. Cela signifie qu'il prend la forme d'un jugement logique (en l'occurrence, le jugement de la causalité est de forme hypothétique : Si A Alors B) qui pour subsumer du réel doit attendre la donation empirique d'une existence réelle, *i.e.* une sensation perceptive.

La nature formelle n'est pas la nature matérielle car sinon cela voudrait dire que penser la nature (nature formelle définie par les principes) permet de la connaître (nature matérielle déterminée) et que la pensée construit la science par ses seules ressources sans faire appel à l'expérience, ce qui serait retomber dans la métaphysique cartésienne. On dit que la nature formelle n'a donc affaire qu'à des objets *indéterminés* quant à leur existence empirique, car non déterminés à partir d'une sensation empirique donnée par l'intuition sensible. Pour passer de la nature formelle à la nature matérielle, il faut donc déterminer ces objets quant à l'existence empirique. Une existence empirique donnée par le choc de l'expérience et donnant la matière à une détermination scientifique permet donc ce passage, *i.e.* le passage entre la science possible et la science réelle, entre l'expérience possible et l'expérience réelle.

Il faut donc un concept empirique donné, c'est-à-dire une existence spatio-temporelle posant un divers, à partir duquel les objets sont construits. Dans le cas des sciences physiques qui intéressent Kant, ce concept empirique est celui de mouvement :

> « La détermination fondamentale d'un quelque chose qui doit être un objet des sens externes, devait être le mouvement car celui-ci seul peut affecter ces sens. » in (Kant, 1990, p.20)

Par conséquent, à partir d'une existence empirique donnée, l'entendement applique ses principes pour construire une science matérielle de la nature. Cela signifie que, une existence étant donnée, les principes peuvent s'appliquer et la déterminer en objet.

Mais, comment faire face à la diversité des sensations, des contenus empiriques ? Comment être universel et nécessaire dans la contingence de l'existence empirique ? Comment savoir effectuer une synthèse valant universellement et nécessairement, *i.e.* une synthèse objectivante, sur la diversité empirique et contingente de l'expérience ? Car on retrouve le même problème qu'avec le temps : le temps est invisible, et donc il est impossible d'effectuer une synthèse temporelle. Mais l'espace est invisible lui aussi, parce que pur également. Il serait donc impossible d'effectuer une synthèse spatio-temporelle.

La solution vient en remarquant que l'existence empirique s'inscrit dans le cadre universel et nécessaire de l'intuition pure, le cadre spatio-temporel. Si l'on possède des connaissances sur ce cadre, ces connaissances seront universelles et nécessaires, et détermineront le contenu de l'expérience. Réciproquement, le cadre spatio-temporel étant invisible, mais seulement révélé à travers l'expérience empirique, il faut pouvoir remonter à partir de cette dernière au cadre spatio-temporel, donc remonter de l'intuition concrète à l'intuition pure, pour effectuer une synthèse spatio-temporelle. Il faut donc trouver une telle connaissance permettant d'une part d'appliquer l'universel scientifique à la contingence empirique, et d'autre part de retrouver cet universel dans l'expérience.

1.3.2 Le rôle des mathématiques

La réponse de Kant est que les mathématiques sont cette connaissance. En effet, les mathématiques recèlent en elles le principe de la solution au kantisme. Le problème est que le temps et l'espace purs ne sont jamais rencontrés en eux-mêmes, mais toujours à l'occasion d'expériences empiriques. Par conséquent, les connaissances sur l'intuition pure ne peuvent être données d'emblée pures à l'esprit, mais doivent être *déduites* des connaissances empiriques effectives possédant un contenu sensible. Le possible n'existe qu'en vue du fait dont il est la possibilité : le possible kantien n'est pas un virtuel contrefactuel, à la manière d'une monde possible logique dont la seule exigence est qu'il ne soit pas logiquement contradictoire; le possible kantien est la condition de possibilité déduite du fait : jamais premier, seul le fait l'est, il en est une condition posée après coup. C'est dire qu'on ne va pas d'un virtuel logique à l'existence factuelle, mais bien plutôt de l'existence de fait au potentiel qui rend compte de sa possibilité. Kant l'exprime ainsi :

> « Or, les choses dans l'espace et dans le temps ne sont données qu'autant qu'elles sont des perceptions (des représentations accompagnées de sensation), et par conséquent au moyen d'une représentation empirique. Par conséquent, les concepts purs de l'entendement même quand ils sont appliqués à des intuitions *a priori* (comme dans la mathématique) ne procurent de connaissance qu'autant que ces intuitions, et par suite aussi par leur moyen les concepts de l'entendement, peuvent être appliqués à des intuitions empiriques. Par conséquent, les catégories ne nous fournissent au moyen de l'intuition aucune connaissance des choses, si ce n'est par leur application possible à l'*intuition empirique*, c'est-à-dire qu'elles ne servent qu'à la possibilité de la *connaissance empirique*. Par conséquent, les catégories n'ont d'usage pour la connaissance des choses qu'autant que ces dernières sont regardées comme des objets d'expérience possible. » (Kant, 1980, p.169-170)

C'est la notion même d'expérience possible qui transforme radicalement la notion de possibilité, la faisant basculer du métaphysique pour lequel être, c'est ne pas être contradictoire, au transcendantal pour lequel être, c'est exister de fait.

Le problème est donc de pouvoir remonter du contenu empirique à son cadre *a priori*, c'est-à-dire de remonter du fait à son principe de possibilité. Il suffit de pouvoir passer du phénomène concret au phénomène pur car, ce dernier étant *a priori*, la spontanéité du concept peut opérer et appliquer les principes de l'expérience.

Or, dans les mathématiques, on raisonne de manière générale sur des représentations sensibles concrètes, comme par exemple une figure géométrique dessinée sur le papier. Car on ne peut accéder à l'intuition pure, sinon par l'intuition concrète. Mais, réciproquement, en face de l'intuition concrète (la figure sur le papier), comment revenir à son cadre spatio-temporel ? En effectuant un raisonnement mathématique. Les mathématiques permettent de raisonner formellement [15] sur du concret, en abstrayant du concret à son cadre spatio-temporel.

Cela s'explique par le fait que les concepts mathématiques sont par définition une méthode indiquant comment construire leur objet. Ainsi, le concept de cercle est-il défini comme un ensemble de points à une distance identique d'un point appelé centre : cette définition indique comment construire un cercle. C'est pour cela que les mathématiques permettent de raisonner de manière générale sur des cas particuliers, car on raisonne sur la méthode de construction, et non sur son résultat sensible et concret :

> « La figure singulière qu'on a dessinée est empirique et cependant elle sert à exprimer le concept malgré sa généralité, parce que dans cette intuition empirique on ne considère jamais que l'acte de la construction du concept, auquel beaucoup de déterminations comme celles de la grandeur, des côtés et des angles, sont tout à fait indifférentes et qu'on fait, par suite, abstraction de ces différences, qui ne changent pas le concept du triangle. » *in* (Kant, 1984b, p. 494)

15. Il faut prendre ici la notion de « formel » dans un sens non syntaxique : il s'agit de la forme « géométrique » du phénomène, et non de sa forme syntaxique ou conceptuelle.

Cette caractéristique des mathématiques joue un rôle immense dans l'analyse kantienne de la science de la nature, car elle permet à la fois de comprendre que toute science de la nature ne peut être qu'une science mathématisée, et que les mathématiques ne sauraient être à elles seules une physique. Comme on le sait, le principe de cette solution repose sur le fait que les vérités mathématiques portent sur l'intuition pure : antérieures à toute expérience concrète, elles s'appliquent à toute expérience.

D'une part, en effet ,les vérités mathématiques sont synthétiques *a priori*. Ne portant pas sur l'existence empirique mais sa forme seulement, les mathématiques enrichissent indépendamment de toute expérience ma connaissance de concepts comme *triangle* ou *cercle*. Ce qui me permet de sortir de mon concept de triangle pour l'enrichir de nouvelles connaissances, c'est sa figuration dans l'intuition pure, *i.e.* le raisonnement général sur une figure particulière. La figure est une donation sensible non conceptuelle qui ajoute au concept mais ne l'explicite pas. Les mathématiques étudient donc la forme pure de l'intuition, la forme de toute donation empirique. Si l'on considère que ce qui vaut pour la forme vaut pour le contenu, toute vérité mathématique valant pour la forme spatio-temporelle de l'intuition vaudra pour toute sensation empirique. Cela signifie que les mathématiques déterminent *a priori* quel *peut* être le contenu de l'expérience, sa réalité possible, *i.e.* son essence.

D'autre part, elles ne peuvent déterminer l'existence empirique d'un objet, car l'existence ne peut être construite mathématiquement *a priori*, mais seulement donnée perceptivement *a posteriori*. *A priori*, on ne dispose pas de contenu empirique, mais de l'intuition pure, *i.e.* le cadre spatio-temporel. Si on détermine ce cadre, on obtient les mathématiques. Mais la physique porte sur l'existence des choses, et les rapports d'existence entre les choses, *e.g.* la causalité. *A priori*, elle n'a que le concept et pas le contenu intuitif correspondant. Donc elle n'a rien à déterminer *a priori*. La physique ne pourra donc être effective qu'*a posteriori*, *i.e.* avec une existence empiriquement donnée. C'est en cela que, du point de vue de l'*a priori*, les connaissances mathématiques sont immédiates et intuitives (opèrent directement dans l'intuition), mais ne portent que sur l'essence des choses, et que les connaissances physiques sont médiates et discursives (faute d'une existence empirique, c'est la médiation nécessaire, elles restent conceptuelles ou discursives), car elles portent sur l'existence des choses.

Par conséquent, la détermination est soit mathématique, et alors elle porte sur l'essence de l'objet et non sur son existence, soit physique, et alors elle porte sur l'existence de l'objet. Dans le premier cas, la détermination est *a priori*, et constructive : elle construit *a priori* l'essence de l'objet. Ce sont les principes mathématiques de Kant, axiomes de l'intuition et anticipations de la perception, tirés des catégories de qualité et quantité. Dans le second cas, la détermination est *a posteriori* car repose sur une donation empirique de l'existence. Cela correspond aux principes dynamiques de Kant, analogies de l'expérience et postulats de la pensée empirique, tirés des catégories de relation (rapport d'existence entre les objets de l'expérience) et modalité (rapport d'existence entre le sujet et ses objets d'expérience), qui *a priori* ne portent que sur l'expérience possible et non réelle, et qui pour porter sur cette dernière, doivent attendre une existence empirique[16]. Elles ne peuvent construire *a priori* l'existence de leur objet (alors que les catégories mathématiques peuvent construire *a priori* l'essence de leur objet). Par exemple, je ne peux poser *a*

16. « Dans l'application des concepts purs de l'entendement à l'expérience possible, l'emploi de leur synthèse est ou *mathématique* ou *dynamique* ; car il porte en partie simplement sur l'*intuition*, en partie sur l'*existence* d'un phénomène en général. » *Ibid.* p. 163.

priori qu'un événement A possède une cause précise B, mais qu'il doit en posséder une nécessairement. Car le concept d'existence de A, même non instancié par l'expérience, me permet de déduire le concept de l'existence d'un autre événement appelé cause dans le cadre de la nature formelle, *i.e.* de l'expérience possible. Mais, pour en savoir plus, pour les déterminer, je dois recourir à l'expérience concrète.

1.3.3 La science mathématique

La conception kantienne des mathématiques entraîne, comme nous l'avons dit, plusieurs conséquences. La première est que toute science ne peut être que mathématique : elle doit en effet construire ses concepts, et seules les mathématiques permettent de le faire. Etant donné que le divers est inscrit dans un continuum d'espace et de temps, la construction ne peut être que mathématique. Le schématisme transcendantal (correspondant à une nature en général) devient alors construction mathématique (pour caractériser cette nature-ci). Il n'y a donc de science que de science mathématique comme le dit Kant lui même :

> « Or, je soutiens que dans tout théorie particulière de la nature, il n'y a de science proprement dite qu'autant qu'il s'y trouve de mathématique ... la théorie de la nature ne renfermera de véritable science que dans la mesure où la mathématique peut s'y appliquer. » in (Kant, 1990)

Les mathématiques détrônent donc la logique de leur rôle d'outil privilégié de la science. Pour Kant, la logique est analytique : elle n'ajoute aucune connaissance aux concepts manipulés. La raison en est que toute connaissance doit avoir un contenu. Or, la logique formelle manipule les concepts en faisant abstraction de tout contenu, car elle ne retient que la forme (on dirait aujourd'hui la syntaxe) des concepts. La logique ne peut assurer que la légalité formelle des manipulations conceptuelles : la non-contradiction n'est qu'un critère négatif permettant de reconnaître qu'un jugement ne pourra jamais véhiculer de connaissance, mais n'est pas un critère positif permettant d'augmenter la connaissance. La logique est un *canon*, une légalité de la pensée. Comme nous l'avons dit, les mathématiques sont un *organon* : elles permettent d'augmenter la connaissance comprise sous un concept, elles sont synthétiques.

L'objet kantien est donc construit à partir des phénomènes dans lesquels les catégories introduisent la nécessité universelle permettant d'effectuer des synthèses valables pour tous et partout. Les catégories ajoutent au donné de l'expérience la nécessité dont elle a besoin pour devenir connaissance. Contrairement à l'épistémologie logique où les mathématiques expriment seulement l'expérience, étant le langage de la science, dans le dispositif transcendantal, les mathématiques sont l'outil par lequel les objets sont construits à partir du divers. Cela signifie que les phénomènes, le divers, ne sont pas une manifestation d'un objet réel sous-jacent, des apparences, mais des apparitions à partir desquels la détermination mathématique donne le statut d'objet. La perception ne donne pas d'objets, car ces derniers sont construits, mais seulement leur matière empirique.

Le dispositif transcendantal peut donc être compris comme une tentative de conciliation entre la nécessité de la pensée et la contingence du réel, entre l'essence et l'existence, cette dernière échappant toujours aux lois de l'essence. Mais, pour que l'essence soit une existence possible, et qu'il n'y ait pas de divorce entre la pensée et la connaissance scientifique, il faut que l'essence soit mathématique. En effet, la construction mathématique des concepts permet d'*interpréter* les concepts dans le contexte de l'expérience, de leur donner un contenu empirique. Ce sont donc les mathématiques qui *contextualisent* les concepts et lois physiques : l'interprétation mathématique *est* l'interprétation en contexte ; elle donne le contenu gnoséologique de la connaissance, en quoi elle est une connaissance, elle est l'essence du transcendantal car elle exprime en quoi la connaissance renvoie au réel, à l'expérience concrète. Le dispositif transcendantal explique donc de la manière la plus fine comment la science fonctionne. En résumé, on peut distinguer les deux démarches fondamentales suivantes dans toute activité scientifique :

la modélisation qui consiste à caractériser au niveau de l'intuition pure le phénomène concret, sensible donné dans l'expérience. La modélisation donne une description mathématique des phénomènes qui, pour être dérivée de la sensation, n'en donne pas moins l'essence, la nature mathématique, au niveau *a priori* de l'espace/temps ; la modélisation repose donc sur une idéalisation du phénomène qui le rapporte à la forme pure de l'intuition et à sa détermination mathématique ;

la schématisation consiste à traduire en lois mathématiques les concepts et lois formulés au niveau de l'entendement. La schématisation interprète, détermine, mathématiquement les concepts, les contextualisent dans l'expérience pure. Or, cette expérience pure est précisément celle résultant de l'étape de modélisation.

La modélisation permet de légaliser mathématiquement le phénomène concret en y introduisant l'universalité et la nécessité. La schématisation permet de déterminer ce que signifie les concepts dans le cadre de l'expérience scientifique. Cela veut simplement dire que la signification du mot « gravitation » dans le cadre de l'expérience doit être recherchée dans les formulations mathématiques newtonniennes ou relativistes, et non dans le dictionnaire[17]. Mais cette formulation opposant modélisation et schématisation n'est pas strictement kantienne dans la mesure où il faut, dépassant ainsi la lettre kantienne, interpréter la schématisation du concept scientifique comme une construction mathématique de l'objet. C'est à Jean Petitot (Petitot, 1991a) que l'on doit cette interprétation qui, pour être d'inspiration transcendantale, n'est pas kantienne.

L'existence empirique est donc réduite au simple choc de l'expérience, au simple ébranlement du travail de l'entendement sur les phénomènes. En affirmant que l'essence des phénomènes est de nature mathématique, le contenu même des phénomènes est en quelque sorte un produit *a priori* de l'esprit. Seul le choc de l'expérience, impulsion donnée par la réception du phénomène, échappe à l'esprit et constitue sa finitude. L'existence, la contingence du réel, est donc réduite à sa pure dimension de *hic et nunc*, au fait qu'il est, puisque sa nature, ce qu'il est, n'est déjà plus de l'ordre de la contingence, mais de la légalisation effectuée par la raison. On peut donc dire avec Gilson que :

> « L'idéalisme critique s'accorde donc un donné pour être sûr de ne pas penser à vide, mais il prend aussitôt les mesures nécessaires pour que ce donné ne réserve jamais aucune surprise à la pensée qui l'accueille. Depuis l'instant même où il le réfracte dans les formes *a priori* de la sensibilité, jusqu'au moment où il l'organise en science grâce aux catégories de l'entendement, l'esprit jouit sans réserve du privilège d'exploiter un donné qui soit un réel, et qui ne contienne pourtant absolument rien que ce qu'y met l'esprit. » (Gilson, 1981, p. 202)

17. Même si il est parfois bien plus éclairant de regarder dans le dictionnaire ! Mais, dans ce cas, ce n'est pas l'interprétation dans le cadre de l'expérience possible qui est recherchée, mais l'interprétation dans le cadre de l'échange linguistique.

Mais ce que ne dit pas Gilson, c'est que ce qui permet à Kant d'agir ainsi, c'est que précisément « prendre les mesures nécessaires » consiste à constater que le cadre *a priori* des connaissances, l'espace et le temps, est en lui-même le principe de nouvelles connaissances sur l'existence si bien que le cadre spatio-temporel et son étude mathématique révèlent bien des surprises.

Mais il faut sans doute rendre justice à Gilson que l'expérience possible kantienne n'est peut-être pas la totalité de l'expérience humaine, qu'il faut certes l'identifier à l'expérience scientifique dont elle donne sans doute la meilleure définition, mais qu'il ne faut probablement pas réduire toute expérience, toute connaissance, à l'expérience et connaissance scientifiques. En d'autres termes, l'espace / temps de l'esthétique transcendantale n'est peut-être pas notre seule manière de fonder notre rapport au monde et il faudra certainement faire appel à une attitude philosophique où, par un retour aux choses mêmes telles qu'elles sont données, l'existence, le monde, ne se voit pas réduit à n'être qu'un pur pointeur dans l'espace/temps, mais qu'un contenu, un objet, une chose, puisse être donné. C'est donc à une réforme de l'intuition, cette ouverture sur l'existence, qu'il faudra se résoudre si jamais il faut sortir du cadre mathématique strict de la connaissance scientifique. Alors que chez Kant :

> « Ainsi donc, il y a de l'existence, mais tout se passe heureusement, pour la connaissance sensible, comme s'il n'y en avait pas. » (Gilson, 1981, p. 205)

il faudra rendre compte du fait que pour nombre de connaissances, il y a de l'existence, et qu'on ne peut faire comme s'il n'y en avait pas. Mais ces considérations sont seulement liminaires dans ce chapitre, et elles seront reprises dans les deux chapitres suivants.

1.4 Kant et nous

Ce long détour par Kant nous a permis de définir ce qu'il fallait comprendre dans la notion d'« objet scientifique » et de « contenu scientifique » d'un concept. L'intérêt qu'il y a à éclaircir et préciser de telles notions excède largement les préoccupations d'une épistémologie générale ou théorie générale de la science. C'est directement l'épistémologie expérimentale, ou étude expérimentale de la connaissance qui est ici concernée. En effet, dans le projet d'étudier scientifiquement la pensée et les connaissances, il importe de se donner une caractérisation précise de ce qu'il faut entendre par « étude scientifique ». La lecture kantienne de la science permet d'affirmer, puisque toute science est mathématique, qu'une étude scientifique de la pensée doit correspondre à une mathématisation de la pensée.

Dans quelle mesure est-il possible de mathématiser la pensée? Ce paraît être une préoccupation moderne. Il n'en est rien puisque Kant s'est lui-même explicitement posé la question de la possibilité d'une psychologie scientifique. Kant conclut à une impossibilité pour des raisons propres à sa conception de la science. Mais, si comme nous allons le soutenir, la théorie kantienne de la science est toujours valable, il faut donc que les approches scientifiques de l'esprit et de la pensée puissent être confrontées à l'impossibilité kantienne d'une psychologie scientifique et qu'elles en reçoivent un supplément d'intelligibilité : c'est en effet en comprenant pourquoi ces approches se prétendent scientifiques dans le cadre d'une théorie kantienne de la science et de son interdit concernant la psychologie que l'on prend la mesure de leurs innovations théoriques, de leurs présupposés, et de leurs lacunes.

Nous reprenons donc le projet d'une psychologie scientifique et analysons les raisons pour lesquelles Kant en nie la possibilité. Puis, nous revoyons rapidement les théories contemporaines de l'esprit pour en caractériser la nature au vu de l'analyse kantienne.

Mais l'intérêt d'un exposé de la théorie transcendantale de la science excède le problème de la psychologie scientifique. Dans le cadre de l'intelligence artificielle en particulier et du cognitivisme computationnel en général, l'approche scientifique est endossée par des disciplines comme la logique et l'informatique. Il est donc nécessaire de préciser quels sont les contenus scientifiques mobilisés par ces disciplines pour évaluer leur prétention à rendre compte de l'esprit et de la pensée. Il faut donc résoudre le problème de concevoir une *épistémologie transcendantale de la logique et de l'informatique.* Ce n'est qu'à cette condition qu'il sera possible de comprendre quel type de contenus signitifs peut être appréhendé par une approche computationnelle de l'esprit.

2 Le projet d'une psychologie scientifique

2.1 L'héritage cartésien

Jusqu'à Descartes, les conditions de possibilité de la psychologie scientifique ne pouvaient pas poser de problèmes. En effet, la science, encore d'inspiration aristotélicienne, voit dans la psychologie l'étude des facultés de l'âme, l'âme étant le principe de tout « vivant animal ». La psychologie est donc une branche des sciences naturelles, et plus particulièrement, de la biologie. S'appuyant traditionnellement sur de *De Anima* d'Aristote, cette science du vivant culmine dans les facultés intellectives, dont la prééminence est également étayée par la thèse d'anthropologie philosophique générale que « l'homme, en tant qu'homme, n'est rien d'autre qu'intellect », tirée de l'*Ethique à Nicomaque.*

Mais, Descartes[18], voulant fonder métaphysiquement la science, bouleverse de fond en comble ces données traditionnelles. En affirmant un dualisme *ontologique*, dans les *Méditations métaphysiques*, de la matière et de l'esprit, dont l'essence est respectivement l'étendue et la pensée, il oppose d'une part le monde physique dont la science ne peut être qu'une géométrie, et d'autre part le monde spirituel dont l'étude ne peut être menée que par les idées claires et distinctes[19].

En opposant l'esprit à la matière, Descartes écarte la possibilité que l'on puisse étudier le premier à l'aide des outils forgés pour connaître la seconde. Par conséquent, la possibilité d'une étude de l'esprit par des moyens scientifiques, « scientifique » étant pris dans le sens de « science de la matière physique » s'anéantit. En sortant l'esprit de la nature matérielle, Descartes l'a rendu hors de portée pour les sciences dites « de la nature ». Il y avait donc place pour des sciences dites « de l'esprit », dont les méthodes et l'objet n'avaient rien à emprunter à leurs consœurs de la nature.

En particulier, Descartes ayant posé que l'essence de la matière est l'étendue et que

18. Dont les œuvres complètes sont publiées chez Garnier (Descartes, 1963; Descartes, 1967; Descartes, 1973).

19. Le dualisme cartésien se trouve être ainsi le problème fondateur des sciences cognitives et c'est bien comme cela qu'elles le considèrent, comme en témoignent (Haugeland, 1989; Dennett, 1991; Narayanan, 1990; Bechtel, 1988) et (Jackendoff, 1987).

l'esprit est immatériel, la principale irréductibilité de l'esprit à la matière devait se manifester par sa non-spatialité. Par conséquent, si l'esprit ne peut être étudié par les sciences de la nature, c'est tout simplement parce qu'il n'est pas spatial. Or, cette argumentation va être reprise par Kant en y apportant toute la force de l'analyse transcendantale.

2.2 Le problème kantien

Pour bien comprendre que Kant refuse la possibilité d'une psychologie scientifique du fait de la non-spatialité de l'esprit et de la nécessaire spatialité de la science, il faut remarquer que toute synthèse objectivante doit porter sur des phénomènes spatio-temporels et pas seulement temporels. En effet, une psychologie scientifique se définit en tant que science comme une activité de synthèse objectivante de phénomènes. Mais, en tant que psychologie, les phénomènes qu'elle synthétise ne peuvent être que temporels, car le temps est la forme du sens interne. La psychologie scientifique doit donc être la synthèse objectivante de phénomènes temporels ; mais la science ne peut être la synthèse objectivante que de phénomènes spatio-temporels ; donc la psychologie scientifique doit choisir entre le fait d'être un science ou d'être une psychologie.

En effet, une synthèse objectivante ne peut s'effectuer sur des phénomènes déterminés seulement quant au temps : il ne peut exister de phénomènes seulement temporels. Une manière de le comprendre est de considérer ce que cela signifie de synthétiser un divers. Dans la déduction transcendantale subjective de la première édition, Kant distingue trois synthèse, la synthèse de l'appréhension, la synthèse de la reproduction et la synthèse de la recognition dans le concept. Ces trois synthèses constituent la synthèse objectivante. La synthèse de l'appréhension rassemble le divers successif des phénomènes en une appréhension puisque, le temps étant la forme du sens interne, le divers phénoménal est donc nécessairement donné selon la forme du temps, *i.e.* successivement :

> « Toute intuition contient en soi un divers, qui ne serait pas cependant représenté comme tel, si l'esprit ne distinguait pas le temps dans la série des impressions successives ; car en tant que *contenue dans un moment,* toute représentation ne peut jamais être autre chose qu'une unité absolue. Or, pour que de ce divers sorte l'unité de l'intuition (comme par exemple dans la représentation de l'espace), il est nécessaire d'abord de parcourir le divers et puis de le rassembler, acte que j'appelle la *synthèse de l'appréhension,* parce qu'il porte directement sur l'intuition qui présente sans doute un divers, mais qui ne peut jamais, sans l'intervention d'une synthèse, produire ce divers comme tel, et précisément comme contenu *dans une représentation.* » A 99 (Kant, 1980, p.712-713)

Cette appréhension suppose, pour pouvoir être réalisée, que l'imagination puisse reproduire les instants écoulés, les maintenir présents, pour que la synthèse puisse s'effectuer. Par ailleurs, cette reproduction doit être sélective, s'effectuer selon une règle, car il n'est pas question de tout retenir dans l'appréhension sous peine d'avoir non une perception mais une cacophonie d'impressions. Or, cette reproduction sélective correspond à la recognition dans le concept, puisque c'est le concept qui donne la règle de rétention. Tout cela est assez évident comme le souligne Kant :

> « Le mot concept pourrait déjà nous conduire par lui-même à cette remarque. En effet, c'est cette conscience une [synthèse de la recognition], qui réunit [synthèse de l'appréhension] en une représentation le divers, intuitionné successivement et ensuite reproduit [synthèse de la reproduction]. » A 104 (Kant, 1980, p.715)

En résumé, l'intégration ou synthèse consiste à rassembler les diverses apparitions ou phénomènes en un objet : il faut donc que la succession temporelle des apparitions puisse se rapporter à une même présence pour que l'intégration soit possible. Or il ne peut y avoir dans le sens interne plusieurs représentations mentales se rapportant à une même présence puisque par définition le sens interne est pure succession temporelle. Cela implique qu'il doit y avoir une présence hors du sens interne à laquelle se rapportent les représentations du sens interne, *i.e.* une simultanéité. La simultanéité ne peut prendre place que dans l'espace, puisque l'espace consiste dans la donation de choses présentes en même temps. Donc, toute synthèse objectivante, *i.e.* scientifique, ne peut porter que sur des phénomènes non seulement temporels, mais aussi spatiaux :

> « Pour Kant, la saisie de l'objectivité suppose les synthèses, c'est-à-dire une série de phénomènes extérieurs et saisie comme un objet parce que la synthèse de l'appréhension transforme la continuité dans le temps en simultanéité dans l'espace. » *in* (Rivelaygue, 1992, p. 312).

Poser ensemble des impressions successives, c'est là l'affaire de l'imagination. Mais la simultanéité ainsi obtenue des impressions nécessite un renvoi au sens externe, c'est-à-dire à l'espace. Cela signifie que l'imagination, pour assurer sa fonction de reproduction, ne peut opérer que sur des phénomènes spatio-temporels, correspondant à une présence extérieur.

> « Prendre pour une perception l'imagination d'objets extérieurs (rêver...), voilà qui est certes possible, mais à la seule condition de supposer un sens externe, c.-à-d. un rapport de notre intuition externe à des objets situés effectivement hors de nous, sans quoi toutes ces intuitions, en tant qu'au fond simplement internes, n'auraient que la forme et la dimension du temps, et non celles de l'espace, et cette forme n'est pas pensée, mais intuitionnée, c.-à-d. immédiatement rapportée à un objet, *bien que nous ne sachions pas ce que celui-ci est en soi,* mais seulement comment il nous apparaît. Si cela n'était pas, nous n'aurions même pas d'imaginations, car celles-ci ne sont que des *reproductions,* quant à la forme, d'*intuitions sensibles* d'objets extérieurs qui certes peuvent être des fictions, mais non *au sens où elles n'ont aucunement d'objets extérieurs.* » (Lose Blätter 7) *apud* (Eisler, 1994, p.520, art. imagination)

Il existe donc une corrélation profonde entre reproduction par l'imagination et sens externe, ainsi qu'en témoigne la célèbre réfutation de l'idéalisme : il ne peut y avoir de sujet interne que s'il y a une présence externe. En effet, le moi se perçoit comme une continuité au sein d'une succession temporelle. Or, la succession ne peut être perçue sans référence à la simultanéité, sans quelque chose qui demeure, sans quelque chose de permanent. Or le sujet, forme vide et non substance (cf. paralogisme de la substance), ne peut être ce point fixe et permanent. Etant lui-même la succession, le flux, le sujet ne peut être le principe d'unité et de continuité de ce flux, notamment pour que dans ce flux des appréhensions puissent s'effectuer. C'est donc que la permanence vient des données spatiales, l'espace étant le principe de la permanence et de la simultanéité. Les données spatiales dont il s'agit doivent correspondre à de véritables données empiriques concrètes, et non pas seulement la forme pure de l'espace. En tant que telle, celle-ci ne peut m'affecter : les formes pures chez Kant ne sont rien par elles-mêmes et n'existent qu'avec et que par leur contenu. En rapprochant ainsi la réfutation de l'idéalisme de la synthèse de l'appréhension, on s'aperçoit que l'imagination comme faculté de « tenir ensemble » des impressions déjà écoulées, comme faculté de maintenir une fenêtre[20] de simultanéité dans le flux de la succession temporelle, est partie prenante avec l'espace dont elle tire son principe de possibilité. C'est dire que le temps comme succession ne peut se laisser penser que comme une coprésence spatiale sur un axe orienté, et que le concept de temps est par conséquent impensable et irreprésentable. Nous retrouverons une semblable conclusion chez Husserl où le temps, selon nous, ne se laisse penser que par l'intermédiaire de l'espace.

Il est en effet impossible de « poser ensemble », de « syn-thétiser » des phénomènes dans le temps : pure succession, la prise en compte d'un élément du divers temporel implique nécessairement l'oubli de ce qui est avant et l'ignorance de ce qui vient après. Le temps étant en effet un continuum muni d'un ordre total, ce qui ne coïncide pas avec l'élément du divers soumis à l'actualité de la synthèse est obligatoirement avant ou après, donc inaccessible à la synthèse. C'est donc pour cela qu'il est impossible de distinguer l'ordre du temps du cours du temps : alors que le premier indique qu'un événement est *objectivement* avant ou après quelque chose d'autre, le second implique que nous percevons *subjectivement* quelque chose avant ou après quelque chose d'autre, sans que cet ordre subjectif de la perception implique quoi que ce soit dans la réalité temporelle objective. Kant l'explique lui-même très clairement :

> « Je vois, par exemple, un bateau descendre le cours d'un fleuve. Ma perception de la position qu'il occupe en aval du courant du fleuve est postérieure à la perception de la position qu'il occupait en amont, et il est impossible que, dans l'appréhension de ce phénomène, le bateau puisse être perçu d'abord en aval et ensuite en amont du courant. L'ordre dans la série des perceptions qui se succèdent dans l'appréhension est donc ici déterminé et cette appréhension est liée à cet ordre . Dans l'exemple précédent d'une maison, mes perceptions dans l'appréhension pouvaient commencer au sommet et finir au sol ; je pouvais aussi les faire partir du bas et m'arrêter en haut, et également appréhender par la droite et par la gauche le divers de l'intuition empirique. Dans la série de ces perceptions il n'y avait pas d'ordre déterminé qui m'obligeât à commencer par un côté ou par un autre l'appréhension pour lier empiriquement le divers. Cette règle se trouve toujours dans la perception de ce qui arrive et elle rend nécessaire l'ordre des perceptions qui se succèdent (dans l'appréhension de ce phénomène). » (Kant, 1984b, p. 185)

Or, ce qui permet de distinguer l'ordre du temps du cours du temps, c'est la simultanéité comme présence-ensemble des phénomènes dans l'espace. C'est parce que les appréhensions du tableau visent des phénomènes permanents et non changeants que la succession inhérente à l'appréhension subjective n'est pas attribuée à l'objet : il est possible de modifier le cours de l'appréhension en une perception ultérieure et néanmoins de percevoir le même objet. Lorsqu'une nécessité se glisse dans le cours de l'appréhension subjective, c'est-à-dire qu'il est impossible de percevoir le même objet quand le cours du temps se modifie, c'est alors que la succession est une propriété objective. Cela repose sur le fait que le temps lui-même ne change pas : il n'est que la forme du changement. Le temps ne peut donc être perçu que sur la base d'un substrat soumis au changement. Or un changement n'est perceptible que si du permanent en est le fondement : la succession n'est connue comme telle qu'en face du permanent. Si le temps est le principe de la succession, et l'espace, le principe du permanent, le temps ne peut être donné que par l'espace. Comme le résume Rivelaygue :

> « Pour les phénomènes du sens interne, qui n'ont qu'une dimension, notre mode subjectif d'appréhension est le seul possible ; il n'est pas constitutif d'une appréhension objective :

20. Qu'on nous pardonne cette métaphore d'informaticien !

> l'*ordre* du temps et le *cours* du temps sont indiscernables, ce qui signifie qu'il n'y a pas de synthèse objectivante ; de telles synthèses posent l'unité de l'objet à partir de ses apparitions successives et différentes, alors que les phénomènes du sens interne, n'ayant qu'une dimension et un flux continu, ne peuvent se présenter sous des faces et des profils différents, qu'il faudrait rassembler ensuite sous un concept unifiant. Au contraire, la face sous laquelle ils apparaissent est la seule possible ; leur unité est unicité, et jamais unité d'une diversité : cela signifie que, pour les phénomènes du sens interne, l'impossibilité d'une synthèse objectivante interdit la position comme objet (c'est-à-dire principe d'unité de phénomènes divers et successifs) ; » *in* (Rivelaygue, 1992, p. 313)

Cela se concrétise par le fait que les phénomènes de la conscience, du sens interne, ne constituent pas une phénoménalité, *i.e.* des phénomènes propres à la construction scientifique. Une phénoménalité se caractérise en ce que on sait distinguer ce qui vient de phénomène de ce qui est introduit par l'observateur, que ces phénomènes sont productibles à volonté, et qu'ils ne sont pas perturbés par l'observateur. Or, il est impossible d'établir des distinctions entre les phénomènes psychologiques qui soient objectives, et pas simplement des vues de l'esprit ; cela est la conséquence de l'impossibilité de faire une synthèse objectivante dans le temps seul. En reprenant l'analogie avec le jugement physique, on n'a que des jugements de perception et pas de jugements d'expérience. Par ailleurs, on ne peut faire d'expérience permettant de produire ces phénomènes à volonté et de telles expériences perturberaient de toutes manières ce qui serait à observer. Comme le remarque Kant :

> « La psychologie empirique est encore plus éloignée que la chimie même du rang d'un science naturelle à proprement parler, premièrement parce que les mathématiques ne peuvent s'appliquer aux phénomènes du sens interne et à leurs lois, car il faudrait alors tenir compte de la seule loi de continuité dans l'écoulement de ses modifications internes ; mais ce serait une extension de la connaissance qui se comporterait par rapport à celle que les mathématiques procurent à la théorie des corps à peu près comme la théorie des propriétés de la ligne droite se rapporte à la géométrie toute entière. En effet, la pure intuition intérieure où doivent se construire les phénomènes de l'âme, est le temps qui n'a qu'une dimension. Elle ne pourra jamais même se rapprocher de la chimie comme art systématique d'analyse ou théorie expérimentale, parce que le divers de l'observation interne n'est séparable que par une simple division en idée, mais ne peut se conserver à l'état séparé, ni se combiner de nouveau à volonté, de plus, il n'est pas possible de soumettre une autre sujet pensant à des expériences convenant à nos fins et l'observation elle-même altère et défigure déjà en soi l'état de l'objet observé. » *in* (Kant, 1990, p. 13)

Comme nous le verrons avec Husserl, les phénomènes du sens interne ne constituent pas une phénoménalité mais une phénoménologie. C'est donc à cette dernière que doit revenir le droit et la capacité d'étudier les connaissances phénoménologiques qui ne peuvent donc se prêter à un traitement scientifique.

2.3 Le problème kantien et la psychologie contemporaine

Cependant si l'on tient résolument au programme cognitiviste, que l'on veut être scientifique en matière de pensée et d'esprit, il faut tirer les conclusions qui s'imposent. Il n'y a donc de science possible, *i.e.* fondée sur des synthèses objectivantes du divers, que de science de l'espace. Donc, pour le projet d'une science de l'esprit, il ne sera possible de l'envisager qu'à la condition de *spatialiser l'esprit.*

Mais si l'esprit n'est pas spatial selon l'héritage cartésien des philosophies de la conscience, il faut choisir entre l'impossibilité d'une science de l'esprit et la spatialisation de l'esprit. Descartes et Kant, en refusant de spatialiser l'esprit, ont refusé de le naturaliser en en permettant l'étude par la science qui, de manière constitutive, est science de la nature. Dans une démarche inverse, on peut montrer facilement que les figures psychologiques adoptées par l'IA ont toutes coïncidé avec une spatialisation de l'esprit, la nature précise de l'espace variant selon l'attitude théorique adoptée.

Le behaviorisme[21] a spatialisé l'esprit dans le comportement, externe et public. La théorie de l'identité le spatialise dans le cerveau. Le fonctionnalisme le spatialise dans la spatialité du calcul, par l'espace mémoire de la machine de Turing. Enfin, le cognitivisme[22] physique tente de mathématiser les structures de l'esprit à partir de la perception, montrant comment le temps cérébral est un codage de situations spatiales perçues (Petitot, 1993). Ces cognitivismes ont de fait spatialisé l'esprit en recherchant tout simplement une phénoménalité sur laquelle asseoir leur légalité scientifique. Comme toute phénoménalité doit être spatiale, ils ont de fait spatialisé l'esprit. En avaient-ils le droit ?

Pour que cette spatialisation de l'esprit réussisse, il faut montrer comment les termes intentionnels caractérisant l'esprit se construisent comme des objets scientifiques à partir du divers phénoménal, et s'assurer que le terme ou concept ne suffise pas à créer l'objet, dans un fonctionnement métaphysique. C'est-à-dire qu'il faudra veiller à ce que les termes intentionnels, la notion d'esprit, ne soient pas des mots vidés de leur substance, de leur signification, pour désigner des objets scientifiques construits à partir de la phénoménalité spatiale que l'on se sera donnée. Ainsi, si un but, terme intentionnel, devient le nom d'une structure de programmation, il faudra justifier en quoi cette structure est vraiment un but et ne pas se contenter de notre volonté d'y voir un but.

La difficulté réside dans le fait qu'il est très délicat de définir ce qu'est un phénomène psychologique sans l'investir de l'objectivité qu'il faut précisément construire, sans donner au concept psychologique une existence toute métaphysique parce que non construite dans le cadre de l'expérience possible. Ainsi, de même que les psychologues cognitifs reprochaient aux behavioristes d'utiliser des termes cognitifs dans la description physique des stimuli / réponses, on peut reprocher aux cognitivistes d'utiliser les lois du sens dans la description de leurs expériences cognitives, si bien que leurs résultats présupposent précisément ce qu'il y a construire. Ce qui signifie qu'on ne peut pas observer les lois de sens pour en faire une théorie prédictive, on ne peut que les utiliser. Autrement dit, on ne peut pas décrire le sens d'une expression symbolique pour quelqu'un, typiquement un sujet d'expérience, indépendamment du sens que cette expression possède pour nous. Il n'existe pas, pour le sens, de principe de relativité qui, comme cela se produit en physique mathématique, résout le problème de la subjectivité de l'observateur.

Mais en outre, il s'agit de comprendre de quelle nature est la phénoménalité spatiale

21. Il n'est pas dans notre propos de présenter ici ces théories cognitives de l'esprit. (Lycan, 1990) présente les principaux textes fondateurs. (Bechtel, 1988) en donne un point de vue d'ensemble sur ces théories.

22. Nous appelons (Bachimont, 1994) « cognitivisme » toute tentative de naturaliser l'esprit, c'est-à-dire d'appliquer les méthodes expérimentales et théoriques des sciences de la nature à l'étude de l'esprit. Le cognitivisme peut être computationnel quand l'approche adoptée est informatique, physique quand l'approche est physique, etc. Naturaliser l'esprit laisse indéterminée la question de savoir si l'esprit est réellement ou non un objet de la nature : lui appliquer les méthodes des sciences de la nature n'implique pas qu'il soit « naturel ». Cette question n'est d'ailleurs pas une question relevant des sciences cognitives à proprement parler, mais plutôt de philosophie cognitive.

mobilisée dans le cognitivisme computationnel, de quel type sont les objets construits à partir de cette phénoménalité, pour en évaluer la portée signitive. Ainsi, indépendamment du fait de savoir si le cognitivisme computationnel parvient à spatialiser *l'esprit* ou non, il n'en demeure pas moins qu'il s'inscrit dans un *espace*. C'est cet espace qu'il faut étudier si l'on veut d'une part reconnaître la nature scientifique de l'informatique, et d'autre part y reconnaître l'esprit.

3 Epistémologie transcendantale de l'informatique

Nous avons d'une part souligné la nécessité d'élaborer une épistémologie de la logique [23] et de l'informatique, et d'autre part indiqué notre choix de recourir à une analyse de type transcendantale pour mener à bien cette élaboration.

3.1 Héritage kantien et informatique

Kant a élaboré son dispositif transcendantal principalement pour rendre compte de la physique newtonnienne, en montrant comment ses objets devaient être construits par l'application des lois mathématiques correspondant aux catégories de l'entendement. Il montre notamment comment les mathématiques sont l'*organon* de la physique.

Notre propos est de tirer les conséquences de ce dispositif pour l'informatique selon les idées maîtresses suivantes : l'informatique est une physique, une étude de la nature, mais d'un genre particulier, car elle porte sur les signes. Elle étudie les lois effectives de calcul sur les signes. Cette physique des signes possède son propre *organon*, la logique mathématique. Cela entraîne des conséquences inhabituelles en épistémologie. Contrairement à la tradition remontant à Kant, la logique mathématique n'est pas analytique, mais synthétique *a priori* ; organon et non langage, elle permet à l'informatique de construire ses objets. Nous présentons les principaux points de notre argumentation ci-après.

3.1.1 Logique formelle et logique transcendantale

On sait comment pour Kant la logique formelle était *geschlossen und vollendet :* terminée et complète. Reprenant la logique aristotélicienne, Kant voyait dans la logique formelle la logique de la pensée, la doctrine pour construire des jugements et des raisonnements. Cette logique formelle ne pouvait qu'être analytique, car penser un objet n'est pas le connaître. Il n'est en effet pas possible de donner un contenu objectif à un concept uniquement par la pensée. Il faut en effet une donation par la perception pour que la déter-

23. La logique abordée ici est la partie de la logique mathématique portant sur l'informatique théorique et la théorie de la calculabilité puisque, pour nous, la logique mathématique est l'*organon* mathématique de la physique du calcul qu'est l'informatique. Pour une bonne introduction à la logique mathématique dans une perspective informatique, on pourra consulter l'excellent (Lalement, 1990) ainsi que (Gochet & Gribomont, 1990) ou encore (Lassaigne & de Rougemont, 1993). (Yasuhara, 1971) est une bonne introduction aux problèmes de la calculabilité ainsi que (Mendelson, 1987). Autrement dit, nous limitons nos considérations concernant la logique mathématique à la théorie logique de l'informatique. Nous laissons donc en suspens par exemple la question de savoir si la théorie des catégories (Lambek & Scott, 1986) est concernées ou non par nos considérations.

mination mathématique puisse s'exercer. Par conséquent, la logique formelle, si elle donne les lois de la pensée, ne peut donner les lois de l'objet et de sa construction. Formelle, la logique est par définition indifférente à son contenu et donc elle ne peut avoir de valeur objective :

> « Or, on peut prendre note de cela, comme d'un avertissement sûr et utilisable : que la logique générale, *considérée comme un organon*, est toujours une logique de l'apparence, c'est-à-dire une dialectique. Car, comme elle ne nous enseigne absolument rien sur le contenu de la connaissance, mais seulement les simples conditions formelles de l'accord avec l'entendement, qui sont d'ailleurs tout à fait indifférentes à l'égard des objets, prétendre l'employer comme un instrument (organon), pour agrandir et élargir sa connaissance, du moins à ce que l'on allègue, ne peut aboutir à rien d'autre qu'à du bavardage, où l'on affirme avec quelque apparence, ou bien où l'on conteste à son choix tout ce que l'on veut. » A 61 / B 86 (Kant, 1980, p.126)

C'est pour cette raison que Kant lui opposait la logique transcendantale qui étudie les lois de la pensée sous l'aspect de leur valeur objective : à quelle condition une loi logique de la pensée peut-elle posséder un contenu objectif par lequel un objet n'est pas simplement pensé mais connu. D'après ce que nous avons rappelé plus haut, la réponse de Kant à la question d'une logique transcendantale est le schématisme et les principes de l'expérience.

3.1.2 Logique formelle et logique mathématique

La logique formelle de Kant, héritée d'Aristote, n'est pas la logique mathématique que nous connaissons aujourd'hui. Jusqu'à Boole et ses *Lois de la Pensée*, la logique reste l'affaire des philosophes. Avec Boole, elle devient une algèbre et dès lors, sa grande affaire est le calcul symbolique et ses règles d'inférence. Mathématiser la logique a longtemps signifié pour les logiciens le recours à davantage de précision, l'utilisation d'un langage univoque et précis, au service du même objectif que la logique formelle classique : les lois de la pensée. Le titre de l'ouvrage de Boole en témoigne manifestement, et la *Begriffsschrift* de Frege également : une écriture des concepts ou idéographie, où le formalisme syntaxique d'une langue univoque vient purifier les lois de la pensée des imperfections de la logique formelle classique.

Une telle perspective où se succèdent les logiques classiques et mathématiques dans le but de fixer les lois de la pensée nous paraît cependant discutable. A notre point de vue, la logique mathématique contemporaine ne traite pas plus des lois de la pensée que les mathématiques. En ce sens, la logique mathématique n'est pas une logique formelle : son objet n'est pas la pensée, mais le calcul. Pour comprendre la logique mathématique comme une logique de la pensée, il faut auparavant réduire la pensée à un calcul. La logique mathématique n'est pas la figure moderne de la logique formelle classique, mais un outil mathématique portant sur le calcul symbolique.

Par conséquent, nous ne rendons pas synthétique la logique que Kant appelait formelle. Mais nous considérons que le qualificatif de « mathématique » de la logique contemporaine est lourd de conséquence, et qu'au lieu de faire partager aux mathématiques l'analyticité de la logique, il convient au contraire d'étendre la synthéticité des mathématiques à la logique. Mais, puisqu'il ne s'agit pas de retomber dans la métaphysique d'où Kant nous avait fait sortir, de faire procéder l'être du penser, il est nécessaire de distinguer l'ordre de la pensée, qui possède sa logique, logique formelle donc, de l'ordre du calcul, domaine étudié par la logique mathématique et opérationnalisé par l'informatique.

De plus, la logique formelle ne peut être comprise comme une logique de la pensée, abstraction faite de tout contenu. La logique refléterait ainsi les lois de la pensée pure, c'est-à-dire pure de tout contenu. Nous pensons qu'il n'en est rien et qu'une logique de la pensée pure qui en dresserait les lois n'est pas possible, ne peut exister en tant que telle. Et cela, d'une part parce qu'il est phénoménologiquement impossible d'accéder à la pensée pure, on pense toujours à quelque chose, on pense toujours un contenu dont la nature prescrit la manière de le penser, et d'autre part parce qu'il n'existe pas d'exemple, et pour cause, d'une semblable logique de la pensée.

Dans le chapitre suivant, nous montrons l'aide de la phénoménologie husserlienne que tout contenu pensé est constitutif des lois concernant la manière de le penser. En d'autres termes, il n'y a pas de logique formelle, mais que des logiques régionales.

Par ailleurs, la logique formelle kantienne ne doit pas être assimilée à une logique de la pensée pure. La logique formelle kantienne est une logique de la science qui pose les lois, valables dans le cadre de la connaissance scientifique, concernant les objets scientifiques et leurs concepts. En effet, la logique formelle ne porte pas sur les phénomènes donnés à la sensibilité : en ce sens, elle est pure. Mais elle porte sur les lois unissant les contenus objectifs construits à partir du divers phénoménal.

Par conséquent, la logique formelle kantienne est la logique régionale de la science. Mais, contrairement à la logique mathématique, elle ne légalise pas des phénomènes, mais des concepts. En effet, nous allons montrer que la logique mathématique est l'outil mathématique de l'informatique et légalise en algorithme les phénomènes temporels. La logique mathématique est un outil de modélisation des phénomènes temporels, tout comme la géométrie permet de légaliser, modéliser, les phénomènes spatiaux. Pour dire cela en une formule lapidaire, l'informatique est la géométrie du temps, l'algorithme étant la figure géométrique temporelle. En revanche, la logique formelle définit des rapports entre objets et non des modèles de phénomènes. Elle n'est donc pas au même niveau que la logique mathématique.

Prenons un exemple informatique pour fixer les choses. La logique formelle établit des lois valant entre des contenus objectifs ou conceptuels sans préciser comment se réalisent dans le cadre de l'expérience ces lois. En revanche, la logique transcendantale précise comment ces lois se réalisent. On retrouve ces deux niveaux dans la programmation orientée objet (Masini *et al.*, 1991). Dans ce type de programmation en effet, on dispose d'une part de messages adressés à des structures informatiques, analogue de la logique formelle, et d'autre part de méthodes propres à ces structures, qui leur sont locales ou encapsulées, et qui exécutent le message reçu. Les méthodes donnent un contenu effectif aux ordres reçus : l'articulation des messages et des méthodes correspond à la logique transcendantale. Par exemple, un message peut contenir un message ordonnant une impression. Le message est conceptuel dans la mesure où il ordonne l'impression sans préciser à l'objet qui reçoit cet ordre comment il doit l'exécuter : c'est en effet l'objet qui possède une méthode spécifique lui indiquant à quoi cela correspond, pour lui, que d'imprimer. Ce qui est particulier dans la programmation objet, c'est qu'il est possible de dégager un niveau conceptuel où c'est le *même* concept qui s'applique à toutes sortes d'objets, comme par exemple le concept d'impression, et c'est l'objet qui réalise localement le rapport à l'expérience correspondant. Les messages, à l'instar de la logique formelle, peuvent donc être universels, mais ils restent « formels » dans la mesure où ils ne peuvent préciser par eux-mêmes quelle expérience les réalise. Par ailleurs, les méthodes de chaque objet correspondent à une logique transcendantale permettant de relier les concepts à un contenu. D'une certaine manière, on peut dire que l'épistémologie transcendantale est dans ces conditions une épistémologie « orientée objet » (Petitot, 1995) puisque la logique transcendantale est la méthode locale à un cadre esthétique particulier dans lequel réaliser les concepts universels.

Il en ressort que la logique formelle n'est pas universelle, mais ne vaut que pour les contenus objectifs élaborés par la construction scientifique. La logique formelle kantienne est donc la logique de la science, et même de l'histoire de la science, dans la mesure où chaque catégorie correspond à une étape du développement de la science physique : la catégorie de la quantité correspond à la cinématique cartésienne, la catégorie de qualité correspond à la dynamique leibnizienne, et la catégorie de relation correspond à la mécanique newtonnienne.

Comme l'analyse finement Philonenko, il ne faut pas tomber dans l'erreur consistant à classer Kant parmi les métaphysiciens de la pensée qui en dresseraient les lois indépendamment de tout contenu. La logique kantienne est une logique particulière, régionale, celle de la région scientifique :

> « Abusés par l'exposé systématique de la *Critique de la Raison pure*, allant du principe de l'ordre – le jugement et ses fonctions – aux catégories, de nombreux philosophes et critiques se sont imaginé que Kant « déduisait » les catégories ou formes de la pensée de la logique générale et formelle, revenant ainsi à la scolastique, dégageant d'une simple théorie de la grammaire les fonctions des sciences – on a cru que la logique transcendantale était dérivée de la logique générale. C'est tout-à-fait inexact. Kant est parti de la réflexion sur les sciences et ne s'est servi de la logique générale que comme principe d'ordre et cela est à tel point vrai qu'il fut obligé, étant donné les résultats obtenus dans la réflexion sur les sciences, de corriger les affirmations de la logique formelle. Voilà pourquoi il nous dit qu'il disposait du travail des logiciens, mais que celui-ci n'était pas impeccable. *Loin de dépendre de la logique formelle, la logique transcendantale la conditionne.* » (Philonenko, 1989, p. 111 - 112)

La logique formelle kantienne, celle dont les lois sont fixées dans la table des jugements et celle des catégories, provient donc d'une réflexion sur les sciences, et non sur les lois de la pensée pure. Mais, l'informatique et la logique mathématique s'inscrivant dans le contexte scientifique, la même logique formelle pourra convenir à l'élaboration transcendantale de la physique et des mathématiques ainsi qu'à l'élaboration transcendantale de l'informatique et de la logique mathématique. Par conséquent, tout comme J. Petitot dans sa reprise transcendantale de l'épistémologie des mathématiques, nous conserverons la même table des catégories que Kant, non parce qu'étant formelle, elle est universelle comme la pensée pure dont elle pose les lois, mais parce que nous restons dans l'aire régionale de sa validité.

3.2 Principes d'élaboration de notre épistémologie

L'élaboration d'une épistémologie transcendantale de l'informatique fondée sur le rôle synthétique *a priori* de la logique mathématique est à la lettre non kantienne dans la mesure où, comme on le sait, la logique formelle est analytique chez Kant. Mais, nous

nous inscrivons dans la perspective transcendantale ouverte par J. Petitot selon laquelle :

- on peut généraliser le dispositif transcendantal à la physique moderne, ainsi que les néokantiens (Philonenko, 1989) l'ont montré, à l'instar de l'analyse transcendantale de la relativité par Cassirer (Cassirer, 1957) ;
- il est également possible de généraliser le dispositif transcendantal aux mathématiques elles-mêmes (Petitot, 1991b; Petitot, 1992a), en rompant avec la philosophie kantienne des mathématiques ; en appliquant une interprétation transcendantale des formalismes proche de celle de Hilbert (*Sur l'infini* (Hilbert, 1972)), J. Petitot a montré (Petitot, 1991b) que l'on obtient une philosophie des mathématiques assez proches de Wittgenstein (Shanker, 1987), qui n'est donc plus du tout kantienne même si elle reste transcendantale (Wittgenstein, 1967) en son principe.

Notre thèse est que cette généralisation du transcendantal comme logique d'objectivation peut s'appliquer non seulement à la physique moderne, non seulement aux mathématiques conçues dans une optique hilbertienne, mais aussi à la logique elle-même, à partir du moment où sont mis en avant les problèmes d'effectivité et de calculabilité, c'est-à-dire un rapport à une expérience réelle.

La logique devient une région de l'objectivité, en fait celle de l'effectivité comme le montre l'informatique. Il est dès lors justifié de faire la logique transcendantale de la manière dont l'informatique rend objective, c'est-à-dire dans son contexte effective, la logique formelle. A partir du moment où, pour reprendre l'expression de M. Nivat, on s'aperçoit que l'informatique est une *physique* des signes, elle a dès lors droit, en tant que physique, à sa logique transcendantale.

C'est donc dans une optique transcendantale et kantienne qu'une épistémologie de la logique mathématique et de l'informatique est élaborée ici.

3.2.1 L'ANALYSE TRANSCENDANTALE DES MATHÉMATIQUES

La philosophie kantienne s'est construite, comme nous l'avons rappelé, à partir de la physique newtonnienne. Les mathématiques apparaissaient alors comme l'outil de la science physique : les mathématiques permettent d'expliciter, par l'étude de l'espace, le savoir *a priori* sur l'expérience.

Le principe de base est d'appliquer le dispositif transcendantal kantien, élaboré et conçu autour de la physique, aux mathématiques. Il doit y avoir par conséquent un parallélisme à établir entre le dispositif élaboré pour la physique et son application aux mathématiques. Si l'on se souvient que le dispositif transcendantal se décline en trois parties, l'esthétique transcendantale qui montre comment l'expérience est donnée, l'analytique transcendantale qui montre comment l'expérience est pensée, et le schématisme transcendantal qui montre comment les relier, il y aura une esthétique transcendantale en mathématiques consistant dans la donation des symboles concrets sur le papier en analogie avec le divers spatio-temporel, une analytique des concepts reprenant celle de Kant puisqu'elle constitue une logique de la science et enfin un schématisme consistant dans la démonstration de l'objet mathématique, la démonstration étant l'analogue mathématique de la construction de l'objet en physique.

Quand une proposition est écrite en langage mathématique et qu'elle n'est pas encore démontrée, elle n'est qu'une conjecture, où l'on se fie à son sens linguistique, en attente de la détermination objective obtenue par la démonstration. Le sens linguistique d'une proposition mathématique est l'analogue du jugement de perception en physique dans la mesure où ni l'un ni l'autre ne garantit un contenu objectif. La démonstration, en donnant un contenu proprement mathématique à la proposition, en déterminant son sens mathématique, est analogue à la subsomption du jugement de perception sous un concept de la raison pure, *i.e.* une catégorie: le sens linguistique de l'énoncé est subsumé par une catégorie de la raison, déterminant ce sens linguistique en sens mathématique. Par conséquent, le sens mathématique d'un énoncé mathématique démontré correspond, selon les analogies dressées, au jugement d'expérience.

Appliquer aux mathématiques le dispositif transcendantal revient, *ipso facto*, à leur conférer un contenu objectif, à poser une objectivité mathématique. Les mathématiques possèdent une objectivité scientifique transcendantalement construite via la construction de concept comme démonstration de proposition. Mais, dans la philosophie kantienne, il n'y rien qui ressemble à une *objectivité* mathématiques: en effet, les mathématiques ne sont conçues que comme détermination de l'intuition pure en vue de la construction de l'objectivité physique. Dès lors qu'une objectivité mathématique est dégagée se pose la question de son articulation avec l'objectivité physique, objectivité pour laquelle les mathématiques jouent le rôle que l'on a vu.

Une proposition est de fonder l'articulation de ces deux objectivités sur le continu (Petitot, 1992b). Le continu est problème immense qui suscite de nombreux travaux tant en mathématiques qu'en philosophie des mathématiques (Salanskis, 1991; Salanskis & Sinaceur, 1992; Salanskis, 1992; Salanskis & Panza, 1995) (Barreau, 1989). La proposition consiste à poser que le continu reste pour les mathématiques l'horizon transcendant de la chose en soi, l'Idée limite que les structures mathématiques permettent de penser sans jamais le construire. Comme le souligne (Salanskis, 1991) à propos de la question de l'espace, les constructions mathématiques n'épuisent pas le problème du continu et ne font que reformuler autrement cette question en approfondissant ce qu'elle peut signifier.

Par ailleurs, l'esthétique transcendantale de la physique s'installe dans le continu comme donnée immédiate et originaire, le requiert dans sa définition de l'espace et du temps comme continua. La physique excède les mathématiques en ce que ses objets se construisent à partir du continuum spatio-temporel dans l'horizon de la chose réelle en soi tandis que, les objets mathématiques se construisant à partir de la littéralité symbolique, ce continuum ne peut être pour les mathématiques qu'un horizon. Le continu possède donc une « bimodalité objective » en physique et en mathématiques, et c'est cette bimodalité qui articule l'objectivité mathématique à l'objectivité physique :

> « Si l'on s'astreint (comme il se doit) à ne poser les problèmes d'être, d'existence, de réalité, de vérité et de nécessité qu'à l'intérieur d'une doctrine de l'objectivité, alors le continu est, nous venons de le voir, doublement objectif.
>
> « Relativement à l'objectivité symbolique des mathématiques formelles il n'est pas complètement déterminable. A ce titre, comme Husserl l'a déjà profondément noté, il possède le statut d'une Idée au sens kantien du mot (d'horizon inaccessible pour la détermination complète) et son traitement comme objet conduit nécessairement à des antinomies dialectiques (les paradoxes de l'infini). Mais, relativement à la physique, c'est un donné originaire conditionnant universellement les phénomènes. Comme fondement d'une Esthétique transcendantale, le continu apparaît ainsi comme une « réalité externe » que les mathématiques

formelles doivent déterminer.» in (Petitot, 1991a, p.271).

3.2.2 L'analyse transcendantale de la logique

Nous articulons de façon semblable la logique à l'informatique. Pour cela, nous considérons l'opposition que nous avons déjà longuement fréquentée chez Kant du temps et de l'espace. Nous pensons que l'opposition entre mathématique et logique, et physique et informatique, se fonde sur l'opposition espace / temps. C'est-à-dire que le dispositif transcendantal mis en place pour comprendre les rapports entre les mathématiques et la physique portent essentiellement sur l'espace dans la mesure où le temps est toujours spatialisé, constituant une dimension géométrique de l'espace, tandis que ceux unissant logique mathématique et informatique portent sur le temps. Néanmoins, cette opposition ne peut être poussée jusqu'au bout pour la raison que nous avons déjà évoquée de l'impossibilité d'une science du temps pur. Par conséquent, c'est d'un point de vue spatial qu'il faudra aborder l'informatique et la logique comme construites autour du temps.

3.2.3 Figure et algorithme

Pour justifier ces assertions, il faut, nous semble-t-il, revenir à la déduction transcendantale subjective et considérer les rapports existant entre les concepts et l'espace et le temps de la sensibilité, notamment les rapports qui se dévoilent dans la synthèse d'appréhension, synthèse la plus importante dans la mesure où les autres ne sont posées que pour rendre celle-ci possible.

Considérons tout d'abord l'espace, et interrogeons nous sur ce que peut être l'appréhension d'un divers spatial, c'est-à-dire des impressions appréhendées comme un divers spatial, rassemblées dans l'unité d'un divers. Cela revient à se poser la question de déterminer en quoi consiste l'unité d'une intuition pure de l'espace. Or, chez Kant, l'intuition pure de l'espace correspond à la *figure* : quand de la perception on enlève l'unité conceptuelle, c'est-à-dire la détermination apportée par l'entendement, et la matière de la sensation, c'est-à-dire la donation empirique d'un divers, ce qui reste est l'intuition pure, c'est-à-dire la contribution des cadres *a priori* de la perception à la perception proprement dite. Or, cette intuition pure participant à la perception, c'est la *figure*. C'est en effet la détermination de l'intuition pure par l'imagination pour obtenir une perception, abstraction faite de la matière et du concept :

> « Quand, dans la représentation d'un corps, je sépare ce que l'entendement en pense, comme la substance, la force, la divisibilité, etc., ainsi que ce qui appartient à la sensation [...], il me reste encore quelque chose de cette intuition empirique, à savoir l'étendue et la figure. Celles-ci appartiennent à l'intuition pure, qui se trouve *a priori* dans l'esprit, comme une pure forme de la sensibilité, indépendamment même de tout objet réel des sens ou de toute sensation. » A 20 / B 35 (Kant, 1980, p. 89)

Par ailleurs, cette appréhension du divers permet en effet de comprendre le rapport entre les catégories et la sensibilité, l'espace et le temps. D'une part, l'espace renvoie à la catégorie de l'homogène : l'unité du divers spatial, *i.e.* la figure, abstraction faite de la sensation et du concept, correspond à une sommation de l'homogène dans l'espace.

C'est la détermination de l'espace par la catégorie de la quantité. Les mathématiques, en l'occurrence la géométrie, correspondent donc à la détermination de l'espace par la catégorie de l'homogène.

> « Quand donc de l'intuition empirique d'une maison, par exemple, je fais une perception, par l'appréhension du divers de cette intuition, l'*unité nécessaire* de l'espace et de l'intuition sensible externe en général me sert de fondement, et je dessine en quelque sorte sa figure conformément à cette unité synthétique du divers dans l'espace. Or, cette même unité synthétique, si je fais abstraction de la forme de l'espace, a son siège dans l'entendement, et elle est la catégorie de la synthèse de l'homogène dans une catégorie en général, c'est-à-dire de la *quantité*, à laquelle cette synthèse de l'appréhension, c'est-à-dire la perception, doit donc être entièrement conforme. » B 162 (Kant, 1980, p.180)

Par ailleurs, le temps renvoie à la catégorie de la relation : le temps se détermine en rapports de causalité (permanence et réciprocité) :

> « Lorsque (pour prendre un autre exemple) je perçois la congélation de l'eau, j'appréhende deux états, (celui de la fluidité et celui de la solidité), comme se trouvant l'un par rapport à l'autre dans une relation de temps. Mais dans le temps, que je donne pour fondement au phénomène, en tant qu'*intuition* interne, je me représente nécessairement une *unité* synthétique du divers, sans laquelle cette relation ne pourrait être donnée dans une intuition *façon déterminée* (par rapport à la suite du temps). Or, cette unité synthétique, comme condition *a priori*, sous laquelle je lie le divers d'une *intuition en général*, si je fais abstraction de la forme constante de mon intuition interne, le temps, est la catégorie de la *cause*, par laquelle je *détermine*, en l'appliquant à la sensibilité, *tout ce qui arrive, dans le temps en général, selon sa relation.* L'appréhension dans un tel événement, et par suite cet événement lui-même, relativement à la perception possible, est donc soumise au concept du *rapport* des *effets* et des *causes*, et il en est de même dans tous les autres cas. » (Kant, 1980, p.180-181)

Si l'on cherche à définir l'équivalent de la figure dans le cas de l'intuition pure temporelle, on constate que Kant réserve un traitement différent au temps et à l'espace. Il n'existe pas en effet de notion équivalente pour le temps à celle de figure pour l'espace. Si on veut néanmoins dégager une telle notion équivalente, il faut faire abstraction de tout contenu pensé, c'est-à-dire de tout concept, et de tout contenu sensible, c'est-à-dire de toute sensation, matière empirique de l'intuition, et comprendre ce qui fait l'unité de l'intuition temporelle ainsi épurée. La synthèse de l'appréhension pour le divers temporel, *i.e.* celui d'une pure succession rassemblée en unité, correspond à l'unité d'un processus. Alors que la figure fixe dans l'espace un ordre et une manière déterminée de considérer la juxtaposition des points dans l'espace, la « figure temporelle » correspond à l'ordre ou à la manière déterminée de considérer la succession des instants dans le temps. Alors que la figure spatiale détermine une certaine configuration spatiale entre les points, déterminant les relations générales topologiques comme « à côté de », « à droite », etc., la figure temporelle détermine une certaine configuration temporelle entre les instants, déterminant les relations générales temporelles comme « avant », « après ».

Notre hypothèse, celle sur laquelle nous voulons fonder notre analyse transcendantale de l'informatique, est que la figure temporelle correspond à la notion d'algorithme ou de calcul : un manière de déterminer la succession temporelle. L'unité d'un divers temporel, abstraction faite de la sensation et du concept, est l'algorithme.

On comprend bien que cette unité du processus, que nous appelons algorithme, ne saurait se passer de la notion d'espace : car il faut bien maintenir présent l'instant passé, comme étant passé, pour que l'ensemble des instants de la succession puissent être rassemblés pour constituer une unité. La notion d'espace intervient ici également à titre de synthèse de l'homogène : les instants sont sommés comme étant homogènes les uns aux autres, c'est pourquoi c'est bien une figure qui est obtenue, mais une figure temporelle et non une figure spatiale. La notion même de figure renvoyant à l'espace, cela signifie que le temps est ici compris depuis le point de vue de l'espace.

3.2.4 Une mécanique sans étendue

Rattacher le temps aux catégories de relation permet de mieux comprendre la nature de la logique et de l'informatique, sciences de la succession et des algorithmes ou figures temporelles, par opposition aux mathématiques (géométrie) et la physique, sciences de l'étendue et de la figure spatiale.

Caractériser ainsi les figures temporelles en correspondance avec les catégories de relation renvoient au rapport existant entre les catégories et les étapes et domaines des sciences mathématiques et physiques. Kant considère en effet que les catégories de quantité renvoient à la cinématique ou physique cartésienne, les catégories de la qualité à la dynamique ou physique leibnizienne, et les catégories de la relation à la mécanique ou physique newtonnienne.

Or, l'algorithmique est la science du calcul : un calcul est un mécanisme abstrait, sans matière car indépendant de tout substrat matériel ou physique, mais néanmoins effectif : il détermine toute une classe de mécanismes concrets capables de réaliser en un temps fini le calcul défini par l'algorithme. L'algorithmique est donc une mécanique sans matière, sans substrat physique. Rapportée à la pure notion d'effectivité, elle détermine la succession de manière nécessaire et universelle par des lois ne portant que sur le temps, les lois du calcul.

L'algorithmique est donc la science de la causalité mécanique déterminée d'un point de vue purement immatériel, et seulement temporel. Le point de vue spatial inhérent à l'algorithmique est introduit par la notion de nécessité et de détermination : pour qu'un processus soit identique à lui-même quel que soit son exécution, il faut nécessairement un espace d'inscription où les deux instances du processus puissent être rapportées l'une à l'autre. Sans un tel espace, la notion même d'identité est vide.

Si l'on veut marier les points de vue de l'espace et du temps, c'est-à-dire envisager une physique matérielle, une mécanique concrète et non abstraite, on obtient les catégories de la qualité qui adoptent un point de vue spatial à propos de phénomènes temporels, et un point de vue temporel à propos de phénomènes spatiaux. Par les anticipations de la perception, Kant montre que si l'on tient compte simultanément de la continuité de l'espace et du temps, les catégories obtenues sont celles du calcul différentiel, si bien que les mathématiques de la physique seront celle du calcul différentiel[24]. Il est frappant de voir que les mathématiques sans matière, c'est-à-dire sans étendue, sont le calcul : l'algorithmique. Mais dès que l'on réintroduit le point de vue du substrat, c'est-à-dire la notion d'un espace infini et continu, on obtient une mathématique du continu.

24. Si l'on s'inspire de l'histoire des sciences (Koyré, 1966), il apparaît très nettement que Galilée et Descartes se sont précisément heurté au problème de spatialiser le temps, c'est-à-dire de comprendre la continuité de temps à partir de la continuité de l'espace.

catégorie	science kantienne	informatique	physique
quantité	cinématique	temps	espace
qualité	dynamique	symbole	variation
relation	mécanique	algorithme	force

FIG. 5.1 –*Les catégories et leurs sciences associées*

Pour Kant, les catégories de la relation suppose celles de la quantité et de la qualité, celles de la qualité les catégories de la quantité ; la physique leibnizienne présuppose celle de Descartes, la mécanique newtonnienne présuppose la cinématique cartésienne et la dynamique leibnizienne. De cette manière, on fonde les lois de la mécanique sur celles du calcul différentiel et ces dernières sur la géométrie ou espace. On peut tenter semblable construction pour l'informatique : alors que l'espace est le point de vue dominant pour la physique, il s'agirait donc du temps à la base de l'informatique. Par ailleurs, la dynamique en informatique serait le niveau du symbole. En effet, un symbole en informatique est avant tout un pointeur, c'est-à-dire un élément indiquant un emplacement en mémoire. Le pointeur est dynamique dans la mesure où par essence, il n'existe que pour donner accès à ce qu'il point, donc pour être activé et parcouru. Enfin, les lois mécaniques régissant les lois de la dynamique sont les algorithmes, qui déterminent les dynamiques entre pointeurs pour en faire des programmes (voir la figure 5.1).

Alors que la géométrie est le point de vue spatial des processus dynamiques, l'informatique en constitue le point de vue temporel. C'est pourquoi, de la même manière qu'il est toujours possible de donner une description géométrique des processus, il est toujours possible également d'en donner une formalisation algorithmique. A cet égard, l'informatique est la science des processus matériels considérés en tant qu'ils sont effectifs, c'est-à-dire mécanique et temporellement fini [25].

3.2.5 Infini temporel et infini spatial

Sur la base des déterminations précédentes, il est désormais possible de caractériser la logique et l'informatique à partir des rapports établis par Petitot entre mathématiques et physique. Pour cela, on l'aura compris, nous considérons que ses analyses valent pour l'infini spatial, et qu'il faut, pour les adapter au logico-informatique, partir de l'infini temporel. Avant de nous lancer dans la traduction temporelle des propriétés spatiales, rappelons nous que chez Kant, comme nous venons de le revoir avec les anticipations de la perception, ces deux formes *a priori* de la sensibilité possèdent une structure continue infinitésimale : entre deux points de l'espace ou du temps, existe une infinité de points.

Le rapport mathématique/physique se construit autour d'une double objectivité du continu : donné en acte en physique, il reste une Idée, un horizon pour les mathématiques. Notre analyse consiste à affirmer que l'on retrouve une double objectivité équivalente pour le continu temporel entre la logique, pour lequel il reste un horizon, et l'informatique, pour lequel il est donné en acte.

25. De cette manière, en abordant le point de vue des sciences cognitives, on peut comprendre qu'un processus considéré comme étant « naturel » comme la pensée peut être compris comme un calcul ou comme une transformation géométrique : les deux approches possèdent leur propre légitimité.

En effet, alors que la physique se construit à partir de l'infini spatial donné en acte, la mathématique le poursuit à travers ses constructions, comme le calcul différentiel. Cela ne signifie pas bien sûr que le temps soit absent de la physique et ignoré par la mathématisation: mais le temps est inclus dans l'espace et on en retient le minimum de ses caractéristiques pour le fondre dans l'espace.

En revanche, l'informatique se construit à partir de l'infini temporel, donné en acte par la notion d'intervalle temporel réalisé, comme le suppose l'exécution pas à pas : on passe d'un instant ou point du temps à un autre en « sautant » l'infinité (non dénombrable) des points intermédiaires. On se donne en acte cette infinité intermédiaire, car, si elle était potentielle, comme le montre tous les paradoxes antiques, il serait impossible de passer d'un point du temps à un autre. Cet infini temporel en acte possède un nom : l'effectif ou le fini. De même que la physique veut maîtriser l'infini spatial par la mesure finie (par exemple le problème des intégrales divergentes dans les problèmes de renormalisation), l'informatique veut maîtriser l'infini temporel par la mesure finie (par exemple, la cadence de l'horloge qui donne le rythme des pas d'exécution). Et de même que les mathématiques poursuivent l'infini que se donne la physique pour l'arraisonner à ses raisonnements finis, la logique mathématique poursuit l'effectif ou l'infini temporel actualisé que se donne l'informatique pour construire ses algorithmes, qui se saute d'instruction en instruction, abolissant la distance temporelle entre les deux. L'informatique a principalement affaire au temps dont elle opérationnalise les modèles mathématiques : la logique est une théorie formelle du temps que l'informatique expérimente. J.-L. Laurière remarque ainsi :

> « [L'informatique] est alors un prolongement des mathématiques qui met l'accent sur le phénomène temps et les processus évoluant dans le temps. La solution d'un problème ayant été analysée mathématiquement, elle est traduite par un programme exécutable par un ordinateur. » (Laurière, 1987, p. 1)

De même qu'il ne faut pas hâtivement conclure de notre présentation des rapports entre physique et mathématique que le temps en est exclu, il ne faut pas conclure qu'il n'est pas question d'espace entre logique et informatique. Mais, de même que le temps est réduit à sa plus simple expression pour être rendu spatial : on en prend l'unidimensionalité pour rajouter une dimension à l'espace, l'espace est réduit à sa plus simple expression pour être rendu temporel : on le réduit à l'unidimensionalité pour y inscrire les phénomènes temporels. C'est ainsi que l'informatique réalise précisément ce que Kant pensait impossible : une physique de phénomènes inscrits dans un cadre unidimensionnel (cf. section 5-2).

Faut-il en conclure que l'on a alors levé la principale objection concernant la possibilité d'une psychologie scientifique ? Il faut alors s'interroger sur le fait que le temps considéré ici n'est pas le temps psychologique, phénoménologique, mais le temps du monde. C'est la raison pour laquelle que, s'il existe une physique du temps, c'est-à-dire une physique du calcul, cette physique du temps n'en est pas autant une psychologie authentique qui serait computationnelle.

Le doute concernant la pertinence psychologique d'une physique du calcul doit être éveillé par la définition même du temps telle que l'esthétique transcendantale la donne : infini et continu. Il est clair que ces caractérisations ne résultent que de l'attribution à la succession des propriétés de l'espace, c'est-à-dire que, alors que l'espace est compris chez Kant comme l'*espace* de la simultanéité, le temps est compris comme l'*espace* de la succession. Séparer pour les opposer le temps et l'espace revient à marquer profondément que l'espace est simultanéité, et donc que l'infini est celui d'une *juxtaposition*, et que le temps est succession, et donc que l'infini est celui d'une *succession*. Et, de même que la physique, comme physique de l'étendue ou géométrie, depuis sa fondation cartésienne jusqu'à ses figures les plus modernes [26], est donc une *physique de la juxtaposition*, l'informatique, comme physique de l'effectif, est une *physique de la succession*.

Ces physiques de la juxtaposition et de la succession reposent dans leur possibilité sur la donnée d'un espace de la simultanéité (*i.e.* l'espace géométrique) et de la succession (*i.e.* espace calculatoire). C'est donc la nature de cet espace premier, permettant de comprendre l'espace géométrique et l'espace du calcul, qui conditionne l'intelligibilité du dispositif transcendantal. Nous y reviendrons dans la dernière partie de ce chapitre.

Enfin, notons pour conclure cette section, une dernière analogie. De même que la notion de figure spatiale possède son énigme pour l'entendement, à savoir l'apparence morphologique des juxtapositions ou choses spatiales, la figure temporelle ou algorithme possédera sa propre énigme, l'apparence *mélodique* des successions ou choses temporelles. On retrouve ainsi la notion de *forme* au sens de la *Gestalt* qui, comme on le sait, traitait tant les formes spatiales que les formes temporelles. Cela signifie par parenthèse qu'il est plus pertinent de s'interroger sur les rapports entre l'informatique et la musique qu'entre l'informatique et la psychologie. Faut-il formuler le projet d'une morpho-dynamique de l'informatique dont la tâche serait d'expliquer la mélodie [27] alors que la morpho-dynamique physique tâcherait d'expliquer les morphologies de l'apparence ? L'une s'occuperait de l'ouïe et l'autre de la vision ? Il est probable que les tenants de la morphodynamique physique réfuteraient un tel partage des tâches pour la raison simple qu'ils estimeraient que la mélodie comme forme n'est pas différente en nature de la morphologie spatiale. En effet, cette morphodynamique, comme son nom l'indique, intègre une authentique dimension temporelle, même si elle est intégralement pensée comme espace. Le problème de fond vient de la matière : comme nous l'avons souligné plus haut, la prise en compte de la matière, de l'étendue implique une physique géométrique. Dès l'instant que la mélodie doit être pensée comme forme d'une matière, il sera nécessaire d'adopter une approche géométrique, morphodynamique de ce problème. Car rappelons le, ce qui fait la force, et la faiblesse de l'informatique, c'est qu'il s'agit d'une physique *immatérielle*.

3.3 Logique mathématique et informatique

L'informatique est la science du calcul effectif dont la théorie est la logique mathématique. Les objets informatiques sont les objets calculés effectivement. Ils s'installent dans l'effectivité, dans l'actualité du calcul. La logique mathématique en revanche s'installe dans la manipulation potentielle de la littéralité symbolique, ce qu'elle partage avec les mathématiques. Mais, alors que celles-ci sont pensées dans l'horizon du continu spatial, donné en acte dans la physique, la logique est pensée dans l'horizon du continu temporel, donné en acte dans l'informatique. Et alors que les mathématiques sont l'organon d'une physique du continu spatial, de l'étendue géométrique, la logique mathématique est l'organon d'une physique de l'étendue temporelle, de l'effectif et donc du discret, *i.e.* l'informatique. Tant la physique que l'informatique s'installent dans l'actualité : l'actualité de l'étendue spatiale

26. « Physics is geometry », comme le souligne Clifford.

27. Les travaux informatiques menés à l'IRCAM plaiderait en faveur d'un tel projet. Par ailleurs, toute une communauté s'est constituée autour de cette thématique ; on peut consulter un dossier spécial du *Bulletin de l'AFIA* à cet égard.

infinie en acte, l'actualité du calcul effectivement réalisé. Cette actualité reste inaccessible au logico-mathématique, traduisant ainsi l'excès de la donation empirique sur la pensée. La logique mathématique n'est donc pas cette rêverie des philosophes n'ayant qu'un rapport réflexif avec les « véritables » mathématiques, mais un outil formel puissant et productif pour élaborer une authentique physique expérimentale de l'effectivité calculatoire :

> « La logique part, en effet, d'une réflexion sur l'activité mathématique, et une réaction épidermique courante du mathématicien est de dire : « A quoi bon tout cela ? nous ne sommes pas des philosophes, et ce n'est pas en se cassant la tête sur le modus ponens ou le tiers exclu que l'on résoudra les grandes conjectures, ni même les petites ».
>
> « Voire ...
>
> « Cependant un élément nouveau, et de taille, est venu clore ce débat un peu byzantin sur l'intérêt de la logique : l'explosion de l'informatique, dans tous les domaines de la vie économique et scientifique, dont l'onde de choc a fini par atteindre les mathématiciens eux-mêmes. Et petit à petit, une évidence se fait jour : pour cette nouvelle science en train de naître, les bases théoriques ne sont autres que cette discipline si discutée : la logique mathématique. » (préface de J.-L. Krivine *in* (Cori & Lascar, 1993)

Cette physique de l'effectivité calculatoire qu'est l'informatique est une physique des signes, de leur dynamique effective en tant que types effectivement réalisés dans un ordinateur.

Pour proposer une épistémologie transcendantale de la logique mathématique et de l'informatique, notre tâche est double puisqu'il faut déterminer tant pour la logique que pour l'informatique ce qui tient lieu de phénomènes donnés à l'intuition, des formes *a priori* de la réceptivité, *i.e.* l'esthétique transcendantale, des catégories, et de leur application à l'expérience, *i.e.* les principes. Nous commençons par l'informatique puisque, comme science expérimentale, elle se prête d'emblée au dispositif transcendantal qui, rappelons-le, fut principalement conçu avec la physique newtonienne en perspective.

3.3.1 L'informatique comme physique des signes

Nous prenons comme caractérisation d'un programme le fait qu'un programme correspond à un ensemble de variables tels que les instructions du programme reviennent à définir leur assignation à des valeurs. Un programme est donc un ensemble de symboles, certains étant des variables, d'autres des valeurs, gérant les assignations des secondes aux premières. Cette caractérisation est classique et reprend celle qui est utilisée pour la logique dynamique (Harel, 1978).

3.3.1.a Esthétique transcendantale

Les phénomènes donnés dans l'informatique sont les calculs effectifs, les exécutions effectives des programmes. Ce sont les expériences menées en informatique, c'est-à-dire les mesures expérimentales effectuées. L'esthétique transcendantale de cette informatique correspond d'une part à la littéralité symbolique, le symbole comme type susceptible de recevoir une valeur, et d'autre part à l'effectivité comme forme temporelle infinie bornée en acte.

Cette esthétique transcendantale est commune à la logique et à l'informatique dans la mesure où toutes deux opèrent sur des signes. Mais, le rapport aux temps est différent : alors que dans la logique, le rapport à l'effectivité n'est que potentiel, il est actuel en informatique. Les démonstrations de la logique veulent circonscrire l'effectivité du calcul menant au résultat, mais seule la réalisation informatique permet de l'obtenir.

L'espace de présentation des phénomènes pour l'informatique est l'espace des signes, espace que nous appellerons « sémiotique ». Dans le cas de la logique, un espace que tous le monde connaît est la feuille de papier, espace bidimensionnel dans lequel les signes sont assemblés. Dans le cas de l'informatique, l'espace est l'espace mémoire de l'ordinateur, la bande infinie de la machine de Turing. Le temps de cette esthétique est bien sûr celui du parcours discursif de cet espace sémiotique, temps de la pensée pour la logique, temps du calcul effectif pour l'informatique. Nous dirons que l'espace et le temps, cadres transcendantaux de cette esthétique transcendantale du symbolique logico-informatique, sont la forme d'un divers sémiotico-discursif. L'espace sémiotique comme espace d'une esthétique remonte à Hilbert:

> « La condition préalable de l'application des inférences logiques et de l'effectuation d'opérations logiques est l'existence d'un donné dans la perception: à savoir l'existence de certains objets concrets extra-logiques qui en tant que sensations immédiates précèdent toute pensée. Pour que le raisonnement logique soit sûr, il faut que ces objets soient perçus dans toutes leurs parties et que leur occurrence, leur caractère distinct, leur succession ou leur juxtaposition se présentent en même temps que ces objets, comme quelque chose d'immédiat et qui ne se réduit pas ou n'a pas besoin d'être d'être réduit à quoi que ce soit d'autre. Telle est la conception philosophique fondamentale qu'à mon sens exigent les mathématiques et d'ailleurs toute pensée, toute compréhension et toute communication scientifiques. En ce qui concerne particulièrement les mathématiques, l'objet de notre étude sera donc les signes concrets eux-mêmes dont nous savons, du point de vue que nous avons adopté, distinguer et reconnaître la forme. » in *Sur l'infini* (Hilbert, 1972, p.228)

Hilbert évoque donc un espace sémiotique où les signes concrets sont des types de symboles: les occurrences matérielles, les tokens, se présentent directement comme des types de symboles puisque l'on peut percevoir « leur caractère distinct, leur succession et leur juxtaposition ». Ce n'est pas seulement les traces sur une feuille de papier dont il s'agit, mais des types de symboles configurés spatialement dans un espace sémiotique[28]. Dans cet espace sémiotique et ce temps discursif se présentent donc les divers phénoménaux qu'il nous faut considérer. Dans cet espace sémiotique, l'imagination empirique synthétise en ce qu'on peut appeler, par analogie avec les jugements visant la nature tels que les exposent les *Prolégomènes*, des « jugements de perception », exprimés en langage naturel. Mais ces jugements n'ont pas de contenu objectif véritable. Pour qu'ils en aient un, il leur faut être subsumés par les catégories de l'entendement. Ces jugements deviendront des jugements scientifiques lorsque, soumis aux catégories, ils seront obtenus par les principes de l'expérience et les schèmes de l'imagination. Ces principes et schèmes correspondent aux règles de démonstration et de calcul: un jugement de perception est un jugement scientifique quand il est démontré ou calculé. Une telle assertion est bien conforme à ce que l'on peut demander d'une épistémologie de la logique et de l'informatique en ce qu'elle ne heurte pas la pratique scientifique: une formule est utilisée avec certitude quand elle est démontrée; sinon elle reste à l'état de conjecture et le terme de théorème ou de résultat calculé est alors impropre pour la désigner.

28. Il faut remarquer que cela correspond directement à la différence type/token de l'informatique, où

Au jugement de perception, où le sens et le contenu objectif est de nature purement linguistique, en faisant confiance au sens des mots utilisés pour l'exprimer, on substitue un contenu logique et mathématique d'où le sens linguistique est exclu:

> « Avec le calcul logique, nous disposons d'une langue symbolique qui est capable de représenter les théorèmes mathématiques par des formules et d'exprimer les inférences logiques par des processus formels. Exactement comme nous l'avons fait quand nous sommes passé de l'arithmétique à contenu à l'algèbre formelle, nous considérons les signes ou les symboles d'opération du calcul logique détachés de leur sens contentuel. Le résultat est que nous obtenons finalement, à la place de la science mathématique à contenu, science dont l'instrument de communication est la langue usuelle, un stock de formules constituées de signes mathématiques et logiques enchaînés à la suite des uns des autres selon des règles définies ... la transition du traitement naïf au traitement formel est donc maintenant achevée : elle est accomplie d'une part pour les axiomes qui étaient à l'origine, dans une vision naïve, tenus pour des vérités fondamentales, mais qui depuis déjà longtemps dans l'axiomatique contemporaine, n'étaient plus en fait que des relations entre des concepts ; elles est accomplie d'autre part pour le calcul logique qui à l'origine n'était qu'une nouvelle sorte de langue. » *Sur l'infini in* (Hilbert, 1972, p.233-234)

Comme nous le verrons en examinant le cas de la logique, l'objectivation d'un énoncé logique se fait en deux étapes. Premièrement, la démonstration logique le constitue en objet logique. Deuxièmement la réalisation informatique lui donne l'existence concrète dans l'univers du calcul effectif. Le calcul logique, présenté à l'origine comme une langue, est en fait l'organon des jugements informatiques. Non pas analytique comme la langue, mais synthétique a priori comme organon, le calcul logique est ce qui permet de construire des objets à partir du divers sémiotico-discursif.

3.3.1.B Les catégories

Tout d'abord, nous reprenons la table des catégories kantienne. Pour les raisons invoquées plus haut, la logique formelle kantienne est une logique objective de la science : propre à la région scientifique, cette logique établit les lois valant entre les objets et concepts scientifiques, comme par exemple le concept de cause, substance, etc. Cette logique n'est pas mathématique dans la mesure où elle ne légalise pas les phénomènes temporels en algorithme mais conceptualise les algorithmes en termes de rapports objectifs. La logique formelle est par conséquent un *canon* objectif et conceptuel valant tant pour la légalisation géométrique des phénomènes spatiaux, c'est-à-dire considérés en tant qu'ils sont donnés dans l'espace, que pour la légalisation algorithmique des phénomènes temporels, c'est-à-dire considérés en tant qu'ils sont donnés dans le temps. Cette logique caractérise le raisonnement scientifique mais ne reflète pas les lois pures de la pensée. Si nous la conservons, c'est parce que l'informatique respecte ce canon formel en tant qu'elle est une science, et non parce que tout domaine, tout contenu doit s'y plier puisque tout contenu pensé doit se plier, tant que *pensé*, aux lois de la pensée. Comme nous le verrons plus tard, il ne peut exister de lois *pures* de la pensée, une logique posant en termes conceptuels ce qui est légalisé au niveau phénoménal dans le domaine. Mais le domaine est toujours constitutif de la logique qui y règne : il n'existe donc pas un rapport type / instance entre la logique

---

les algorithmes ne portent pas sur les occurrences matérielles implémentées physiquement, mais sur les symboles en tant que types de ces occurrences.

formelle et les domaines matériels. Il n'y a bien plutôt que des rapports de constitution où la forme est une abstraction de la matière.

3.3.1.c Les principes de l'expérience

Les axiomes de l'intuition correspondent à la catégorie de la quantité, c'est-à-dire les catégories de l'homogène. Ils caractérisent la mesure quand elle est effectuée : mesure dans l'espace et dans le temps dans la physique kantienne. En informatique, on dira que les axiomes de l'intuition correspondent au temps d'exécution. Il s'agit bien en effet d'une sommation dans le temps d'instant homogène. Par ailleurs, ce qui est sommé, ce n'est pas un temps physique, *i.e.* celui du chronomètre, mais celui du programme, *i.e.* le nombre de pas d'exécution. En effet, il faut se rappeler que le temps d'exécution d'un algorithme n'est pas le temps mesuré sur l'ordinateur, qui est un temps physique mesurant la vitesse du support matériel, mais le nombre d'exécution donné par l'algorithme. Ce nombre d'exécution n'est pas la complexité du programme, qui est une caractéristique intrinsèque du programme et non des exécutions, mais le nombre d'instructions exécutées pour une exécution donnée du programme. En d'autres termes, il s'agit, pour un jeu de données déterminées, d'établir le nombre de pas d'exécution effectués par le programme, étant entendu que pour un jeu de données, le programme n'effectuera pas forcément les mêmes opérations que pour un autre jeu de données.

Les anticipations de la perception caractérisent la genèse de la mesure spatio-temporelle. Pour Kant, toute sensation, empirique par définition, échappant donc à toute caractérisation transcendantale, se définit par un degré : toute sensation est une grandeur intensive dont la genèse est de nature différentielle. Toute conscience est conscience de quelque chose : l'espace et le temps ne sont pas perçus pour eux-mêmes mais à l'occasion d'un divers délivré dans l'espace et le temps. Prenons le cas limite d'une conscience pure, *i.e.* sans sensation empirique. Cette conscience pure n'est pas consciente (au sens psychologique) puisque l'espace et le temps ne peuvent être perçus. Kant pose que la variation entre la conscience pure et la conscience effective d'un contenu empirique est continue. Cela repose sur la continuité du temps : entre le moment où il n'y a pas de conscience (conscience pure) et celui où il y a conscience, il n'y a pas de saut de temps, du temps est parcouru de façon continue. Mais le temps ne peut être perçu pour lui même, il doit donc y avoir une donation empirique, dans l'espace. Par conséquent, l'irruption à la conscience d'un donné empirique implique qu'il a été construit de manière différentielle depuis une conscience pure. Si bien que toute sensation est une grandeur intensive. En informatique, on posera que la genèse différentielle correspond à la complexité du programme. L'idée est que quand une formule est présente comme objet, c'est qu'elle s'est construite depuis un néant de conscience : un système déductif (axiomes + règles) contient en puissance ses théorèmes (néant de conscience) que la dérivation effective permet de faire apparaître (conscience de l'objet).

Les anticipations de la perception se rapportent aux catégories de la qualité comprise comme intensité. Le principe est que la matière du phénomène correspond à un degré : par exemple, la vitesse est une intensité associée à chaque point d'une trajectoire. La vitesse d'un programme est donnée par sa complexité : le nombre d'instructions moyen qu'il reste à exécuter pour obtenir le résultat. Plus ce nombre est élevé, plus la vitesse est faible. La complexité rend compte de la genèse du signe produit, du signe résultat. Autrement dit, pour conserver l'analogie avec la physique, il s'agirait du nombre moyen d'itérations qu'il faudrait encore faire pour aboutir à un résultat.

Les analogies de l'expérience correspondent chez Kant aux concepts de substance, cause et réciprocité. La substance en informatique correspond à une assignation d'une valeur à une variable restant vraie au cours de l'exécution d'un programme. La causalité correspond à une instruction d'assignation. La réciprocité correspond à l'interassignation de variables.

Les postulats de la pensée empirique correspondent aux modalités: le possible, le réel et le nécessaire. Pour Kant, est possible ce qui est cohérent avec les conditions transcendantales de l'expérience, réel ce qui est donné dans les conditions transcendantales, nécessaires ce qui est impliqué dans les conditions transcendantales. En informatique, on dira que la possibilité correspond à un état d'assignation des variables d'un programme résultant d'une exécution possible d'un programme. La réalité correspond à l'état d'assignation réellement produit, et la nécessité à un état produit par toutes les exécutions possibles d'un programme.

3.3.2 La logique comme organon

Détaillons en quoi consiste le dispositif transcendantal dans le cas de la logique. Comme pour l'informatique et pour les mêmes raisons, nous reprenons les catégories kantiennes. Les phénomènes donnés ne sont plus les calculs effectifs réalisés sur ordinateur, mais les formules logiques écrites sur le papier. L'espace sémiotico-discursif de la logique ne possède pas l'effectivité de l'espace sémiotico-discursif de l'informatique. Alors que l'informatique possède une extériorité radicale dans la mesure où l'essence du calcul n'est pas son existence physique, la logique repose, comme les mathématiques, sur la spontanéité raisonnante et conceptuelle de l'esprit.

La légalisation des phénomènes, c'est-à-dire la construction de l'objet scientifique, logique en l'occurrence, correspond aux règles de la démonstration des systèmes formels. Est un objet logique une formule démontrée. Tant qu'elle ne l'est pas, tout sens qui lui est attaché correspondrait avec ce que, dans le cadre physique, Kant appelle le jugement de perception, et conférer une quelconque valeur objective à ce jugement relève de la métaphysique.

On peut se demander à présent ce qui tient lieu de médiation entre la logique et l'informatique. Si nous suivons l'analogie que nous avons dressée entre notre dispositif articulant logique et informatique et celui de J. Petitot entre mathématiques et physique, il faudrait adopter comme instance médiatrice un principe donné en informatique mais restant un horizon pour la logique, de la même manière dont J. Petitot propose de voir dans le continu une objectivité donnée en physique et inatteignable en mathématiques.

Or nous avons proposé de considérer que l'informatique se donne l'infini temporel dans la cadence d'horloge à partir de laquelle elle construit l'effectivité de ses programmes. Par conséquent, c'est l'effectivité qui constituerait l'horizon de la logique. En d'autres termes, cet horizon correspondrait au fait que les démonstrations ne sont jamais totalement détaillées dans tous leurs pas d'inférences ou de calcul : elles ne sont pas implémentées. L'objet en soi échappant à la logique correspondrait ainsi à l'effectivité dans la mesure où elle n'est jamais donnée en logique, mais toujours recherchée.

Réciproquement, il est toujours fascinant que quasiment tous les programmes intéressants, *i.e.* les phénomènes informatiques, ne possèdent pas de justification logique. Cela signifie que l'effectivité donnée par définition par ces programmes, puisqu'ils « tournent », échappe à la logique, même si tous ses efforts correspondent à circonscrire cette effectivité. Tant que les programmes feront infiniment plus de choses que les systèmes formels, tant que les informaticiens sauront avec évidence qu'ils pratiquent une science expérimentale, l'effectivité ne sera pas la réalité construite de la logique mais son horizon.

Comme Jean Petitot le souligne quant au continu spatial en mathématiques et en physique, il faut donc noter que le continu (l'effectivité) est un contenu constituant l'horizon pour les mathématiques (la logique) pour devenir la forme actuelle de donation des phénomènes de la physique (informatique). Le parallélisme se poursuit alors jusque dans la bimodalisation de l'effectivité entre logique et informatique.

3.3.3 Effectivité et logique

A l'instar de la physique kantienne, nous n'avons défini que les conditions d'une physique en général, et non les caractéristiques de la physique concrète et empirique qui correspond à la nature réelle. Il ne s'agit que d'une physique du calcul en général. Pour obtenir une physique concrète, une nature matérielle au lieu d'une nature simplement formelle, il nous faut un concept empirique le plus général possible, donnant une existence spatio-temporelle sur laquelle appliquer les constructions logico-mathématiques d'objets. Kant propose le concept de mouvement pour la physique newtonnienne. Voulant construire une physique des signes, nous allons poser un concept empirique de mouvement : mouvement non pas spatio-temporel mais sémiotico-discursif.

Il s'agit d'un concept de mouvement comme pas de calcul, inférence élémentaire. Plusieurs physiques pures peuvent alors s'édifier. Si on prend la résolution logique comme mouvement élémentaire, on obtient la programmation logique. Si on prend l'application du lambda-calcul, on obtient la programmation applicative. Et de même pour tout paradigme de calcul effectif. Ces concepts empiriques de mouvement donnent lieu à des modèles de calculabilité, et à des réalisations concrètes, en l'occurrence les machines virtuelles.

Le fait remarquable de l'informatique, ce qui lui donne une autonomie théorique forte, est que tous ces paradigmes sont équivalents, et se ramènent à la machine universelle de Turing. Il y a donc bien une physique authentique du mouvement symbolique comme calcul.

3.3.4 Informatique et métaphysique

Il faut retenir de notre épistémologie transcendantale de l'informatique que les expressions d'un langage informatique ne sont des objets que s'ils sont dérivés effectivement selon les critères des catégories schématisées et déployées en principes d'expérience. Il existe une phénoménalité intrinsèque à l'informatique qui correspond aux mouvements sémiotico- discursifs. Les concepts de l'informatique possèdent donc un sens d'objet qui leur est propre. Le contenu des concepts informatiques est obtenu en écartant le sens linguistique qu'il possèdent quand ils sont énoncés subjectivement (jugement de perception de Kant, formulation naïve chez Hilbert) et en lui substituant le sens computationnel obtenu par calcul effectif. Si on considère le sens des énoncés de l'informatique selon les critères de la langue commune, on hypostasie le sens linguistique en contenu objectif. Or, comme le dit excellemment Jean Petitot:

> « En physique [BB: informatique], en dehors d'un dispositif de détermination objective, un jugement ne possède pas de sens objectif. Il ne possède que son sens linguistique. Croire que ce sens est fondé en réalité et peut s'identifier à son sens objectif, c'est introduire subrepticement la thèse métaphysique que l'objectivité peut rejoindre l'ontologie d'une réalité en soi, indépendante. Il en va de même en mathématiques [BB: logique] pour la distinction entre conjecture et proposition. » in (Petitot, 1991b, 259)

Et nous ajouterons, il en va de même en informatique et en logique pour la distinction entre expression conjecturée (logique) ou posée (informatique) et expression démontrée (logique) ou calculée (informatique). Petitot poursuit:

> « Le sens mathématique [BB: logique et informatique] d'une proposition ne peut pas lui préexister car il est le résultat de sa détermination objective (de sa démonstration). En faire un sens préexistant c'est croire que la compréhension de la réalité mathématique peut transcender son objectivité (ce qui est la négation même du point de vue transcendantal). » Ibid.

Une première conclusion est donc que le sens informatique est autonome et construit. Une seconde conclusion, corrélée à la première, est que l'informatique n'est pas une théorie représentationnelle. En effet, de la même manière que, selon Wittgenstein (Shanker, 1987), les mathématiques ne représentent pas, ne décrivent pas, mais construisent et prescrivent le sens de leurs objets, l'informatique ne représente rien d'une réalité préexistante mais construit ses objets et leur sens.

Si l'on revient à l'IA dans la mesure où elle mobilise des représentations symboliques, cette dernière conclusion permet de préciser notre contribution méthodologique de la première partie. En effet, il faut distinguer deux sens selon lesquels l'IA manipule des représentations symboliques : d'une part en tant qu'elle manipule dans ses programmes des expressions symboliques comme assemblages construits syntaxiquement à partir d'un alphabet fini de symboles, d'autre part en tant qu'elle manipule des représentations formelles de connaissances d'un domaine. Ce que nous venons de voir au terme de notre analyse transcendantale de l'informatique, c'est que seules les représentations symboliques au premier sens possède une signification informatique. En revanche, l'informatique ne manipule pas de représentations du domaine.

Cela signifie qu'en IA on parle de représentation symbolique dans un sens impropre : strictement parlant, les expressions informatiques ne représentent rien. Cela signifie que leur statut de représentation dont on leur prête ordinairement n'a aucun rapport avec le contenu informatique dont elles seront pourvues dès qu'elles seront dérivées et intégrées dans un processus computationnel. Cela compromet la démarche méthodologique habituelle adoptée en IA que l'on peut résumer ainsi :

- des connaissances qualitatives relevant d'un domaine non mathématisé, se présentant en langue naturelle, permettent de résoudre un problème.
- bien qu'exprimées dans la langue naturelle, ces connaissances sont effectivement des connaissances et non du verbiage car (i) elles constituent la seule expression

opérationnelle, effective des connaissances du domaine, et (*ii*) elles correspondent à l'expression linguistique d'une pratique concrète, systématisée et rationalisée par une tradition technique et sociale, permettant d'assurer le succès et d'atteindre le but visé ; l'exemple prototypique de ces connaissances est la pratique médicale ;

- étant donné que toute connaissance n'a de contenu gnoséologique que dans la mesure où elle coïncide avec une représentation formelle (formalisme comme principe de modélisation), ces connaissances linguistiques doivent être formalisées si l'on veut dégager d'une part le véritable contenu gnoséologique, par delà les variations dues à la langue naturelle, et d'autre part, en vertu du rapport dénotatif entre la connaissance et le monde, dégager la structure de la réalité telle qu'elle se présente dans le domaine abordé ;
- ces connaissances qualitatives exprimées en langue naturelle peuvent se formaliser et dénoter le monde tout comme les connaissances scientifiques « officielles » ;
- l'IA est donc la science des modèles qualitatifs programmés pour la résolution automatique de problèmes ;

Mais, comme nous venons de le voir, le sens intrinsèque des expressions informatiques est indépendant de leur compréhension comme représentation de connaissances. Le niveau des connaissances, comme niveau d'explication scientifique, est donc une hypostase linguistique qui, comme le rappelle Petitot, repose sur la thèse métaphysique que l'objectivité construite rejoint une réalité préexistante linguistiquement posée.

En considérant le sens intrinsèque des objets construits informatiquement comme légalisation de phénomènes temporels (*i.e.* pris comme succession) en algorithme, l'analyse transcendantale montre que l'informatique ne possède rien en particulier qui lui confère un privilège épistémologique pour modéliser l'esprit, pour donner une psychologie computationnelle, ou pour modéliser le monde, pour donner l'IA qualitative. Si l'informatique peut intervenir pour atteindre ces deux objectifs, c'est en tant qu'outil opérationnalisant les calculs, réalisant effectivement les calculs. C'est-à-dire, pour la même raison qui suscite son implication dans toutes les disciplines modélisatrices élaborant des représentations/légalisations de phénomènes réels.

En se fiant au sens linguistique dont on revêt les objets informatiques, l'IA, dans sa figure psychologique ou qualitative, adopte une démarche métaphysique puisqu'elle réifie le sens linguistique en objets dénotés qu'elle dit alors représenter par ses programmes au lieu de s'attacher à l'objectivité effectivement construite par la science informatique.

Par conséquent, la démarche méthodologique ne peut reposer sur le principe suivant lequel des représentations symboliques, parce que d'une part elles véhiculent un sens linguistique interprétable dans le domaine et d'autre part elles possèdent un contenu opératoire parce qu'elles sont informatiques, constituent un modèle du domaine. En effet, leur contenu opératoire est indépendant de leur interprétation linguistique. Par conséquent, si néanmoins le seul moyen de rapporter des représentations au domaine est d'en passer par une telle interprétation linguistique, il faut s'assurer explicitement que le contenu opératoire, effectif, des représentations symboliques en tant qu'instructions informatiques traduisent et se traduisent dans le contenu interprétatif des représentations symboliques en tant qu'interprétables en connaissances du domaine.

Néanmoins, la valeur épistémologique du contenu linguistique phénoménologique n'est pas nulle. En effet, elle fournit à la proposition son rôle de conjecture qui guide la construction scientifique pour en faire un objet scientifique. Le contenu linguistique a donc valeur heuristique pour guider la construction : il est régulateur. Quand il est utilisé de manière déterminante, en posant comme objets scientifiques les objets linguistiques qu'il subsume, on est dans la métaphysique. Quand il est utilisé de manière régulatrice ou réfléchissante, on est dans l'heuristique de la recherche scientifique. Il en est donc du linguistique comme du biologique vis-à-vis du mécanisme physique : il est métaphysique si on utilise les concepts de vie, organisme, etc., comme des concepts déterminants, mais ils doivent guider heuristiquement la recherche des mécanismes physiques responsables de ces phénomènes. Le contenu linguistique des expressions informatiques résulte d'une interprétation effectuée par un utilisateur qui leur affecte un sens, une signification. Or, les lois de l'attribution du sens ne sont pas des lois de la nature ni de l'informatique, mais des lois sémiotiques.

4 La spontanéité de la pensée

Notre thèse est donc que l'informatique est une science expérimentale de la nature, en particulier une physique des signes, et qu'elle repose sur un outil théorique mathématique, la logique, tout comme la physique classique du mouvement repose sur l'outil mathématique qu'est le calcul différentiel. Par ailleurs, de même qu'il est possible de concevoir pour les mathématiques une objectivation de type transcendantal, la logique peut recevoir une telle analyse.

Mais un doute s'élève quant à la légitimité de ce dernier point. En effet, les différences essentielles distinguant la physique des mathématiques ne semblent pas autoriser une analyse transcendantale de même type et notre analyse n'a pas montré comment elle pouvait ignorer de telles différences.

Il est possible de formuler la difficulté d'un point de vue purement kantien. La physique est une objectivation de l'objet construit à partir du divers spatio-temporel délivré par la perception. Elle repose donc sur une réception passive de ce divers, sur lequel la spontanéité de la pensée va appliquer ses règles de synthèses. Pour qu'une objectivation de type transcendantal puisse être reconnue tant dans la logique que dans les mathématiques, il faut qu'il y ait une réception passive, une perception dans ce sens. Et c'est là que le bât blesse. Car, s'il est bien possible de voir dans le signe le divers spatio-temporel perçu, l'analyse effectuée par la suite n'a pas explicité comment cette passivité était constitutive de l'objectivation.

Si une telle explicitation n'est pas faite, on ne montre pas en quoi la science mathématique ou logique n'est pas qu'activité pure du sujet, reposant sur sa spontanéité, où penser l'objet, c'est le connaître, alors que la physique est fondée sur la critique kantienne de la métaphysique, critique construite sur l'opposition du penser et du connaître.

Notre analyse doit donc montrer en quoi nous respectons l'opposition du penser et du connaître, condition *sine qua non* de la légitimité à considérer la logique et les mathématiques comme des discours d'objectivation, c'est-à-dire comme des physiques.

Il faut donc montrer en quoi penser un objet, ce n'est pas déjà le connaître, dès lors qu'il s'agit d'un objet logique ou mathématique. Il faut montrer qu'en logique ou en mathématique, la spontanéité de la pensée n'entraîne pas celle de la connaissance.

5 La passivité de la pensée

Notre approche de la logique et de l'informatique sera donc légitime que si nous montrons que la pensée logique possède une passivité, une réceptivité qui justifie que l'on applique à son égard le dispositif transcendantal. Il nous faut donc revenir au jugement mathématique lui-même et interroger son caractère tant synthétique qu'*a priori*.

Notre conclusion sera qu'il faut se démarquer de l'approche formaliste et analytique : non seulement les jugements logiques et mathématiques sont synthétiques, ce en quoi nous révoquons l'approche analytique, mais ils le sont *a posteriori*, dans un sens à préciser, puisqu'il est patent que l'expérience mathématique se distingue de l'expérience physique ou empirique. Ces jugements reposent sur la médiation du signe externe, la médiation du sémiotique, qui permet de donner un troisième terme à l'alternative du transcendantal et de l'empirique.

5.1 Mathématiques et jugement synthétique a priori

Nous avons rappelé comment Kant a vu dans les jugements mathématiques les jugements synthétiques *a priori* nécessaires à la constitution de jugements d'expérience : c'est-à-dire à la fois nécessaire (c'est leur caractère apriorique) et empirique (c'est leur caractère synthétique : ils enrichissent notre connaissance du monde). C'est donc à ces jugements qu'il faut revenir pour comprendre d'une part leur rôle dans l'économie du dispositif transcendantal, et d'autre part s'il est possible de leur ré-appliquer ce qu'ils autorisent pour la compréhension de la nature.

Dégager le fait qu'il existe des jugements synthétiques *a priori* a permis à Kant d'une part de rendre compte de l'existence des sciences et d'autre part de définir le problème de comprendre comment de tels jugements, et donc la science, sont possibles.

> « Le trait essentiel de la connaissance mathématique pure, celui qui la différencie de toute autre connaissance *a priori*, c'est qu'elle doit progresser non par concepts, mais toujours seulement par la construction de concepts. Donc, puisque, dans ses propositions, elle doit sortir du concept pour atteindre ce que contient l'intuition qui lui correspond, alors, ses propositions ne peuvent et ne doivent jamais non plus être produites par une décomposition de concepts, c'est-à-dire analytiquement, et elles sont par suite, dans leur totalité, synthétiques. »
> Proleg, 2, II p. 663 KL

Comme nous l'avons déjà souligné, l'intuition pure correspond à une figuration dans l'espace perceptif prise en général : c'est une intuition du particulier considéré dans sa généralité. Cela signifie que le jugement mathématique, notamment la démonstration géométrique (dans le sens que cela pouvait avoir pour Kant, c'est-à-dire avec une figure sur un papier), doit reposer sur une figuration reproduisant dans sa structure les relations formelles et nécessaires imposées par les objets mathématiques mis en jeu. Mais, la synthèse des objets en une même figure, *i.e.* le fait de poser ensemble, de réunir ensemble des objets sur une même figure en fonction de la définition propre à chacun, permet de faire constater les relations entre ces objets, ces relations étant nécessaires car synthétisées sur la figure du seul fait des définitions nécessaires des objets. Comme le dira si bien Peirce [29] :

> « The first is mathematics, which does not undertake to ascertain any matter of fact whatever, but merely posits hypotheses and traces out their consequences. It is observational, in so far as it makes constructions in the imagination according to abstract precepts, and then observes these imaginary objects, finding in them relations of parts not specified in the precept of construction. This is truly observation, yet certainly in a very peculiar sense ; and no other kind of observation would at all answer the purpose of mathematics. » (Peirce, 1960, 1.240)

Ces relations figurées trouvées sur la figure mettent en relief le caractère iconique et diagrammatique du raisonnement mathématique, qui puise sa synthéticité dans la figuration diagrammatique de ses objets. Le raisonnement mathématique comme raisonnement iconique que nous analysons à la lumière de l'intuition pure kantienne est un thème particulièrement fort chez Peirce.

Peirce a non seulement insisté sur l'iconicité des raisonnements mathématiques, mais en a également tiré la leçon : tout raisonnement mathématique repose sur *l'observation,* c'est-à-dire sur un élément d'empiricité. Il faut donc comprendre l'intuition pure non comme une antériorité à l'expérience, une aprioricité s'opposant radicalement à l'aposteriori de l'expérience physique ou sensible, mais comme une expérience physique particulière, celle où l'on interprète la figure comme une traduction figurée de la nécessité inhérente aux objets mathématiques.

L'introduction de l'observation comme un élément fondamental de la pensée mathématique ne consiste pas en un retour à un scepticisme empirique ni à une théorie inductive des mathématiques à partir de l'expérience physique. Elle se veut expliquer, ce fait également pressenti par Kant, l'étonnante innovation propre à chaque raisonnement mathématique alors que l'on suit fidèlement des lois nécessaires. Les mathématiques permettent de sortir de la répétition en introduisant de la différence tout en restant nécessaires. C'est ainsi que Peirce introduit l'observation comme responsable de la synthéticité des mathématiques pourtant déductives :

> « On s'est demandé longtemps comment il était possible, d'une part que les mathématiques soient par nature purement déductives et tirent apodictiquement leurs conclusions et,

29. Certains thèmes peirciens sont récurrents dans notre travail. Il n'est pas possible dans les limites de ce mémoire de les détailler. L'œuvre de Perice est colossale et encore peu connue. L'édition de référence de ses travaux est *The Collected Papers of C.S. Peirce* en 6 volumes édités par Ch. Hasthorne et P. Weiss (1931 - 1935) et 2 volumes édités par A. Burks (1958), entreprise à l'initiative de l'université de Harvard. Cette édition est progressivement supplantée par le monumental *Writings of C.S. Peirce: a chronological edition* édité par l'université de Indiana. Seulement 5 volumes sur 30 annoncés sont déjà parus. Egalement disponible, (Peirce, 1992), récemment traduit en français au Cerf, rassemble les conférences données dans le cadre des *Cambridge Conferences Lectures* présentant les dernières réflexions de Peirce sur la logique. Les citations, quand elles sont formulées en français, sont tirées de (Peirce, 1984) et (Peirce, 1978). Bien que les études peirciennes n'en sont qu'à leurs débuts, quelques essais font références : (Hookway, 1985), (Tiercelin, 1993; Tiercelin, 1993), (Chauviré, 1995). (Thibaud, 1975) fait le point sur la logique de Peirce et ses graphes existentiels, repris dans le cadre du formalisme des graphes conceptuels de J. Sowa (Sowa, 1984). Peirce fait l'objet d'un regain d'intérêt de la part des philosophes de la logique (Hintikka, Putnam, etc.) et des mathématiques du continu (Petitot) dans la mesure où la profonde originalité de ses vues permettent de renouveler l'abord de questions difficiles comme la nature des termes généraux ou du continu.

d'autre part, qu'elles offrent une série de découvertes surprenantes aussi riches et apparemment infinies que n'importe quelle science d'observation. [...] La vérité cependant paraît être que tout raisonnement déductif, même le syllogisme, implique un élément d'observation. La déduction consiste en effet à construire une icone ou un diagramme dont les relations des parties doivent présenter une complète analogie avec celles des parties de l'objet du raisonnement, de l'expérimentation sur cette image dans l'imagination et de l'observation du résultat de façon à découvrir des relations qu'on n'avait pas remarquées et qui étaient cachées dans les parties. » (Peirce, 1960, 3.363)

La déduction consiste à introduire une figure sous le regard de l'esprit qui y puise matière à réflexion et surtout à invention. La pensée mathématique est une expérience empirique, et comme le souligne Peirce, toute connaissance vient *in fine* de l'observation :

« The first things I found out were that all mathematical reasoning is diagrammatic and that all necessary reasoning is mathematical reasoning, no matter how simple it may be. By diagrammatic reasoning, I mean reasoning which constructs a diagram according to a precept expressed in general terms, performs experiments upon this diagram, notes their results, assures itself that similar experiments performed upon any diagram constructed according to the same precept would have the same results, and expresses this in general terms. This was a discovery of no little importance, showing, as it does, that all knowledge without exception comes from observation. » (NE 4, pp. 47 - 48)

La contingence n'est pas le fardeau inévitable de toute expérience empirique, et l'expérience mathématique caractérise précisément celle pour laquelle l'observation se conduit selon les règles et les critères de la nécessité. L'analyse de Peirce, contrairement à Kant, semble privilégier davantage l'algèbre que le raisonnement géométrique : les structures formelles de l'objet mathématique peuvent être aussi bien iconifiées par la formule algébrique que par la figure géométrique.

Mais le primat de l'algèbre ne coïncide pas avec un retour au conceptuel : la formule algébrique schématise dans sa structure les rapports mathématiques qu'elle doit représenter. Tout autant que la figure, elle se laisse observer pour que l'on y découvre les nouveaux rapports nécessaires qu'elle tisse entre les objets qu'elle représente :

« Quant à l'algèbre, l'idée même de cet art est qu'elle présente des formules que l'on peut manipuler et que par observation des effets de cette manipulation on découvre des propriétés qu'on aurait pas discerner autrement. Dans cette manipulation, on est guidé par des découvertes antérieures qui sont incorporées dans des formules générales. Ce sont des schémas que nous avons le droit d'imiter dans notre façon de procéder et qui sont *par excellence les icones de l'algèbre.* » (Peirce, 1960, 3.363)

Comme nous y reviendrons abondamment dans le dernier chapitre, l'inscription en une écriture algébrique permet d'une part de capitaliser la connaissance passée, mais aussi de la prolonger en manipulant les signes selon les combinaisons autorisées par les règles constitutives de l'icone. La connaissance ne se laisse représenter et découvrir seulement dans le diagramme ou l'icone, c'est-à-dire que si l'on retrouve dans la structure les rapports de l'objet. Il se suffit pas que ces rapports soient pensés dans une description conceptuelle, il faut aussi qu'ils soient représentés, schématisés, pour que le principe synthétique puisse être mis en œuvre. Si bien que, quand elles sont exprimées de manière conceptuelle, les connaissances mathématiques ne sont pas utilisables ; elles doivent être en effet exprimées de manière schématique, au moyen de formules iconiques :

> « Une formule de ce genre [(x+y)z = xz + yz] pourrait être, il est vrai, remplacée par une règle formulée abstraitement (ici que la multiplication est distributive); mais aucune application de cet énoncé abstrait ne serait possible si l'on ne le traduisait pas en une image sensible. » (Peirce, 1960, 3.363)

En termes kantiens, on dirait qu'un authentique concept mathématique n'est mathématique que dans la mesure où il consiste en une règle de construction, *i.e.* en un schème. C'est le schématisme pur, où le concept est une règle de construction dans l'intuition pure. Cela signifie que l'on raisonne de manière générale sur une figure particulière, car on la considère non comme une construction concrète donnée dans l'intuition empirique, mais comme une construction générale donnée dans l'intuition pure. C'est le caractère intuitif dans l'intuition pure qui permet de parler de découverte mathématique, car l'observation de la figure (mais aussi de la formule) permet de découvrir de nouvelles relations générales, c'est-à-dire qu'elles valent seulement pour la figure effectivement regardée par le mathématicien, mais aussi pour toutes les figures qu'il aurait pu construire avec la même règle de construction.

> « Les icones dont la ressemblance a pour support des règles conventionnelles méritent particulièrement d'attirer notre attention. Ainsi une formule algébrique est une icone devenue telle par les règles de commutation, d'association et de distribution des symboles. Il peut sembler à première vue qu'il est arbitraire de classer les expressions algébriques parmi les icones ; qu'il serait tout aussi bien ou mieux de les considérer comme des signes conventionnels composés. Mais il n'en est pas ainsi, car une des grandes propriétés distinctives de l'icone est que par son observation directe peuvent être découvertes concernant son objet d'autres vérités que celles qui suffisent à déterminer sa construction. [...] Cette capacité de révèle une vérité inattendue est précisément ce en quoi consiste l'utilité des formules algébriques, c'est pourquoi le caractère iconique est leur caractère dominant. » (Peirce, 1960, 2.279)

Il faut bien remarquer que l'observation et la manipulation de l'icone enseignent des vérités non sur l'icone, ce qui serait somme toute banal, mais sur l'objet de l'icone. L'icone, construite par une règle donnée par le concept, est plus riche que cette règle, car exposée dans l'espace et le temps. L'espace et le temps constituent l'intuition pure, c'est-à-dire la forme *a priori* de ce que le sujet connaissant rencontre comme n'étant pas lui, comme étant son extériorité. Par conséquent, la règle mathématique permet de construire quelque chose qui, appartenant à l'extériorité, échappe au sujet, n'est pas le reflet de sa structure subjective, mais qui obéit à des lois objectives, *i.e.* de l'objectivité extérieure. Car est un objet, et non pas seulement une facette du sujet, ce qui reste le même quand le sujet change, ce qui possède ses propres lois de changement auxquelles le sujet ne peut que se plier. Par conséquent, le sujet doit observer pour apprendre les propriétés de ce qui, bien qu'il l'ait construit, n'est pas lui, et donc possède une vie autonome, *i.e.* avec ses propres lois nécessaires et objectives. C'est ainsi que Hookway affirme :

> « The key in iconicity is not perceived resemblance between the sign and what it signifies but rather the possibility of making new discoveries about the object of a sign through observing features of the sign itself. Thus a mathematical model of a physical system is an iconic representation because its use provides new information about the physical system. »

> [c'est idem pour la logique :] « Formal logic takes the structure of a well formed formula as a sign of the structure of the proposition it represents and regards rules of inference as formal rules for syntactic transformation ; it provides an iconic account (or a diagram) of the structure of arguments. » (Hookway, 1994, p. 58 - 59)

Le primat de la construction iconique et diagrammatique dans l'activité mathématique se généralise chez Peirce à la manière dont il faut procéder en matière de mathématiques :

> « établir son hypothèse en termes généraux, en second lieu construire un diagramme, soit une suite de lettres et de symboles auxquels sont associées des règles conventionnelles ou les transformations autorisées, soit une figure géométrique qui non seulement le protège de toutes confusions entre *tout* et (quelque), mais lui met devant les yeux une icone par l'observation de laquelle il découvre des relations entre les parties du diagramme autres que celles qui étaient utilisées dans la construction. Cette observation est la troisième étape. La quatrième consiste à s'assurer que l'on trouverait la relation observée dans toute représentation iconique de l'hypothèse. La cinquième et dernière étape consiste à énoncer la question en termes généraux. » (Peirce, 1967, 1147), (Peirce, 1976, III,II, p. 749) *apud* (Tiercelin, 1993, p.247)

Ceci est à rapprocher des deux modes recensés par Peirce de déduction, une appelée déduction corollarielle, l'autre théorématique. Alors que la seconde est une authentique, la première ne l'est pas dans la mesure où elle ne requière pas de construction spécifique:

> « [Le corollaire] déduit directement de propositions déjà établies sans l'aide de quelque autre construction que celle qui est nécessairement suggérée dans la saisie de l'énonciation de la proposition. » (Peirce, 1976, IV, p.288)
>
> « [Le théorématique] il faut expérimenter en imagination sur l'image de la prémisse afin d'amener les déductions corollarielles, à partir du résultat de l'expérimentation, à la vérité de la conclusion. » (Peirce, 1976, IV, p. 288) *apud* (Tiercelin, 1993, p. 248)

Il ne faut pas exagérer le parallèle avec Kant dans la mesure où l'on risque de se heurter à la critique de l'intuition par Peirce. En revanche, il faut remarquer que si Peirce peut bien se dispenser d'intuition dans sa sémiotique, c'est bien parce que l'iconicité remplit le même rôle que celui dévolu à l'intuition dans l'édifice kantien : à savoir une faculté passive de réceptivité de l'extériorité, lieu donc du synthétique et de la nouveauté. C'est à juste titre d'ailleurs que l'on peut rapprocher cela des modèles mentaux chez Johnson-Laird, où la construction du modèle dans un espace de figuration autorise la découverte synthétique de nouvelles propriétés.

5.2 Jugements analytiques et synthétiques

Avec Peirce, on va au bout de la figuration kantienne, et la logique et les mathématiques deviennent des sciences d'observation. Cela signifie que la question de l'a priori est délaissée au profit de la question du synthétique, en déplaçant les sciences aprioriques vers les sciences de l'observation, même s'il s'agit d'une observation spéciale, gouvernée par l'imagination.

> « Il ne peut y avoir de doute sur l'importance de ce problème. D'après Kant, la question centrale de la philosophie est de savoir « comment sont possibles les jugements synthétiques

> a priori? » Mais auparavant il faut se demander comment sont possibles les jugements synthétiques en général, et plus généralement encore, comment est possible un raisonnement synthétique. Quand on aura la solution du problème général la solution du problème particulier sera relativement simple. C'est le « pont aux ânes » de la philosophie. » (Fondements de la validité des lois de la logique) (Peirce, 1984, p. 260)

Or, en quoi consistaient pour Kant les jugements synthétiques? Ce problème clef de la *Critique de la Raison Pure* repose sur la distinction cardinale entre jugements analytiques et synthétiques, comme nous l'avons rappelé plus haut. Il nous faut revenir avec quelque détail sur cette opposition fondatrice puisque c'est celle que nous allons remettre en cause pour surmonter l'objection que nous avons soulevée précédemment quant à notre analyse transcendantale de la logique et de l'informatique. Kant présente ainsi ces deux types de jugement :

> « Les jugements analytiques (ceux qui sont affirmatifs) sont donc ceux dans lesquels la connexion du prédicat avec le sujet est pensée par identité, tandis que ceux dans lesquels cette connexion est pensée sans identité doivent s'appeler jugements synthétiques. On pourrait appeler les premiers jugements *explicatifs*, les seconds jugements *extensifs*, car ceux-là, par le prédicat, n'ajoutent rien au concept du sujet, mais le décomposent seulement par analyse en ses concepts partiels, qui étaient déjà pensés (quoique confusément) en lui : tandis qu'au contraire les seconds ajoutent au concept du sujet un prédicat qui n'était pas du tout pensé dans le sujet, et qu'aucune analyse de celui-ci n'aurait pu en tirer. » A 7 / B 11 (Kant, 1980, p.71)

Mais comment ajouter quelque chose au sujet d'un jugement, comment ajouter au concept quelque chose qui, bien que non pensé en lui, peut néanmoins lui appartenir et l'enrichir? Il faut donc un principe permettant de faire la synthèse du sujet et du prédicat. Poser ensemble le sujet et le prédicat dans un jugement analytique repose en sa possibilité sur le principe de non-contradiction, plus précisément sur les tautologies du type $A \wedge B \rightarrow A$. Dans les jugements synthétiques, aucune loi logique ne permet de déduire et d'affirmer avec assurance l'adéquation du sujet et du prédicat. Il faut donc trouver un principe de synthèse qui ne soit pas logique. Ce principe, c'est l'expérience :

> Il est clair « 1) que par des jugements analytiques notre connaissance n'est pas du tout étendue, mais que le concept que j'ai déjà est décomposé, et m'est rendu intelligible à moi-même ; 2) que dans le cas des jugements synthétiques, je dois avoir en dehors du concept du sujet encore quelque chose d'autre (x), sur quoi s'appuie l'entendement pour qu'un prédicat ne se trouvant pas dans ce concept-là, soit reconnu comme en faisant quand même partie. [...] Cet (x) est l'expérience complète de l'objet que je pense au moyen d'un concept A, qui ne constitue qu'une partie de cette expérience. » A 8 (Kant, 1980, p.73)

Dans les jugements empiriques, l'expérience est l'expérience empirique. Dans les jugements synthétiques *a priori*, le principe de synthèse, le (x) est l'expérience *a priori* prenant place dans l'intuition pure. En effet, tous les jugements synthétiques « ne sont possibles que par le rapport du concept donné à une intuition » et nous ne pouvons avoir une intuition pure que dans la mesure où :

> « elle consiste simplement dans la forme du sujet et sa réceptivité représentative qui le rend apte à être affecté par des objets, alors l'existence effective des jugements synthétiques *a priori* est déjà en soi suffisante pour démontrer qu'ils ne peuvent porter que sur les objets des sens, et pas plus loin que sur les phénomènes ».

C'est ici que l'on trouve précisément la réponse à notre problème, qui est de savoir où réside la distinction entre le penser et le connaître dans l'activité mathématique qui puisse légitimer l'application à son égard du dispositif transcendantal. En effet, l'intuition pure consiste « dans la forme du sujet et sa *réceptivité représentative*[30] », c'est-à-dire qu'elle repose effectivement sur une passivité et une réceptivité de la pensée. L'intuition pure est encore de l'intuition, et donc passivit000é. Mais, le fait de considérer l'intuition *pure* n'entraîne-t-il pas que cette intuition ne peut que ressortir à la spontanéité du sujet, car appartenant au sujet seul indépendamment de toute extériorité? Le problème serait donc que si l'intuition pure est encore de l'intuition, l'intuition pure n'est plus passive car ne correspond pas à une donation empirique sollicitant la faculté réceptive, mais reposant sur le sujet seul, donc sur sa spontanéité.

On a vu que, pour notre part, nous interprétons l'intuition pure comme une faculté de l'intuition, c'est-à-dire reposant sur des constructions concrètes. De même, Peirce, en interprétant les mathématiques comme une science de l'observation, insiste sur le caractère réceptif des mathématiques, qui échappent par là-même à la spontanéité du sujet. On ne saurait donc affirmer que tant les mathématiques que la logique (mathématique) sont des activités pures du sujet, mais qu'elles sont bien plutôt, comme nous l'avons vu avec Hilbert, et nous le reverrons plus tard, des pratiques sémiotiques.

On tire la conclusion que les mathématiques et l'informatique sont de l'ordre du jugement synthétiques *a posteriori*. Cela signifie que l'esprit découvre en lui des choses qui ne sont pas de lui. Cela implique que l'esprit découvre en lui des interprétations des signes et des symboles dont la dynamique interprétative est davantage du fait du signe lui-même que du sujet connaissant. Cela signifie que le sujet découvre en lui cette dynamique interprétative, et qu'il la subit en quelque sorte. On retrouve ainsi la passivité nécessaire à une authentique physique du signe.

Cela signifie que le sujet découvre en lui une dynamique interprétative, une sémiosis dont il est le siège et à laquelle il assiste. Par conséquent, la pensée nous habite malgré nous, et assister à son procès nous enseigne des jugements à son égard. Mais, ces jugements appartenant à la pensée et non au sujet, il faut donc conclure que nous sommes des pensées en mouvement et non que nous sommes l'auteur et le créateur de pensées dont nous maîtrisons le destin sans pour autant nous confondre avec elle. C'est ainsi que l'on peut affirmer avec Peirce que

> « Par conséquent, exactement comme nous disons qu'un corps est en mouvement, et non que le mouvement est dans le corps, nous devrions dire que nous sommes en pensée et non que les pensées sont en nous. » (Peirce, 1984)

Or, cette conclusion, il est également possible de l'obtenir à partir de la phénoménologie husserlienne, non bien sûr en tant qu'elle l'implique comme l'une des conséquences qu'elle revendique et assume, mais en tant que l'analyse de ses théories et ses apories permet de la déduire.

Le fait que les jugements logiques et mathématiques soient non seulement synthétiques, mais aussi *a posteriori* permet de comprendre comment de tels jugements peuvent être soumis à un devenir historique, dans la mesure même où leur principe synthétique qui

30. C'est nous qui soulignons.

permet d'expliciter en modifiant et enrichissant le sujet du jugement peut lui-même changer et être modifié. En effet, le principe synthétique est un principe renvoyant la pensée à l'extériorité, cette extériorité se concrétisant dans des espaces concrets où elle s'inscrit. C'est ainsi que l'espace mathématique a pu progressivement s'inscrire dans les champs du bord du Nil qu'il fallait réarpenter après chaque crue, dans le ciel des astres, sur les figures dessinées dans le sable, sans que jamais on puisse dire qu'il en soit l'abstraction, ou qu'il en soit perceptivement dérivé.

Mais il faut nuancer la notion d'*a posteriori*. En fait, ce terme est impropre dans la mesure il renvoie de manière privilégiée à ce qui est empirique. Or, l'observation des diagrammes peirciens ou des signes hilbertiens n'est pas seulement une expérience empirique. En effet, il ne s'agit pas de diagrammes ou signes qui se donnent comme de simples traces sur le papier : ils se donnent comme renvoyant d'emblée à une signification. Cela signie que, appartenant à un système d'écriture et ou de diagrammatisation, les signes ou diagrammes renvoient à des significations au sens des opérations que l'on peut leur appliquer au sens du système. Les signes et diagrammes prescrivent un sens qu'on leur prête dans leur observation. Par conséquent, les connaissances obtenues par leur manipulation ne sont ni empiriques, ni transcendantales, ni *a priori*, ni *a posteriori*. Il faut donc dégager un troisième terme correspondant au statut que possède ces connaissances : notre hypothèse, que nous explicitons dans les chapitres suivants, est que ces connaissances sont d'ordre techniques dans la mesure où elles sont obtenues par la médiation de la technique, ici de l'écriture. Ce que permettent d'argumenter la figuration kantienne et le diagramme peircien, c'est que la connaissance procède d'une constitution technique, ici scripturaire, et que le synthétique *a priori* n'est pas résorbé dans l'*a posteriori*, ce qui serait tourner le dos à tous les acquis kantiens, mais dans le technique, dont le statut reste encore largement à expliciter.

Enfin, affirmer que «nous sommes en pensée et non que les pensées sont en nous » implique qu'il faut également remettre en cause la notion de jugement analytique. Une explicitation enrichit toujours le concept explicité : c'est un jugement synthétique dont le cadre empirique est la pensée comme lieu de donation des contenus noétiques auxquels nous assistons. S'il faut toujours parler de spontanéité, ce n'est plus pour y voir la spontanéité contrôlée d'un sujet qui domine ses connaissances, mais pour y reconnaître une spontanéité analogue au choc de l'expérience sensible : de la même manière que l'expérience sensible est un choc mettant en branle le dispositif synthétique de l'entendement, la pensée en nous est un choc mettant en branle nos facultés noétiques. Mais que sont alors ces facultés noétiques : sont-elles des facultés relevant de la spontanéité d'un sujet sous-jacent, si bien que nous n'aurions que déplacé le problème de la spontanéité du sujet sans l'avoir pour autant résolu ? La solution kantienne est encore une fois satisfaisante. En effet, le sujet transcendantal kantien est une forme logique vide qui accompagne toute les représentations, sans en affecter le contenu, ni en tirer lui-même un contenu. Cela permet de comprendre que d'une part le sujet ne contrôle pas le jeu de la pensée en lui et que d'autre part il s'agit de sa pensée, et que les contenus noétiques sont des contenus pour lui. Le sujet est un spectateur quasi-impuissant du jeu des représentations en lui, et la dynamique de la pensée, celle lui permettant d'inférer à partir de diagrammes mathématiques, est une dynamique spontanée au sens où nous l'avons définie, c'est-à-dire une spontanéité relevant du choc de l'expérience.

Par conséquent, il n'est plus possible de tenir comme absolue la loi de l'identité comme la loi de l'être et de la pensée. En effet, puisque la pensée est un choc noétique, elle possède une nécessité à l'instar des jugements logiques et mathématiques, c'est-à-dire une nécessité reposant sur l'enrichissement d'un concept par l'intermédiaire d'une expérience possible, ici de pensée. La répétition tautologique de l'identité, affirmer que $A = A$, n'est donc en rien nécessaire de manière analytique. Et la raison en est simple : cette affirmation s'inscrit dans un temps, celui de la pensée, qui relève de l'expérience et de sa contingence inhérente. Toute expérience est contingente dans la mesure où, synthétique, elle introduit une différence par rapport au sujet du jugement, par rapport à l'objet pensé. Si bien que dans le cas des tautologies identitaires de type $A = A$, il s'agit d'un jugement où la différence introduite par la dimension empirique de la pensée s'avère être ici l'identité. Cela signifie qu'à présent, l'identité est un cas particulier de la différence.

Que l'identité soit un cas particulier de la différence, explique pourquoi toute inférence, toute pensée, tout jugement, puisse être réélaborée, reprise, questionnée, mise en doute. En effet, la pensée, le contenu nous est donné dans une expérience noétique, et nous n'en avons donc que l'apparence phénoménale, telle qu'elle nous apparaît, et sa réalité en soi reste inaccessible. Mais comme la notion même de phénomène implique qu'il renvoie à une réalité en soi dont il est la manifestation, et l'on recherche constamment à épuiser l'essence du phénomène en remontant, en exhibant la réalité, la substance dont il relève. Mais, comme par définition, cette substance échappe et n'est jamais donnée, cette recherche ne peut avoir de fin. Sans cesse, la pensée est reprise, sans cesse elle apporte des éléments nouveaux, sans cesse l'essence de la pensée s'enrichit sans jamais s'épuiser ni se délimiter.

Que l'identité soit un cas particulier de la différence implique que, si l'on réinterroge la pensée affirmant cette identité, si on la reprend, il se pourrait que cette identité ne soit plus. C'est pourquoi il est crucial de déterminer les conditions sous lesquelles la reprise de la pensée maîtrise les ressources de la différence en interdisant que des déterminations contradictoires affectent les mêmes notions. C'est ainsi, nous semble-t-il, qu'il faut comprendre les langages scientifiques et leur capacité d'objectivation. Leur embrigadement, le contrôle imposé à la reprise herméneutique des concepts ou notions, est précisément ce qui leur permet de construire une objectivité.

6 Synthéticité et écriture

Or, comment contrôler la reprise herméneutique ? Il faut pouvoir juxtaposer les différentes déterminations effectuées pour établir leurs éventuelles concordances ou contradictions. Or, juxtaposition dit espace. Il faut donc un espace où les déterminations soient inscrites pour être soumises à l'analyse de leur concordance. C'est dire qu'il faut pouvoir les *écrire* sur un support permettant leur juxtaposition. C'est-à-dire qu'il faut la juxtaposition écrite des jugements pour que leurs déroulements discursifs respectifs puissent être confrontés. Par conséquent, l'écriture constitue donc l'espace sémiotico-discursif nécessaire et suffisant à la construction d'un discours maîtrisé où la répétition sera toujours identitaire, où les expériences noétiques donneront toujours le même résultat.

L'écriture devient donc la condition de la différence comme répétition identitaire, en donnant à la pensée un espace de juxtaposition où elle peut projeter la suite de ses idées, et donc, en avoir. C'est d'ailleurs la raison pour laquelle un théorème, logique ou mathématique, doit être écrit pour exister. Son objectivité idéale commence par son objectivité matérielle sur un espace d'inscription.

On comprend alors aussi pourquoi, le fait de nier la nécessité identitaire de la pensée n'entraîne pas des conclusions fort coûteuses sur le plan épistémologique, à savoir la négation de toute nécessité idéale. Car l'écriture permet précisément la construction, au sens kantien du terme, selon les termes même décrits par J. Petitot pour les mathématiques et que nous avons repris pour la logique, d'objets idéaux : objets car, possédant des lois autonomes qui ne sont pas les lois psychologique de la pensée de l'objet, ce sont les lois de la pensée pensée, de l'objet pensé, et non de la pensée pensante, des lois noématiques et non noétiques, idéaux car non dérivés inductivement de l'expérience sensible. L'écriture permet donc véritablement une expérience idéale, et son évolution historique vers plus de rigueur formelle jusqu'à l'invention des langages artificiels ne fait que confirmer ce *telos* de toute écriture.

Il existe un moyen simple de s'en assurer : c'est la distinction type/occurrence sur laquelle toute langue ou langage semble fondée. C'est parce que l'écriture distingue le type comme classe, comme unité théorique d'un ensemble d'occurrences matérielles, que l'écriture recèle en son principe même l'idéalité. Si bien que toute manipulation formelle des occurrences matérielles est destinée à enrichir la signification du type, l'occurrence engageant dans ses turpitudes le type dont elle relève. C'est parce que chaque occurrence engage son type, que les langages formels en contrôlent l'usage. Les langages de type ne font que thématiser cette distinction, en regroupant les symboles (qui sont donc des types), en classes appelées types, pour contrôler leur usage syntaxique.

L'écriture scientifique diffère de l'écriture de la langue naturelle dans la mesure où, donnant un espace de juxtaposition à la répétition, elle en contrôle les différences pour les annuler. Ou plutôt, pour assurer que la répétition du symbole corresponde non pas un glissement polysémique du sens, mais à un enrichissement progressif du sens. Cela signifie qu'un objet mathématique conserve toujours la même signification, et les propositions qu'il vérifie sont valables pour ses occurrences futures. Mais l'évolution des mathématiques montre qu'à l'occasion de ces occurrences futures, le sens de l'objet peut s'enrichir de nouvelles propositions qu'il vérifie. Ces nouvelles propositions ne peuvent ni ne doivent s'opposer ou contredire les vérités passés : au lieu de s'ajouter simplement aux précédentes, elles correspondent plutôt à leur explicitation, leur mise en perspective, à leur développement.

L'écriture scientifique se caractérise donc par le fait que la répétition doit aller dans le sens d'une conservation d'un sens univoque des objets, tout en autorisant une compréhension de plus en plus profonde de ces objets en enrichissant leur interprétation mathématique au fur et à mesure de leur utilisation et mise en perspective dans les différentes théories mathématiques.

En d'autres termes, l'écriture scientifique est conçue pour que la dérive interprétative des écrits commentant ou explicitant les écrits précédents en une suite sans fin aille dans le sens d'une compréhension plus profonde des notions mises en jeu, et non dans le sens d'une diversification du sens. En un mot, l'écriture scientifique donne la place à une herméneutique de la compréhension : la science peut être univoque et herméneutique, car vouloir dire la même chose n'implique pas que toutes les manières de le dire sont équivalentes ni qu'il existe une meilleure manière de dire qu'une autre, ni même qu'il existe une manière de dire qui épuise le contenu d'une notion. Comme le note J. Petitot :

> « De théories en théories, les mathématiques *s'auto-interprètent* indéfiniment et c'est pourquoi la déduction n'y est que locale [...]. *Il existe donc comme une « herméneutique*

intrinsèque » des mathématiques pures. » (Petitot, 1991a, p. 277)

On peut s'interroger sur la manière dont l'écriture scientifique contrôle la dérive interprétative à laquelle toute écriture s'expose en tant que telle. Nous avons vu, avec l'esthétique kantienne et la sémiotique peircienne, que l'écriture scientifique est d'essence diagrammatique et que la manière d'écrire un objet est une manière constructive de le décrire : le diagramme reflète dans son aspect scripturaire les propriétés de l'objet. L'enrichissement scientifique des notions consiste alors à conserver le renvoi diagrammatique de l'objet tout y ajoutant de nouveaux rapports diagrammatiques; comme on l'a vu, le diagramme soumet à l'observation des rapports présents alors qu'ils n'ont pas participé à la constitution du diagramme lui-même; les expliciter par de nouveaux diagrammes permet alors d'enrichir le sens de l'objet en vue duquel le premier avait été construit.

C'est donc la constructibilité de l'objet qui commande l'écriture scientifique et sa dérive interprétative. Ce qui relance en permanence l'élaboration, c'est la perspective de concevoir de nouveaux moyens de constructions à partir des résultats scripturaires déjà accumulés. La question fondatrice de l'interrogation scientifique est donc la construction de l'objectivité. C'est la raison pour laquelle la logique formelle comprise comme une logique de la science manipule des contenus objectifs, des concepts d'objets, et qu'elle peut être formalisée en suivant les principes d'une sémantique formelle. En effet, bien que l'objectivation scientifique, *i.e.* la construction des objets, ne soit pas un procédé logique, dont la logique formelle puisse rendre compte, son résultat, à savoir les objets, peut être soumis à une telle logique. C'est sans doute la raison pour laquelle la logique formelle a pu et peut encore être prise comme un modèle épistémologique de l'objectivité scientifique.

La contrainte d'obtenir un objet effectif au terme de l'objectivation scientifique entraine une notion de finitude de la construction diagrammatique. En effet, il est nécessaire qu'une connaissance mathématique permette à travers son expression de construire l'objet dont elle est la connaissance. Il faut donc que l'expression renvoie à une méthode finie de construction pour que l'objet reçoive un statut ontologique effectif. La dérive interprétative donnant lieu à un approfondissent infini du sens de l'objet passe par l'enrichissement progressif de ce sens en accumulant les méthodes finies de construction. L'objet garde son sens univoque mais néanmoins toujours ouvert dans la mesure où les énoncés véhiculant sa connaissance doivent, pour être scientifique, permettre sa construction. Mais, on a vu, à l'aide des rapports entre logique et informatique, que la construction logique est *en principe* finie et effective, mais que seule l'informatique en donne la preuve matérielle. En d'autres termes, l'effectivité de la construction reste un horizon.

En conclusion, l'écriture scientifique est une écriture diagrammatique exprimant la connaissance de l'objet. D'une part, l'écriture en tant que telle est la condition de possibilité de la connaissance scientifique dans la mesure où seule elle offre l'espace scripturaire permettant la juxtaposition des occurrences et constituant ainsi le type idéal. D'autre part, l'écriture est diagrammatique dans la mesure où les rapports diagrammatiques des expressions écrites renvoient aux propriétés de l'objet et permettent ainsi de le construire. Enfin, la construction de l'objet entraîne la dérivation de nouveaux énoncés dont l'observation permet de constater de nouvelles propriétés vérifiées par l'objet visé.

Chapitre VI

Sens et Phénoménologie transcendantale

Dans le cadre de ce travail, nous avons proposé de voir dans l'IA la rencontre de l'effectivité computationnelle et de l'interprétabilité (orthothéticité) linguistique, en voyant dans la première la base de ses méthodes et dans la seconde le lien avec un domaine de spécialité. Nous avons vu dans le chapitre précédent, en entreprenant une analyse transcendantale de l'informatique, que cette dernière était une science expérimentale de la nature et qu'elle avait pour objet d'étude empirique les algorithmes. Cela signifie que l'informatique ne permet pas de construire le sens linguistique des propositions qu'elle manipule, mais qu'elle en construit le sens « opératoire » ou « computationnel », c'est-à-dire leur signification en tant qu'elles correspondent à des instructions de programmation. Par conséquent, si les modèles informatiques de l'IA sont des modèles d'un domaine, ce n'est pas parce qu'ils sont par eux-mêmes des modèles, mais parce qu'ils opérationnalisent une représentation du domaine.

Nous avions constaté dans notre introduction et dans la première partie le rôle constructif de l'informatique dans l'élaboration des connaissances du domaine. L'intervention informatique est essentielle dans la mesure où ces connaissances resteraient inexplicitées si l'outil informatique n'avait pas prêté son effectivité pour concrétiser des inférences symboliques dont l'interprétation linguistique dans les termes du domaine correspond à l'explicitation de nouvelles connaissances dans le domaine. Notre analyse transcendantale de l'informatique, que nous avons effectuée dans le chapitre précédent, a tenté de montrer que l'informatique est une science expérimentale, une physique du calcul qui constitue de nouvelles connaissances. Mais ces nouvelles connaissances correspondent à une construction scientifique au sens kantien, comme nous l'avons exposée, c'est-à-dire à de nouvelles connaissances sur les calculs, et n'ont ainsi pas de rapport particulier avec la constitution de nouvelles connaissances dans un domaine non formel. En d'autres termes, s'il apparaît bien que l'informatique n'est pas analytique dans le sens où elle peut être au principe de la découverte et de l'élaboration de nouvelles connaissances, ces dernières correspondent à un savoir computationnel et non à un savoir dans un domaine non formel. Par conséquent, nous sommes toujours confronté au problème de comprendre quel rôle l'informatique joue quand elle permet de constituer les connaissances synthétiques *a priori* d'un domaine à travers l'élaboration d'une ontologie matérielle.

L'objectif de ce chapitre est d'aborder la question de savoir si le rôle de l'informatique est contingent ou nécessaire dans l'élaboration des connaissances du domaine. Autrement dit, notre objectif est de déterminer dans quelle mesure les connaissances reposent pour leur élaboration sur une médiation technique et, en ce qui nous concerne ici, sur une médiation informatique. Si, comme nous allons tenter de l'argumenter, toute connaissance repose pour sa constitution sur une médiation technique, notre hypothèse et notre proposition quant à la nature de l'IA sera que cette dernière est la discipline ayant pour objet la constitution de connaissances à travers la médiation du calcul symbolique.

Dans ce chapitre, nous mobiliserons la phénoménologie transcendantale husserlienne pour étudier les rapports entre l'informatique et la connaissance à l'aune de deux questions directrices. D'une part, nous nous poserons la question de savoir si la possibilité de constituer de nouvelles connaissances grâce à un recours à l'outil informatique ne repose pas sur le fait que nous pensons selon des règles et des principes analogues à ceux de l'informatique. Dreyfus (Dreyfus, 1972; Dreyfus, 1991a) a en effet montré comment la phénoménologie husserlienne pouvait être interprétée comme une démonstration du fonctionnement computationnelle de la pensée. Il nous faudra reprendre son analyse pour la discuter. Il s'avérera qu'il est difficile de trancher et que nous achopperons sur des ambiguïtés que nous ne pourrons tenter de lever qu'en nous reportant aux *Leçons sur la conscience intime du temps*. L'étude de ces *Leçons* nous permettront néanmoins de poser l'hypothèse que la pensée ne possède pas de structure computationnelle mais qu'en revanche elle doit mobiliser, pour sa propre effectivité, un recours à une transcendance externe, l'outil technique. Si bien que l'hypothèse, qu'il nous faudra reprendre dans le chapitre suivant, à laquelle conclut ce chapitre est que le rapport de constitution entre l'informatique et les connaissances repose sur le fait que la pensée trouve dans l'outil informatique des signes lui permettant de constituer ses contenus. Cet outil est particulier puisque, comme on l'a vu, le signe informatique est physique et dynamique. Il nous faudra y revenir.

Le plan du chapitre est donc le suivant. Dans un premier temps (section 6-1), nous nous interrogeons sur la nécessité d'adopter une questionnement de type phénoménologique, et tentons de préciser quelques critères qu'une théorie phénoménologique doit vérifier pour nous assister dans les questions entreprises ici. La tradition philosophique mobilisée habituellement dans les problématiques de type IA renvoie principalement à deux théories phénoménologiques, la théorie husserlienne et la théorie peircienne. Cependant, il nous apparaît que la phénoménologie husserlienne est davantage propre à aborder un problématique de constitution de la connaissance et c'est donc celle-ci que nous mobilisons pour aborder notre problème.

Nous abordons dans un premier temps (section 6-2) la phénoménologie husserlienne sous le point de vue de son assimilation à l'IA, puis dans un second temps (section 6-3), les ambiguïtés rencontrées dans la phénoménologie qui rendent délicate l'interprétation de son éventuel computationnalisme, pour enfin (section 6-4) aborder l'analyse husserlienne de la temporalité pour d'une part tenter de lever les ambiguïtés rencontrées et d'autre part envisager le problème de la constitution des connaissances par une médiation externe.

1 Phénoménologies

Puisque l'IA résulte de la rencontre de l'informatique et de l'interprétabilité en termes de connaissances, nous avons traité de l'aspect informatique dans le chapitre précédent en faisant appel à une théorie transcendantale de la science, et il nous faut à présent traiter du second aspect et pour cela nous avons besoin d'une théorie de la connaissance. La question qui pose alors est de déterminer quel type de théorie de la connaissance il nous faut.

A partir du moment où notre réflexion consiste dans la volonté de comprendre un fait, le constat positif que nous avons efectué dans l'introduction, fait qui ne correspond pas à ce qui est habituellement affirmé et pensé dans le domaine, il importe de se donner les moyens de décrire ce fait pour en prendre la mesure sans le contaminer de perspectives théoriques qui peuvent l'altérer et compromettre ainsi sa compréhension. En d'autres termes, si ce fait consiste dans le fait que l'ordinateur permet de constituer des connaissances, il nous faut une théorie descriptive de la constitution des connaissances.

Une phénoménologie de la connaissance est une attitude philosophique qui consiste par essence à considérer les choses « en elles-mêmes » indépendamment de tout préjugé. La phénoménologie a pour tâche de comprendre ce que cela signifie que de connaître un objet et d'élucider comment la pensée peut s'y rapporter. Connaître un objet, c'est, indépendamment du problème de la véracité de la connaissance, se rapporter à un objet, c'est-à-dire se rapporter à un objet qui n'est pas identique à la pensée qui le pense. Il faut donc s'attaquer au double problème que d'une part les objets de la pensée sont donnés à la pensée, c'est-à-dire qu'il faut les décrire tels qu'ils apparaissent, et que d'autre part, en tant qu'objets, ils possèdent une autonomie, une objectivité qui fait qu'ils sont ce qu'ils sont indépendamment de leur manière d'apparaître. En caractérisant les choses telles qu'elles sont et telles qu'elles apparaissent, une phénoménologie doit comprendre comment les objets de la pensée sont présents *à* l'esprit puisqu'ils sont donnés en effet à l'esprit dans la mesure où ils lui apparaissent, sans être pour autant présents *dans* l'esprit puisqu'ils sont ce qu'ils sont par eux-mêmes, du fait de leur objectivité.

On aura reconnu la problématique de l'intentionnalité[1], qui se trouve donc être au cœur de toute phénoménologie. L'intentionnalité est en effet cette propriété que possède l'esprit de pouvoir viser des objets qui ne sont pas lui, qui lui sont différents. En particulier, le problème de l'intentionnalité est celui de comprendre comment, à partir de sa manière d'apparaître, de la manière dont il se donne à la pensée, la pensée se rapporte à un objet qui, comme objet, est pour lui et non pour la pensée. Il faut donc comprendre comment à partir de la subjectivité de la donation il est possible d'accéder à l'objectivité du donné.

Il en ressort que toute phénoménologie de la connaissance doit être une pensée de la différence et de l'altérité : comment l'esprit peut-il sortir de lui-même pour accéder à ce qui n'est pas lui ? Comment à partir de ce qui lui est propre, la manière d'apparaître, peut-il connaître ce qui lui échappe, l'objectivité ? Autrement dit, toute phénoménologie est une une problématique de la constitution et, puisque la théorie de la connaissance que nous recherchons doit pouvoir rendre compte de la constitution de connaissances par la médiation informatique, c'est une phénoménologie de la connaissance qui nous est nécessaire. Précisons à présent ce que doit vérifier une phénoménologie pour être adaptée à nos objectifs.

Une phénoménologie de la connaissance doit rendre compte qu'une connaissance n'est telle que si elle vaut pour tous. Une connaissance se veut de droit universelle et nécessaire,

1. L'intentionnalité est un problème fondamental rencontré par toutes les disciplines qui, d'une manière ou d'une autre, ont affaire avec une théorie du sens. Nous n'en parlerons que depuis le cadre de la phénoménologie husserlienne. Mais bien d'autres manières d'en parler et de le théoriser sont possibles, voir par exemple (Pacherie, 1993).

même si de fait elle ne l'est pas. C'est ce que Kant a découvert dans le jugement de goût : bien que non scientifique, le jugement de goût est posé *comme* un jugement scientifique dont il veut partager l'universalité et la nécessité. Une connaissance ne peut être privée : elle doit pouvoir être accessible à tous.

Une « bonne » (c'est-à-dire adaptée à nos besoins) phénoménologie de la connaissance doit être en outre une phénoménologie du sens et non du fait positif. Si en effet le projet phénoménologique est un projet descriptif exempt de préjugé, il ne peut être question de se donner un fait préalable ni de renvoyer l'apparition du sens objectif ou de la constitution de connaissances à la positivité de données qui serviraient en quelque de sorte de matériau de construction. Car il faut bien comprendre que tout fait positif est une annulation du sens. Le fait positif *est*, il est là, pour lui-même, indépendamment de tout. La positivité est une notion ontologique : à proprement parler, le seul fait positif possible ne peut être Dieu lui-même. D'un point de vue phénoménologique, le fait ne peut être positif dans la mesure où l'enjeu de la description phénoménologique est précisément de montrer comment l'esprit accède à un objet comme à un fait positif, comment il peut poser un objet comme étant un fait positif. La description phénoménologique décrit le fait positif comme étant « positif pour l'esprit », comme étant ce qui apparaît à l'esprit comme fait positif. En d'autres terme, la positivité, pour la phénoménologie, est un mode d'apparaître à l'esprit, et non une caractéristique ontologique. Il y a donc une tension entre phénoménologie et ontologie, dans la mesure l'absolu ontologique devient une modalité de l'intentionnalité phénoménologique.

En conclusion une bonne phénoménologie doit :

- être une phénoménologie de la conscience ou de l'esprit, qui décrit comment s'élabore la connaissance comme rapport à l'altérité et la différence ;
- rendre compte de la publicité des connaissances, c'est-à-dire de leur prétention à être d'une part nécessaires et universelles et d'autre part accessibles à tous ;
- être une phénoménologie du sens et non du fait ; le fait n'a pas de sens, il ne peut être objet de connaissance puisqu'il n'existe que pour lui-même indépendamment du reste du monde ; le fait est une monade sans porte ni fenêtre.

Traditionnellement, en IA, deux théories phénoménologiques sont régulièrement mentionnées comme arrière-plan philosophique : la phénoménologie husserlienne et la phanéroscopie peircienne. En effet, outre la phénoménologie husserlienne dont nous allons voir la proximité à l'IA dans la section suivante, la philosophie de Peirce est fortement mobilisée dans les travaux de J. Sowa qui fonde ses graphes conceptuels en référence explicite aux graphes existentiels de Peirce (Thibaud, 1975) et ses travaux ontologiques (au sens IA) (Sowa, 1995) sur les catégories sémiotiques peircienne (Peirce, 1960).

La phénoménologie transcendantale husserlienne et la sémiotique phanéroscopique peircienne sont toutes deux descriptives : elles considèrent toutes deux les choses telles qu'elles sont, l'une en s'attachant aux vécus de la conscience pour élucider comment ils constituent dans leur immanence à l'esprit une objectivité qui le transcende, l'autre en s'attachant à la manière dont les signes, s'interprétant les uns les autres, constituent un sens dont aucun n'est le dépositaire.

La phénoménologie husserlienne possède l'avantage de consister en une problématique explicite de la constitution. Elle nous paraît ainsi plus propre à nous assister dans le traitement de notre problème. En effet, elle est avant tout une phénoménologie de la conscience. Son but est de décrire comment un sens peut se constituer pour la conscience à partir des données qu'elle contient de par sa vie sensible. Le problème est en effet le suivant : comment, à partir des données en quoi elle consiste, c'est-à-dire des vécus, peut-elle constituer des objets qu'elle n'est pas ? L'objectif de la phénoménologie husserlienne est de décrire les choses telles qu'elles sont et telles qu'elles apparaissent, en montrant comment elles se constituent à partir de données *internes* à la conscience, immanentes.

2 Husserl et l'intelligence artificielle

Outre le fait de développer une phénoménologie de la constitution, Husserl est important pour notre discussion dans la mesure où il a été vu à la fois comme le précurseur et comme le penseur du cognitivisme computationnel. Celui qui a le mieux défendu ce point de vue est sans nul doute H.L. Dreyfus [2], qui inscrit sa critique de Husserl et du cognitivisme à la suite de la reprise de la phénoménologie transcendantale dans une perspective logico-analytique par Follesdal.

Le principe de base est que la distinction que fit Frege entre le sens et la référence (Frege, 1971b), appliquée à la langue, correspond à la distinction husserlienne entre la manière de viser un objet, la matière intentionnelle, et l'objet visé, appliquée quant à elle à la pensée. Alors que les distinctions frégéennes sont linguistiques, celles de Husserl seraient mentales et psychologiques. Ce parallélisme avec Frege permet de projeter sur Husserl les catégories conceptuelles élaborées par la postérité frégéenne, *i.e.* les mouvements de pensée philosophique issus du renouveau de la logique suscité par Frege. C'est ainsi que l'on voudra retrouver chez Husserl des caractérisations de la pensée, des contenus mentaux, de la vérité, etc., empruntées au positivisme logique et au cognitivisme computationnel (voir *infra*).

Dans l'analyse critique de Husserl qui lui permet d'y voir la figure continentale de Fodor, Dreyfus considère que les deux phases de la théorie husserlienne de l'intentionnalité, celle des *Recherches Logiques* et celle des *Ideen*, coïncident respectivement avec la théorie représentationnelle de l'esprit, et la théorie computationnelle de l'esprit.

La théorie représentationnelle est la théorie selon laquelle (*i*) avoir des états mentaux, c'est avoir des représentations mentales, et (*ii*) selon laquelle il existe différentes manières d'en avoir, *i.e.* les différentes attitudes que l'on peut avoir à leur propos, le doute, la croyance, la certitude, etc. Par conséquent, les états mentaux se caractérisent comme des *attitudes propositionnelles* qui correspondent à l'*attitude* qu'un organisme adopte à propos d'une *proposition*, représentation ou contenu représentationnel. Cette notion d'at-

2. Husserl a récemment été découvert depuis la philosophie analytique par des auteurs comme Follesdal qui voit en Husserl le Frege des états mentaux : Husserl a appliqué aux états mentaux l'analyse frégéenne de la proposition en sens et référence. Pour mesurer toute la pertinence d'une telle lecture, on consultera (Fisette, 1994) et (Dreyfus & Hall, 1982). Par ailleurs, Dreyfus a fait de Husserl le phénoménologue précurseur du cognitivisme computationnel : on consultera son (Dreyfus, 1984) et surtout ses articles (Dreyfus, 1991a) et (Dreyfus, 1991b) du numéro spécial des *Etudes philosophiques* consacré à la psychologie cognitive. (Janicaud, 1995) confronte ces lectures majoritairement outre-atlantique de la phénoménologie avec des lectures continentales.

titude propositionnelle est nécessaire dans la mesure où elle permet de distinguer les deux caractéristiques fondamentales des états mentaux que l'on veut dégager dès lors qu'on entreprend une étude de l'esprit, qu'elle soit psychologique ou phénoménologique.

> « D'une part, les états mentaux sont distingués par le *contenu* des représentations associées, et nous pouvons par conséquent comprendre la différence entre penser que Marvin est mélancolique et penser que Sam l'est [...] ; et, d'autre part, les états mentaux sont distingués par la *relation* que le sujet entretient avec la représentation associées (et ainsi nous pouvons comprendre la différence entre penser, espérer, supposer, douter et prétendre que Marvin est mélancolique). » (Fodor, 1981)

Il appert que l'opposition *contenu/ relation* se confond avec la distinction husserlienne entre la matière et la qualité intentionnelle :

> « Pour parler plus clairement encore, nous pouvons dire que la matière est cette propriété résidant dans le contenu phénoménologique de l'acte, qui ne détermine pas seulement que l'acte appréhende l'objectité, mais aussi à *quel titre* il l'appréhende, quels caractères, quels rapports, quelles formes catégoriales il lui attribue de par lui-même. Il dépend de la matière de l'acte que l'objet soit pour l'acte tel objet et non tel autre, elle est dans une certaine mesure *le sens de l'appréhension objective* (ou plus brièvement le *sens d'appréhension*) ; » p. [415-416] RL. V

On peut dire que dans une certaine mesure la qualité intentionnelle est la relation de l'esprit à la matière, et que la matière est « *dans l'acte, ce qui lui confère éminemment la relation à une objectité* », c'est-à-dire la relation à l'objet. La qualité indique le comment de la visée par l'esprit de la représentation, et la matière le comment de la visée de l'objectité par la représentation.

Ces notions d'attitudes propositionnelles, ou de qualités et de matières intentionnelles qui leur sont équivalentes, constituent le minimum nécessaire pour aborder une étude de l'esprit en tant qu'esprit, c'est-à-dire pourvu d'états mentaux. Cela signifie qu'il « est difficile de croire qu'une psychologie sérieuse puisse travailler avec moins (ou moins détaillées) de distinctions que celles-ci, et il est difficile de croire qu'une psychologie qui effectue ces distinctions puisse éviter de prendre la notion de représentation mentale au sérieux. » (Fodor, 1981). C'est donc le moins que l'on puisse faire, et donc reconnaître que Husserl adopte une théorie représentationnelle de l'esprit dans les *Recherches Logiques* ne veut pas dire tellement plus que le fait qu'il veut étudier l'esprit en tant que tel. C'est pourquoi Dreyfus parle d'une ontologie minimaliste à propos des *Recherches*, et notamment il n'est fait mention d'aucune notion permettant de concrétiser matériellement les états mentaux. C'est donc une étude descriptive de l'intentionnalité.

Cela voudrait dire que les notions d'esprit et de représentation s'interdéfinissent. Cela paraît effectivement juste dans la mesure où la seule manière de définir l'esprit en tant qu'esprit, c'est-à-dire sans faire appel à d'autre notion ou concept non mentaux ou spirituels (relevant de l'esprit), c'est de le définir en termes de représentations, en termes d'intentionnalité. Cela signifie que la seule manière de définir l'esprit, c'est de revenir à la caractérisation de Brentano, c'est-à-dire que répéter avec Husserl que la caractéristique essentielle de la conscience (de l'esprit) est d'être à propos d'autre chose, de viser autre chose qu'elle.

A rebours, toutes les théories cognitives révoquant la notion de représentation, révoquent également la possibilité de caractériser l'esprit en tant qu'esprit : ce dernier est alors toujours défini à partir d'autres notions, qu'elles soient empruntées à la biologie (l'esprit, c'est une propriété du cerveau), aux théories de l'évolution (l'esprit, c'est la faculté de survivre), etc.

Par conséquent, s'intéresser à l'esprit en tant qu'esprit, c'est considérer une théorie intentionnelle, c'est-à-dire s'intéresser aux représentations. Dans la perspective d'une psychologie de l'esprit, les divergences ne peuvent donc intervenir qu'au delà de la théorie représentationnelle de l'esprit, c'est-à-dire dans les aménagements apportés à la théorie représentationnelle. Or, et c'est cela qui nous intéresse au premier chef, Dreyfus affirme que Husserl a spécialisé dans les *Ideen* la théorie représentationnelle de l'esprit en une théorie *computationnelle* de l'esprit. Comme l'explique Fodor (1981), alors que la théorie représentationnelle de l'esprit considère que « les attitudes propositionnelles sont des relations que des organismes entretiennent avec des représentations mentales », la théorie computationnelle considère que « les processus mentaux n'ont accès qu'aux propriétés formelles (non-sémantiques) des représentations mentales sur lesquelles ils sont définis. ».

Or, selon Dreyfus, l'élaboration de la notion de « noème » dans les *Ideen* marque le passage de la conception descriptive de l'intentionnalité des *Recherches* à une conception formelle dans la mesure où d'une part le noème constitue le ressort de l'intentionnalité puisque c'est par la médiation du noème que l'esprit se rapporte à un objet visé, intentionné, et où d'autre part le noème est une structure formelle non sémantique : l'esprit parvient à viser un objet en manipulant formellement les structures noématiques. En d'autres termes, la théorie noématique rendrait compte de l'intentionnalité en termes non sémantiques, ce qui reviendrait à dire qu'il s'agit d'une réduction formelle de l'intentionnalité à une structure computationnelle.

2.1 Le noème comme structure de calcul

Le noème est une notion introduite dans les *Ideen* pour décrire de manière objective et non pas subjective l'objet en tant qu'il est pensé, l'objet d'après la manière dont on le pense. Le noème est par conséquent un objet intentionnel, c'est-à-dire que sa consistance objective dépend exclusivement de la manière de le viser intentionnellement à travers les noèses, actes intentionnels subjectifs de la pensée.

Le noème est une véritable innovation dans la phénoménologie husserlienne dans la mesure où il constitue un composant non réel du vécu, une partie idéale, c'est-à-dire non hylétique, du vécu. On sait en effet, que Husserl dénommait par « réel » (*reelle*) la caractéristique immanente du vécu, c'est-à-dire qui se trouve « réellement » présente dans le vécu et correspond à un fait effectif, voire empirique, dans la conscience :

> « Il importe maintenant de décrire ce qui subsiste de cette analyse, à titre de résidu phénoménologique, si nous retournons à la « pure immanence », et *ce qui dans ce cas peut compter comme composante réelle (*reelle*) du pur vécu* et ce qui ne le peut pas. Il faut alors apercevoir avec un clarté totale que le vécu de perception pris en lui-même comporte bien dans son essence « l'arbre perçu comme tel », ou le noème complet qui reste intact quand on met hors circuit la réalité de l'arbre lui-même et celle de l'ensemble du monde ; mais d'autre part ce *noème* avec son « arbre » entre guillemets *n'est pas réellement (*reell *contenu dans la perception, pas plus que ne l'était l'arbre de la réalité.* » Ideen I (Husserl, 1950, p 336)

Le noème n'est pas l'objet perçu, mais l'objet en tant qu'il est perçu. C'est ainsi que l'objet perçu peut disparaître sans que le noème soit en rien modifié. Structure de l'objectivité, le noème n'est pas un objet du monde, mais une composante non réelle de la visée intentionnelle, à savoir son pôle objectif. Le noème a un sens par lequel il se rapporte à son objet :

> « Par contenu nous entendons le « sens », dont nous disons que, en lui ou par lui, la conscience se rapporte à un objet en tant qu'il est le « sien ». [...] Tout noème a un « *contenu* », à savoir son « sens » ; par lui le noème se rapporte à « son » *objet*. » Ideen I (Husserl, 1950, p. 436)

Cela fait irrésistiblement penser au *Sinn* frégéen qui correspond au « mode de donation de la référence ». Le noème serait une structure de la conscience sur au moyen de laquelle elle se rapporte à un objet : le noème est le mode de donation de l'objet (Follesdal, 1982; Dummett, 1991). Mais, contrairement à la référence frégéenne, l'objet donné par l'intermédiaire du noème n'est pas l'objet perçu, l'objet du monde, mais l'objet intentionnel, c'est-à-dire l'objet corrélat de la visée et n'existant en tant qu'objet qu'à ce titre. En effet, dans la pensée « à vide » d'un objet, c'est-à-dire lorsque l'on pense à un objet sans qu'il soit intuitivement donné à la conscience, il y a bien un objet donné à la conscience par l'intermédiaire de la conscience : cet objet ne peut donc être la référence frégéenne puisque cette dernière est en effet obligatoirement un objet réel et effectif du monde, contrairement à l'objet intentionnel donné dans le noème.

Comment fonctionne le noème husserlien ? Tout le génie de la phénoménologie husserlienne consiste à distinguer l'objet perçu qui apparaît des données par lesquelles il apparaît : il faut distinguer l'objet apparaissant de son apparition. Ainsi, dans la vision, il faut distinguer la sensation de rouge, qui est une propriété du sujet percevant : il a la sensation de rouge, de la perception du rouge qui caractérise non le sujet de la perception mais son objet. Le rouge senti n'est pas le rouge perçu, mais le moyen par lequel le rouge perçu est donné à la conscience. Comment se passe cette donation ? La conscience contient donc une sensation de rouge, le rouge senti, qui en soi ne renvoie à rien d'autre que lui-même. La conscience, par un acte donateur de sens, lui confère un sens, l'anime d'une visée intentionnelle à travers laquelle elle va reconnaître un rouge perçu. Mais ce rouge perçu n'est pas le rouge senti : il ne peut être dans la conscience comme une donnée, il caractérise en effet un objet perçu qui ne peut être dans la conscience. Le rouge perçu appartient donc à la conscience en tant qu'elle le vise et qu'elle dépasse, en l'animant intentionnellement, le rouge senti pour l'atteindre, et n'y appartient pas dans la mesure où il n'est pas une donnée dans la conscience mais un objet hors de la conscience mais visé par elle. C'est en cela que le rouge perçu est une composante (il appartient à la conscience) non réelle (mais il n'y appartient pas comme le rouge senti).

A présent, comment peut se dérouler l'animation intentionnelle ? Prenons l'exemple d'une donnée sensible correspondant à la face visible d'un maison : elle est reconnue comme étant la vision d'une maison et, conformément à cette vision, on anticipe qu'il y ait des données sensibles correspondant aux faces cachées si l'on tourne autour de cette maison. En animant les données sensibles d'un sens objectif, on anticipe le cours ultérieur de l'expérience sensible. Si un meuble mord sur un tapis rouge, je m'attends à avoir une sensation de rouge en tirant sur le tapis pour dévoiler la partie cachée par le meuble. Il y a donc une structure qui synthétise l'expérience en donnant un statut objectif aux données sensibles et en anticipant, sur la foi de l'interprétation objective de ces données, les futures données sensibles. Cette structure, c'est le noème. Par ailleurs, Husserl appelle *noèse* l'activité subjective de la conscience correspondant à l'animation intentionnelle des données sensibles ou hylétiques. Le noème décrit comment intégrer la multiplicité des données hylétiques pour en faire la visée d'un objet unique.

> « Faisons réflexion sur la sensation, sur les esquisses : ce sont bien des données évidentes que nous saisissons ; et si nous varions l'attitude et la direction de l'attention, nous pouvons, avec une parfaite évidence, mettre également en relation ces esquisses avec les moments objectifs correspondants et les reconnaître comme correspondants : nous voyons alors sans difficulté que les couleurs esquissées qui se rattachent à quelque couleur immuable attribuée à la chose sont dans le même rapport que « l'unité » à une « multiplicité » continue. » *Ideen I* (Husserl, 1950, p. 337)

En d'autres termes le noème est une *règle* permettant de synthétiser et organiser l'expérience. Structure d'anticipation, elle prescrit le cours de l'expérience en animant les données sensibles pour en faire la visée intentionnelle d'un objet. Dans une telle perspective, l'objet est un principe d'unité, autrement dit de synthèse, des données sensibles. Ce qu'elles ont à faire ensemble, leur commensurabilité, est prescrit par un principe d'unité objective : ensemble, elles constituent un objet grâce à l'animation intentionnelle noétique dont le sens est donné par le noème associé. Le noème joue donc le rôle que tient le concept dans la philosophie kantienne : de principe synthétique.

Et tout comme le concept chez Kant, le noème possède un statut logique dans la mesure où il est l'assemblage structuré de prédicats qui déterminent le sens de l'expérience. Le sens noématique est une expression prédicative spécifiant la nature de l'objet visé :

> « Il devient par là-même évident qu'un *statut tout à fait invariable* est délimité en chaque noème. Toute conscience a son « *Quid* » et tout ce qui est visé a « son » aspect objectif ; il est évident que dans le principe on doit pouvoir faire cette description noématique du « Quid », tel exactement qu'il est visé ; l'explicitation et la saisie conceptuelle nous permettent de former un système clos de « *prédicats* » formels ou matériels, effectivement déterminés ou même laissés « indéterminés » (visés « à vide ») ; ces prédicats, considérés dans leur *signification* modifiée, déterminent le « *contenu* » de ce noyau objectif du noème qui est ici en cause. » *Ideen I* (Husserl, 1950, p. 441)

Toute la difficulté est alors de savoir ce qu'il faut comprendre quand on parle de structure prédicative pour le noème. S'agit-il d'affirmer, comme le fait Dreyfus, qu'il s'agit de prédicats strictement *formels*, dont la manipulation est *syntaxique* ?

> « Il est cependant tout à fait évident, comme nous allons le voir, que Husserl considérait les noèmes comme des structures *formelles* complexes, à partir du moment où, bien qu'il ne bénéficie pas de l'aide de l'ordinateur digital pour fournir un modèle pour ses intuitions, il envisageait le noème comme une « règle stricte pour des synthèses possibles ». Et, pour retourner la question, rien n'indique qu'il ait jamais pensé comme *sémantiques* les règles auxquelles il s'intéressait. Elles n'ont pas de conditions *ceteris paribus*, et leur application ne suscite aucun problème d'interprétation. Nous verrons qu'elles fonctionnent plutôt comme des opérations empilées les unes sur les autres et ayant pour fondement ultime des opérations primitives qui organisent les données des sens dans les éléments sur lesquels opèrent les règles d'ordre supérieur. » (Dreyfus, 1991a, p.11)

Si l'on en croit Dreyfus, il faudrait comprendre la phénoménologie transcendantale à l'instar du fonctionnalisme computationnel[3]. On aurait ainsi des données effectives, les matériaux hylétiques ou sensibles délivrés par la sensation. Ces données sont réelles et immanentes : il s'agit du rouge senti, de la sensation de rouge, qui caractérise seulement le sujet percevant, sa propre sensation, et non l'objet qu'il perçoit qui serait, dans une conception naïve, la cause de la présence de ces données sensibles dans la conscience. Ces données hylétiques correspondraient alors aux états physiques permettant, dans le schéma fonctionnaliste, de réaliser les états computationnels.

La noèse correspond au fait que ces données hylétiques se trouvent enrôlées dans la constitution d'un objet de la perception : animés intentionnellement par des actes de la conscience qui leur donne un sens, ces matériaux hylétiques participent à la constitution d'un objet. Mais constituer un objet veut dire ici disposer d'une règle réglant le déroulement des matériaux sensibles et qui en prescrit l'agencement. Constituer un objet, c'est comme nous venons de le rappeler, prescrire ce que les matériaux ont à faire ensemble, notamment en anticipant les matériaux à venir. Or, cette règle correspond à une structure prédicative se comportant comme un programme. Cela signifie que les matériaux hylétiques s'animent intentionnellement tout simplement en *exécutant le programme correspondant à un sens objectif déterminé par un noème.*

Le noème est donc une structure de calcul, un programme, dont l'exécution correspond à la reconnaissance d'un objet et à l'anticipation de l'expérience future sur le fondement de cette reconnaissance.

Une telle interprétation est-elle légitime ? Peut-on affirmer que le noème est un programme ? Comme le remarque justement Dreyfus, « il n'est pas facile de déterminer si Husserl soutenait, en fait, ce que Fodor appelle la théorie computationnelle de l'esprit, – c'est-à-dire si, selon Husserl et à la différence de ce que pensait Frege, le rôle des sens-prédicats est de représenter des objets selon leurs aspects et d'unifier diverses expériences des mêmes objets, sur la base stricte de leur forme (donc, comme un système syntaxique indépendant de toute interprétation) ». Toute la question est en effet de savoir si les prédicats se définissent par leur forme syntaxique ou par leur contenu sémantique.

Il faut tout d'abord s'entendre sur la notion de « sémantique ». Dreyfus, comme toute la tradition anglo-saxonne, adopte une conception référentielle de la sémantique : un prédicat a un sens parce qu'il renvoie à un objet du monde, parce qu'il dispose d'une référence (cf. chapitre 3). C'est l'acception frégéenne dont nous avons dit plus haut qu'elle ne convenait pas pour comprendre la nature de l'objet visé par la conscience chez Husserl. Autrement dit, la conception husserlienne de la conscience est non sémantique au sens frégéen puisque le sens noématique se définit indépendamment de l'existence ou non d'un objet dans le monde. Or, si elle n'est pas sémantique, c'est qu'elle est syntaxique.

Dans la mesure en effet où la référence n'est pas déterminante pour régler le cours de l'expérience, il est tentant de penser que seule est prise en compte la structure prédicative du noème indépendamment de sa sémantique pour synthétiser les matériaux hylétiques.

3. Le fonctionnalisme est une théorie fondamentale de la psychologique cognitive, qui lui fournit son cadre méthodologique et ses principes théoriques. En assimilant l'esprit à un programme qui s'exécuterait sur le cerveau, le fonctionnalisme donne pour tâche à la psychologie de déterminer quel programme est responsable de tel comportement psychologiquement décrit. (Richard, 1990) précise tous ces points. Sur le fonctionnalisme comme théorie philosophique de l'esprit, voir (Fodor, 1975; Fodor, 1983; Fodor, 1987; Putnam, 1989)

En d'autres termes, Husserl affirmerait un *solipsisme* de la conscience où il serait possible d'en établir la nature et les structures indépendamment de l'existence d'un monde ou non. Par conséquent, le computationnalisme de Husserl découlerait de son solipsisme. C'est donc cette question qu'il nous faut aborder pour confirmer ou infirmer la thèse qui voit en Husserl un précurseur illustre de Fodor.

Mais il faudra néanmoins distinguer la question du computationnalisme ou fonctionnalisme syntaxique de Husserl de la question du fonctionnalisme de la conscience. En effet, le fonctionnalisme *syntaxique* se caractérise par son indépendance à la sémantique : c'est là son formalisme syntaxique. Mais il affirme également l'indépendance au substrat physique : la matière physique permettant d'exécuter un programme est nécessaire pour que le programme existe mais ne détermine en rien l'exécution du programme ; les lois de l'informatique ne sont pas réductibles aux lois de la physique. C'est là son caractère proprement *fonctionnaliste*. Si la question du solipsisme permettra d'élucider la question du formalisme syntaxique chez Husserl, elle n'élucidera pas la question de son fonctionnalisme. En termes phénoménologiques, la question est : les matériaux hylétiques déterminent-ils la nature de l'animation intentionnelle qui les enrôle dans la constitution d'un sens objectif? Y-a-il indépendance du sens intentionnel à son substrat hylétique? Or, une telle question, Husserl ne se l'est pas explictement posée. C'est cette question que nous reprendrons pour conclure au caractère non fonctionnaliste de la phénoménologie dans la section 6-3.3.

2.2 Solipsisme méthodologique, doute cartésien et réduction phénoménologique

Un argument souvent invoqué par Dreyfus pour voir dans la tradition rationaliste de Descartes à Husserl les prémisses de l'IA est la similitude, avec le solipsisme méthodologique propre à la théorie computationnelle de l'esprit, des démarches consistant en première approximation à abolir le monde pour le reconstruire par la seule force de l'esprit et de la pensée. C'est ainsi que sont invariablement convoqués le doute cartésien et la réduction phénoménologique husserlienne comme ancêtres philosophiques de la démarche cognitiviste. Cependant, si le fonctionnalisme de Husserl nous paraît être une hypothèse défendable dans une certaine mesure, il nous semble en revanche que ces rapprochements sont fortement discutables.

Fodor (1981) définit la condition de formalité selon laquelle l'esprit ignore la signification des symboles qu'il manipule, comme c'est le cas pour tout ordinateur. En particulier, l'esprit ignore tout de l'éventuelle vérité ou fausseté des propositions, de la référence des représentations sur lesquelles il calcule. Par conséquent, si les calculs sont menés indépendamment de la référence des symboles, c'est qu'ils sont indépendants de la réalité du monde : il peut alors se référer aux *Méditations métaphysiques* pour introduire son solipsisme méthodologique.

Le parallèle entre Fodor et Descartes[4] est discutable dans la mesure où il ignore des aspects fondamentaux de ce qu'est une représentation pour Descartes, en particulier des caractéristiques correspondant à ce qui fonde l'intentionnalité d'une représentation, c'est-

4. Descartes (dont les œuvres ont reçu une édition critique remarquable dans (Descartes, 1963; Descartes, 1967; Descartes, 1973)) est un auteur fondamental pour les sciences cognitives. Nous ne pouvons expliciter ici davantage la philosophie cartésienne : on se reportera à (Rodis-Lewis, 1971; Alquié, 1950; Guenancia, 1986).

à-dire le fait qu'elle représente quelque chose, qu'elle y renvoie. Ce problème se retrouve également dans les lectures analytiques de Husserl. En substance, on peut dire que les difficultés liées à ce parallèle reposent sur un réalisme naturaliste naïf qui se donne pour point de départ précisément ce que les démarches cartésiennes et phénoménologiques veulent retrouver au terme de leur ascèse philosophique.

Qu'est-ce qui fait d'une représentation une représentation? Quelle est son essence? Une représentation est par essence représentation de quelque chose, elle en est la visée. Cela signifie que la représentation est essentiellement intentionnelle, intrinsèquement intentionnelle. Réciproquement, tout ce qui est intentionnel est visée de quelque chose, et donc en est la représentation [5] . Une représentation se définit non par le fait d'être quelque chose, mais par le fait de renvoyer à quelque chose. Une idée de chapeau ne se définit que par le fait de renvoyer à l'objet « chapeau », et non par le fait d'être quelque chose. On pourrait dire qu'une idée de chapeau, c'est d'abord et avant tout une idée, c'est-à-dire un quelque chose, qui renvoie à un objet ou une classe d'objet, les chapeaux. Mais qu'est-ce qu'une idée? Est-ce vraiment un quelque chose? Ne dit-on pas plutôt qu'une idée, c'est avant tout une idée de, une idée sur, une idée à propos de, etc., c'est-à-dire un pur renvoi? Par conséquent, dire qu'une idée de chapeau, c'est une idée, c'est simplement dire que c'est un pur renvoi à quelque chose, mais non que c'est quelque chose.

Ces distinctions ne sont certes pas nouvelles puisque l'on peut retrouver celles que Descartes mobilise dans les *Méditations métaphysiques* pour démontrer l'existence de Dieu. Il distingue dans les idées trois principes : la *réalité subjective*, la *réalité objective*, la *valeur objective*. La réalité subjective d'une idée est sa *réalité formelle*, c'est-à-dire sa réalité effective, c'est-à-dire encore le fait pour elle d'être quelque chose. Or Descartes affirme que l'idée en tant que telle n'a pas de réalité formelle propre, mais seulement celle du *Cogito* auquel elle appartient : être quelque chose pour une idée, c'est être pensée par un *Cogito*. Son existence effective ne peut être considérée indépendamment du *Cogito*, car sans lui elle n'est rien.

La réalité objective d'une idée est son contenu représentationnel, le fait de renvoyer à ceci ou à cela. Une idée représente quelque chose, et c'est la nature de ce qui est représenté qui *type* ou individualise le type de l'idée. Ainsi, ce qui différencie une idée d'une autre idée, ce n'est pas sa nature, puisqu'elle n'est rien par elle-même : appartenant au *Cogito*, sa nature renvoie à celle du *Cogito* si bien que toutes les idées sont indiscernables à ce point de vue. Par conséquent, c'est par ce qu'elles représentent qu'elles peuvent se distinguer. Mais, cela ne veut pas dire qu'elles se distinguent d'après la réalité de ce qu'elle représente : l'idée de *Pégase* est bien différente de celle d'*Ulysse* bien que ni l'un ni l'autre n'existe. C'est là le troisième principe qu'il faut associer aux idées : leur valeur objective. La valeur objective consiste dans la réalité formelle, *i.e.* l'existence effective, de ce qui est représenté ou visé dans l'idée, c'est la réalité formelle de la réalité objective. Si l'on remarque que mon idée de chaise est la même qu'il existe ou non des chaises, cela signifie donc que la valeur objective ne participe en rien dans la définition et la différenciation des idées et que seule la réalité objective compte.

5. En toute rigueur, il n'est pas exact de formuler les choses ainsi dans la mesure où la représentation n'est qu'une manière parmi d'autres de viser intentionnellement un objet : alors que son intentionnalité consiste à le représenter, d'autres structures intentionnelles peuvent viser quelque chose en le jugeant, en l'aimant, en le haïssant, etc., modalités intentionnelles ne reposant pas nécessairement sur des représentations. Husserl s'est posé explicitement la question dans la cinquième *Recherche logique* de savoir si toute relation intentionnelle à un visé devait reposer sur une représentation et y répond affirmativement.

Pour comprendre le solipsisme cartésien, il faut, outre ces distinctions, considérer la théorie cartésienne de la vérité. En effet, pour Descartes, la vérité ne découle pas d'un accord constaté entre ce qui existe et ce qui est représenté ou visé car nous ne possédons pas de moyen sûr et fiable de déterminer ce qui existe. Notre seul moyen de caractériser avec certitude l'existence est d'en passer par la clarté et la distinction de nos idées, comprises ici comme contenus représentationnels. Par conséquent, la vérité pour Descartes n'est pas une vérité correspondance puisqu'il est impossible d'accéder aux deux termes de la correspondance indépendamment de la correspondance elle-même. La vérité cartésienne sera donc une vérité évidence, où l'évidence de la représentation en tant qu'elle est une représentation, c'est-à-dire qu'elle possède une contenu représentationnel clair et distinct, est la garantie de l'existence effective ou réalité formelle de ce qui est représenté.

Maintenant, si l'on dit que le doute cartésien revient à récuser l'existence du monde pour se concentrer uniquement sur le *Cogito* et ses *cogitata*, cela ne compromet en rien le *sens* ni la *vérité* des *cogitata*, puisque ces notions sont définies sur les *cogitata* indépendamment de l'existence du monde.

Et c'est là qu'intervient la principale différence avec Fodor. En effet, ce dernier adopte une théorie de la vérité comme correspondance entre une représentation et la réalité mondaine de ce qu'elle représente. Par conséquent, s'il écarte la réalité du monde, il écarte du même coup toute possibilité de déterminer tant le sens que la vérité d'une représentation. Mais cela a des conséquences très importantes : en écartant le contenu représentationnel comme renvoi, il ne reste plus de moyen de typer les représentations en tant que représentations, c'est-à-dire comme renvoi. Mais, si on caractérise les représentations par leur réalité subjective, elles sont indiscernables : ce sont toutes des *cogitata* du *Cogito*, ou des attitudes propositionnelles de l'organisme. Il faut donc s'y prendre autrement. La solution est alors que considérer que les idées mentales sont des objets en eux-mêmes, et pas seulement un renvoi ou une visée. C'est ainsi que Fodor en fait des expressions syntaxiques de symboles et que les idées se voient conférer une réalité formelle distincte de celle de l'organisme (*Cogito*). Les idées deviennent des objets à part entière, et la théorie computationnelle de l'esprit en fait même des objets physiques matériels puisqu'il s'agit alors de représentations syntaxiques manipulées par un programme.

La première différence entre les solipsismes fodorien et cartésien est que le premier réifie les représentations alors que le second n'en a cure. Ce dernier se contente de considérer le *Cogito* comme une chose qui pense, ce qui posa bien des problèmes pour sa postérité, mais il n'eut point besoin de vouloir faire des idées des objets à part entière. Par ailleurs, appartenant au *Cogito*, elles ne peuvent être matérielles : ce n'est que le naturalisme de Fodor qui l'oblige, puisqu'il n'existe de substance que de substance physique, à considérer les idées comme des types formels d'objets physiques. Cela signifie que le solipsisme fodorien n'est pas un doute radical hyperbolique, car il ne met pas en doute un seul instant l'existence des idées en tant qu'objets. Or, Descartes ne s'accorde jamais une telle facilité puisque, jusqu'à la découverte de l'existence de Dieu, le *Cogito* est la seule existence avérée.

Une seconde différence vient du fait que le solipsisme fodorien ne peut être surmonté depuis l'intériorité qu'il a constituée alors que le solipsisme cartésien possède cette caractéristique unique de se dépasser de l'intérieur pour poser l'existence de Dieu et donc retrouver un monde qu'il avait aboli faute d'avoir la certitude de son existence. Cela signifie que l'abolition du monde, son anéantissement, n'annule pas le sens des représentations si bien que par la seule force du contenu représentationnel, Descartes pose qu'il est nécessaire d'avoir une existence effective autre que celle du *Cogito*. L'idée du parfait possède un tel contenu représentationnel qu'elle est assurée de sa valeur objective. En retour, cette valeur objective vient soutenir la réalité formelle du *Cogito*, en la rendant certaine et permanente, alors que jusque là, son existence ne pouvait être certaine que lorsque j'y pensais.

Le solipsisme cartésien met donc en évidence l'intentionnalité, alors que le solipsisme fodorien l'annule. Descartes rend justice au fait que toute idée mentale est par essence intentionnelle, et supprimer son intention, c'est la supprimer elle-même. Fodor pense pouvoir caractériser les idées indépendamment de leur intentionnalité car il leur confère une réalité formelle excédant leur définition d'idée en tant qu'idée : il introduit un point de vue naturaliste psychologiste selon lequel ce qui est dans la conscience doit avoir une consistance matérielle et une localisation spatio-temporelle. C'est s'interdire dès lors toute possibilité d'étudier l'esprit en tant qu'esprit, puisque l'on veut étudier les conditions sous lesquelles un objet matériel peut avoir des représentations.

Par conséquent, il nous paraît clair que le parallèle qu'établit Fodor entre son solipsisme et celui de Descartes ne peut venir que d'une équivoque verbale et qu'il s'agit là de deux entreprises philosophiques non seulement distinctes, mais opposées dans le principe.

Il en est de même pour le parallèle Fodor/Husserl établit par Dreyfus, pour des raisons assez semblables, puisque cela revient à nier toute caractérisation possible de l'esprit en tant qu'esprit. Ce qui tient lieu de solipsisme méthodologique chez Husserl est la fameuse réduction phénoménologique exposée principalement dans les *Ideen* (Husserl, 1950). Cette réduction consiste non à abolir ou nier le monde dans un doute hyperbolique, mais à suspendre toute croyance en l'existence du monde, toute thèse d'existence.

La réduction consiste donc à réduire le monde à son phénomène, et l'existence à son sens. Dès lors Husserl peut alors distinguer les différents sens de l'être, et distinguer, pour reprendre la distinction la plus célèbre, le monde de la conscience du monde sensible selon le mode d'existence particulier des objets appartenant à ces mondes. Comme dans le doute cartésien, suspendre la croyance au monde ne supprime pas le sens du monde, et c'est ce sens qui permet de retrouver le monde, en fondant la transcendance du monde sur l'immanence du vécu. Ce qui rend cette manœuvre possible, c'est que l'immanence du vécu est une expression incorrecte : le vécu comporte des parties non réelles qui transcende l'immanence des parties réelles. Mais, cette transcendance est celle de l'essence du vécu et de son idéalité, et non celle de l'objet mondain existant [6].

Si bien que les raisons pour lesquelles le solipsisme fodorien ne peut être assimilé à la réduction phénoménologique sont essentiellement les mêmes que celles le distinguant du doute cartésien. Le solipsisme fodorien est une négation naturaliste du monde, car le sujet effectuant la négation est toujours dans le monde. L'originalité des démarches cartésienne et husserlienne est que le sujet effectuant la réduction ou pratiquant le doute n'est pas posé comme un objet du monde, mais comme l'instance à partir de laquelle le monde doit être reconstruit ou constitué.

6. Nous revenons sur ces points fondamentaux dans les sections suivantes.

3 Les ambiguïtés husserliennes

Mais, même si les raisons habituellement alléguées pour reconnaître en Descartes et Husserl de lointains précurseurs du cognitivisme computationnel que peut incarner Fodor s'avèrent non fondées, il n'en demeure pas moins que la phénoménologie husserlienne recèle certaines ambiguïtés qui rendent difficile sa compréhension et délicate son interprétation. La question peut être résumée ainsi : quelles peuvent être les raisons d'imputer à Husserl une théorie fonctionnaliste du sens ? Il faut distinguer dans le traitement de cette question le fonctionnalisme syntaxique d'une part et le fonctionnalisme sémantique, plus exactement le fonctionnalisme intentionnel, d'autre part. Accorder un fonctionnalisme syntaxique chez Husserl, c'est, comme on l'a vu plus haut, reconnaître dans la phénoménologie transcendantale un cognitivisme computationnel avant la lettre. Accorder un fonctionnalisme sémantique, c'est reconnaître une indépendance du sens à la matière sensible qui sert de support à l'animation intentionnelle, sans bien sûr réduire le sens à son véhicule syntaxique – sinon il s'agirait du fonctionnalisme syntaxique.

Ces deux questions doivent être examinées dans le cadre de la question du sens : c'est en effet cette question du sens, compris comme étant le problème de l'idéalité, qui commande la phénoménologie husserlienne de *La philosophie de l'arithmétique* à *L'origine de la géométrie*. Car Husserl, comme tout penseur profond, a compris que la question du sens ne s'élimine pas ni ne se réduit, mais qu'elle doit être comprise à la mesure de son enjeu : celui de la possibilité d'un idéal qui transcende les véhicules empiriques qui lui prêtent leur matérialité pour le réaliser dans l'effectivité. C'est donc cette théorie du sens et de l'idéalité chez Husserl qu'il convient d'examiner dans le cadre de notre réflexion pour replacer dans le cadre qui leur convient la question du fonctionnalisme husserlien.

3.1 Vécu, signification, et représentation mentale

Une théorie du sens et de la signification est possible dès lors que le sens est distingué de la représentation mentale, dès lors que le problème du sens possède une autonomie théorique vis-à-vis de la psychologie.

Les théories du sens qui se sont constituées à l'aube du XX^{e} siècle se sont posées comme des adversaires résolues du psychologisme. Cette autonomie du sens par rapport à la psychologie fut d'abord pensée comme une distinction entre la logique et la psychologie, une théorie du sens ne pouvant qu'être logique, laissant à la psychologie le champ du vécu et de la représentation mentale.

La phénoménologie husserlienne est une théorie de la connaissance qui veut étudier la connaissance comme un phénomène *sui generis* et non comme un phénomène appartenant à un ordre de réalité sous-jacent, que ce soit un substrat psychologique (la connaissance étant rapportée aux faits spatio-temporels de la conscience, le fait de penser ceci ou cela à tel instant du temps objectif et à tel endroit de l'espace objectif), un substrat biologique (la connaissance étant rapportée aux phénomènes biologiques du cerveau), un substrat mondain (la connaissance étant rapportée aux phénomènes du monde objectif externe). L'ennemi, pour Husserl, prenait la figure du psychologisme, où la connaissance est un événement psychologique de la conscience. Husserl s'intéresse en effet à la définition

> « d'une théorie objective de la connaissance, et s'y rattachant intimement, d'une phénomé-

nologie pure des vécus de la pensée et de la connaissance. Celle-ci, comme la phénoménologie pure des vécus en général qui l'englobe, a exclusivement affaire, dans une généralité d'essence pure, aux vécus qu'on peut appréhender dans l'intuition et analyser [phénoménologie], mais non aux vécus aperçus empiriquement comme des faits réels, en tant que vécus d'hommes ou d'animaux les vivant dans le monde des phénomènes, [phénoménalité] posé comme fait d'expérience. » in Recherche Logique I p. [2]

Dans son projet, la phénoménologie que Husserl vise

> « ne veut pas expliquer, au sens psychologique ou psychophysique, la connaissance, l'événement de fait dans la nature objective, mais élucider l'idée de la connaissance d'après ses éléments constitutifs ou encore d'après ses lois ; ce ne sont pas les relations réelles de coexistence et de succession dans lesquelles sont insérés les actes de connaissance accomplis en fait qu'elle cherche à atteindre, mais ce qu'elle veut comprendre, c'est le sens idéal des relations spécifiques dans lesquelles l'objectivité de la connaissance trouve sa légitimité. » in Recherche Logique I p.[21]

Ce naturalisme psychologique critiqué par Husserl n'est pas le naturalisme moderne, car il ne saurait revendiquer les critères de scientificité exigés de la science. Notamment, en réduisant la connaissance à un événement psychologique, on recourt à l'introspection, à des sentiments inobjectivables et privés de la conscience, qui ne possèdent pas les propriétés nécessaires à un phénomène pour appartenir à une phénoménalité scientifique : être publiquement observable, intersubjectivement répétable, objectivement séparable. Le naturalisme psychologique est un naturalisme non scientifique : il est critiqué d'une part par Husserl en tant que naturalisme, et critiqué d'autre part par les sciences modernes en tant que non scientifique, c'est-à-dire que les sciences modernes qui sont par définition même des sciences de la nature et qui constituent par conséquent le naturalisme scientifique, critiquent le naturalisme psychologique contemporain de Husserl en tant que mauvais naturalisme. Il ne faut pas déduire du fait que la cible soit commune, que les intérêts soient communs. Ce que critique Husserl dans le psychologisme, ce n'est pas tant son manque de rigueur, mais son naturalisme.

Mais, si la seule possibilité de penser l'autonomie du sens vis-à-vis de la conscience psychologique, est d'opposer la signification logique au vécu de la conscience, ne risque-t-on pas de penser le sens à partir d'un préjugé calculatoire, d'une conception computationnelle du sens ? Le risque n'est-il pas alors de définir un fonctionnalisme du sens, où le sens, pour être indépendant des événements de la conscience, est rapporté à un calcul exécuté par l'intermédiaire de ces événements ? Il faut examiner cette question, d'autant plus que la tradition analytique anglo-saxonne, comme nous l'avons vu avec Dreyfus, voit dans la phénoménologie husserlienne une figure précurseur du fonctionnalisme computationnel.

3.2 Représentation réelle et signification idéale

Comme nous l'avons souligné, l'acte fondateur d'une phénoménologie de la connaissance qui ne soit pas une étude psychologique de la conscience connaissance est la remarque que la signification n'est ni un fait de la conscience, ni une occurrence psychologique que l'on pourrait situer. La signification n'est pas une partie réelle de la conscience.

> la signification « n'est rien qui puisse valoir, au sens réel, en tant que partie de l'acte de compréhension correspondant. Ce vécu a naturellement aussi ses composantes psychologiques,

il est un contenu et se compose de contenus – au sens psychologique habituel. En font partie avant tout les éléments sensibles de ce vécu, les phénomènes linguistiques selon les contenus purement visuels, acoustiques, moteurs, et aussi les actes de l'interprétation objectivante qui situent les mots dans l'ordre de l'espace et de temps. » RL I

Le fait que la signification ne soit pas un partie réelle, un contenu réel de la conscience psychologique se manifeste dans le fait qu'elle ne coïncide pas avec la représentation mentale que l'on se fait à l'occasion d'un mot ou d'une idée. Cela est d'autant plus évident qu'il serait impossible de comprendre comment un esprit peut accéder deux fois à la même signification alors que les représentations mentales et les vécus réels de conscience sont à chaque fois différents. Husserl note ainsi que la signification d'un mot diffère de l'image mentale qu'il suscite en moi :

« Quand j'entends le nom de *Bismarck*, il est complètement indifférent pour la compréhension de ce mot dans sa signification unitaire, que je me représente le grand homme en chapeau et manteau ou en uniforme de cuirassier, que je me le représente dans mon imagination conformément à tel ou tel de ses portraits. Même le fait de savoir si se présentent ou non en général des images servant d'illustration intuitive ou animant indirectement la conscience de signification est sans importance. » RLI p. 98

Cette distinction est tout à fait fondamentale, et se retrouve également chez Frege. On sait d'ailleurs comment Frege, dans sa critique de la *la philosophie de l'arithmétique* d'Husserl, en avait sévèrement critiqué le psychologisme et de nombreux historiens de la philosophie en font découler le revirement antipsychologique du Husserl des *Prolégomènes*. Frege exprime, avant Husserl, cette distinction en opposant *pensée* et *représentation* mentale :

« La représentation associée à un signe doit être distinguée de la dénotation et du sens de ce signe. Si un signe dénote un objet perceptible au moyen des sens, ma représentation est un tableau intérieur, formé du souvenir des impressions sensibles et des actions externes ou internes auxquelles je me suis livré.[...] Chez le même individu, la même représentation n'est pas toujours lié au même sens. Car la représentation est subjective ; celle de l'un n'est pas celle de l'autre. [...] C'est par là qu'une représentation se distingue essentiellement du sens d'un signe. Celui-ci peut être la propriété commune de plusieurs individus : il n'est donc pas partie ou mode de l'âme individuelle. Car on ne pourra nier que l'humanité possède un trésor commun de pensée qui se transmet d'une génération à l'autre. » (Frege, 1971b, p. 105-106)

En effet, si la signification est la représentation mentale, comment comprendre que l'on puisse saisir deux fois de suite une même signification alors que la représentation mentale est à chaque fois différente ? Comment comprendre que deux personnes différentes puissent saisir la même signification ou pensée alors que leur histoire psychologique personnelle leur font associer des images ou représentations mentales différentes ? Comment comprendre l'intersubjectivité s'il n'existe que la subjectivité psychologique de la conscience ?

« Le donné psychologique est, à ce point de vue, comme on le sait, très divers, variant considérablement suivant les individus ; mais variant également pour le même individu, suivant les moments, et cela en ce qui concerne « un seul et même » mot. RL I [p. 97]

« [...]

> « Mais, en regard de cette multiplicité illimitée de vécus individuels, *ce qui* est exprimé en eux est partout quelque chose d'identique, le *même* au sens le plus strict du mot. La signification de la proposition ne s'est pas multipliée avec le nombre des personnes et des actes, le jugement, au sens logique idéal est un. » RL I p. [99]

Or, comprendre l'intersubjectivité de la signification, c'est comprendre qu'elle domine les variabilités individuelles, les variations factuelles des représentations, des contenus réels. Si le réel coïncide avec la contingence individuelle, l'intersubjectivité de la signification ne peut appartenir à la réalité de la conscience, mais seulement à son idéalité, c'est-à-dire au fait que la conscience est capable de saisir plusieurs fois de suite la même signification :

> « Or, cette véritable identité que nous affirmons ici n'est autre que *l'identité de l'espèce*. C'est ainsi, mais seulement ainsi, qu'elle peut, en tant qu'unité idéale, embrasser la multiplicité dispersée des singularités individuelles. Les singularités multiples formant la signification idéalement une sont naturellement les moments d'acte correspondants du signifier, les *intentions de significations*. La signification se comporte ainsi par rapport à chacun des actes du signifier (la représentation logique par rapport aux actes de représentation, le jugement logique par rapport aux actes de jugement, le raisonnement logique par rapport aux actes de raisonnement), en quelque sorte comme le rouge *in specie* par rapport aux bandes de papier que j'ai devant les yeux qui « ont » toutes le même rouge. Outre les autres moments qui la constituent (extension, forme, etc.), chaque bande a son rouge individuel, c'est-à-dire son *cas singulier* de cette espèce de couleur, tandis que celle-ci n'existe elle-même réellement ni dans cette bande, ni où que ce soit au monde ; ni surtout « dans notre pensée », en tant qu'elle aussi fait partie du domaine de l'être réel, de la sphère de la temporalité. » RLI p. [100-101]

La signification est donc *idéale* et n'appartient pas à la réalité vécue de la conscience. Il faudra désormais parler des réalités idéales non réelles vécues par la conscience : l'idéalité, visée par la conscience, y appartient à ce titre, mais évidemment elle n'en peut être une partie *réelle*, mais une partie seulement *idéale*.

On déduit donc qu'il existe une étude de la conscience qui échappe à la psychologie, l'étude des parties idéales et non réelles. Cette opposition s'articule entre le *ce qui est jugé* qui reste identique à lui-même, et l'*acte qui juge* qui vise ce pôle identique dans la multiplicité des jugements :

> « Mon acte de jugement est un vécu fugitif qui apparaît et disparaît. Tandis que ce qu'énonce l'énoncé, ce contenu *que les trois hauteurs d'un triangle se coupent en un seul point* n'est pas quelque chose qui apparaît et disparaît. Chaque fois que moi, ou qui que ce soit, prononce avec le même sens ce même énoncé, il y a un jugement nouveau. Les actes de jugement sont différents suivant les cas. Mais *ce qu'* ils jugent, *ce que* dit l'énoncé, est partout la même chose. C'est une chose identique au sens strict du mot, c'est une seule et même vérité géométrique. » RL. 1 p.[44]

Il faut donc poser une discipline permettant d'étudier la signification hors de toute considération psychologique qui la réduirait à un contenu réel qui ne pourrait prétendre en guise d'identité qu'à la seule identité numérique avec lui-même. Le projet d'une telle discipline n'est cependant pertinent que s'il est possible de considérer le contenu idéal du vécu indépendamment de son contenu réel. Car il est une chose de dire que le contenu idéal se distingue du contenu réel, il en est une autre d'affirmer que le contenu idéal peut être étudié indépendamment du contenu réel par une discipline autonome possédant ses propres lois : il faut pouvoir montrer que les lois de la signification sont indépendantes des lois du vécu, que les lois légalisant les contenus idéaux ne sont pas les lois légalisant les contenus réels.

C'est à cette condition qu'il sera possible de construire une grammaire et une logique, la première étant la théorie des conditions de possibilité de la signification, la seconde la théorie des conditions de possibilité de la vérité. Une phénoménologie du sens et de la signification, pour comporter une grammaire des significations et une logique de la vérité, doit montrer son indépendance au substrat réel. Or cette question est celle du fonctionnalisme intentionnel.

3.3 Le fonctionnalisme intentionnel

Il semble donc que les significations s'articulent indépendamment des contenus réels qui permettent de les viser ; les lois grammaticales et logiques de la signification ne sont pas des lois psychologiques décrivant les contenus réels des vécus, les lois du sens ne sont pas des lois du substrat réel. Il est donc tentant de définir le sens véhiculé par un contenu réel de la conscience comme une fonction intentionnelle remplie par ce contenu ; cette fonction qui, pour être dépendante du substrat dans la mesure où elle ne peut exister sans être véhiculée par lui, en est néanmoins indépendante dans la mesure où ce qui est véhiculé n'est pas conditionné par ce qui véhicule.

On retrouve donc les thèmes du fonctionnalisme en psychologie, à la nuance près, et qui est d'importance, où ce qui est véhiculé n'est pas une fonction syntaxique dans un programme, mais une fonction sémantique. Mais, si cette fonction sémantique obéit à des lois logiques et grammaticales, n'est-on pas autorisé à considérer qu'*in fine* ce fonctionnalisme sémantique ou intentionnel se ramène à un fonctionnalisme syntaxique, étant attendu que ce dernier possède toujours une irréductibilité au substrat ? Certains auteurs se sont risqués à accomplir un telle réduction, comme Hubert Dreyfus nous en a montré un exemple ci-dessus. Cela est pour notre propos de première importance car c'est précisément parce qu'il voit dans les lois de l'intentionnalité un fonctionnalisme sémantique réductible à un fonctionnalisme syntaxique que Dreyfus voit en Husserl un précurseur de l'IA, et donc son philosophe par excellence. Mais, avant d'évaluer la critique dreyfusienne, considérons ce qui chez Husserl autorise à parler d'un relatif fonctionnalisme sémantique ou intentionnel.

Sans prétendre être exhaustif en aucune façon, de nombreux écrits husserliens se prêtent à une telle interprétation. Nous retiendrons ici la VIe *Recherche logique* et les *Ideen I.*

Dans cette recherche, Husserl élabore une terminologie complexe et précise pour analyser la notion de représentation et de vécu intentionnel. Notamment, on trouve une opposition entre *représentant* et *représentation*, où le représentant, pour devenir une représentation, doit être animé par une intention pour viser l'objet signifié. Or, ce qui est fondamental pour nous, le représentant est dans sa nature même indifférent à ce qui est visé à travers lui. Son statut intentionnel ressortit entièrement à l'intention elle-même, sans dépendre de ce qu'est le représentant. Ainsi, s'il faut un représentant, son rôle de représentant ne dépend pas de lui. C'est là proprement un fonctionnalisme. Reprenons donc cette terminologie et distinguons *représentant* et *représentation.*

Les recherches logiques sont bâties sur la notion d'acte de la conscience, et de vécu.

Husserl définit le vécu à partir du *cogito* cartésien :

> « Prenons pour point de départ la conscience, entendue au sens fort, celui qui se présente d'emblée, et que nous désignerons de la façon la plus simple par le *Cogito* cartésien, par le « je pense ». On sait que Descartes l'entendait en un sens si large qu'il y incluait tout ce qui s'énonce par les formules : « Je perçois, je me souviens, j'imagine, je juge, sens désire, veux », et de même tous le autres vécus subjectifs semblables, avec leurs ramifications innombrables et fluantes. [...] Considérons les vécus de conscience, *avec toute la plénitude concrète* selon laquelle ils s'insèrent dans leur contexte concret – *le flux du vécu* – et s'y adjoignent en vertu de leur propre essence. »

En particulier, un acte est un vécu intentionnel où la conscience vit à travers son vécu dans la visée d'un objet, un vécu où la conscience est la conscience de quelque chose :

> « Par « *actes* » on doit entendre les vécus du signifier, et ce qui est significatif dans chaque acte particulier doit précisément résider non dans l'objet mais dans le *vécu d'acte*, il doit résider dan ce qui fait de celui-ci un vécu « *intentionnel* », « *dirigé* » sur des objets. » RL V p.[344]

Dans la visée intentionnelle d'un objet, il faut distinguer la matière intentionnelle d'une part et la qualité intentionnelle d'autre part. La matière intentionnelle consiste dans la manière de viser l'objet en tant que quelque chose, par exemple Napoléon (objet visé) en tant que vainqueur d'Iéna[7], la qualité intentionnelle dans la modalité associée à la visée : *je crois* que Napoléon est le vainqueur d'Iéna, ou *j'en suis sûr*. Ainsi, la matière intentionnelle est l'objet visé dans le comment de sa visée, tandis que la qualité est la thèse que je pose à propos de cette manière de viser.

Ces matières et qualités intentionnelles sont une visée intentionnelle animant des contenus réels de la conscience, contenus primaires dans la terminologie des *Recherches logiques*, data hylétiques de sensation dans celle des *Ideen*. Ces contenus primaires constituent le représentant que doit animer une visée intentionnelle le rapportant à une objectité. Dans la terminologie des *Ideen* :

> « il faut pourtant distinguer deux choses :
>
> 1. Tous les vécus qui dans les *Etudes logiques* étaient désignés du nom de « contenus primaires » ;
> 2. Les vécus ou moments du vécu, qui portent en eux la propriété spécifique de l'intentionnalité.
>
> Au premier groupe appartiennent certains vécus « *sensuels* » formant une unité en vertu de leur genre suprême : les « *contenus de sensation* » tels que les data de couleur, les data de toucher, les data de son, etc., que nous ne confondrons plus avec les moments des choses qui apparaissent, avec la qualité colorée, la qualité rugueuse, etc., qui au contraire « s'esquissent » de manière vécue au moyen des précédents. De même les sensations de plaisir, de douleur, de démangeaison, etc., ainsi que les moments sensuels de la sphère des « impulsions ». De tels vécus concrets rentrent comme composantes dans un nombre beaucoup plus vaste encore de vécus concrets qui, considérés comme totalités, sont intentionnels, en ce sens que par delà ces moments sensuels on rencontre une couche qui pour ainsi dire « anime », *leur donne un*

7. Pour reprendre un exemple de Frege...

> *sens* (ou qui implique essentiellement une donation de sens) ; c'est par le moyen de cette couche, et à partir de *l'élément sensuel qui en soi n'a rien d'intentionnel*, que se réalise précisément le vécu intentionnel concret. » *Ideen I* (Husserl, 1950, p. [172)

Mais cette animation peut revêtir deux modes principaux : le mode signitif et le mode intuitif. Dans le premier cas, l'objet visé est visé à vide, sans qu'il soit présent ou donné en personne, dans le second cas, la visée est remplie car l'objet visé est présent, donné en personne dans le vécu intentionnel :

> « Je pense, par exemple, à un toit couvert de tuiles rouges, sans le voir, sans même l'imaginer [intention signitive, à vide]; je lève les yeux, et je vois ce toit tel que je l'ai pensé. Il y a ici une réalisation de ma pensée première – le toit, tel que je l'ai pensé, est directement devant mes yeux, est présent *en personne*. [intention intuitive remplie] » (Lévinas, 1988, p.113)

Ces deux types d'animation intentionnelle reposent sur deux types de représentants, les contenus intuitifs et les contenus signitifs :

> « Or, nous trouvons, d'une façon générale, en présence de la possibilité de la distinction phénoménologique suivante :
>
> 1. le *contenu purement intuitif* de l'acte, comme étant ce qui, dans l'acte, correspond à la totalité des déterminations de l'objet qui « entrent dans le phénomène » ;
> 2. le *contenu signitif* de l'acte, correspondant parallèlement à la totalité des autres déterminations qui sont, sans doute, visées dans le même temps, mais qui elles-mêmes n'entrent pas dans le phénomène.
>
> » RL VI, p.[80]

On a donc un contenu signitif permettant de viser à vide un objet intentionnel et un contenu intuitif permettant de le viser en plein. La question que l'on peut se poser est de savoir s'il existe une relation entre le contenu signitif et l'objet intentionné d'une part, et le contenu intuitif et l'objet intentionné d'autre part. En clair, la question que l'on se pose est de savoir s'il existe une dépendance entre le sens d'une représentation et la nature du représentant. Or, Husserl écrit :

> « Nous partirons de cette remarque que la *représentation signitive* établit, entre la matière et le représentant, une *relation extérieure contingente*, tandis que la représentation intuitive établit une *relation essentielle, interne*. La contingence consiste, dans le premier cas, en ce qu'une signification identiquement la même doit être présumée attachée à chaque contenu quel qu'il soit. La *matière significative* a seulement besoin, d'une manière générale, d'un contenant lui servant de support, mais, *entre son individualité spécifique et sa composition spécifique propre, nous ne trouvons aucun lien de nécessité*. La signification ne peut pas être, en quelque sorte, suspendue en l'air, mais, pour ce qu'elle signifie, le signe dont nous disons qu'elle est la signification est absolument indifférent. » RL VI p. [92]

Considérons précisément la dernière phrase de l'extrait cité : « La signification ne peut pas être, en quelque sorte suspendue en l'air ». Cela signifie qu'il ne peut exister de représentation sans représentant, d'animation intentionnelle sans contenu primaire ou matière hylétique, pas de *fonction intentionnelle* sans support réel, dans le sens que Husserl donne à ce terme, c'est-à-dire un support effectivement présent dans la conscience, que pourrait étudier la psychologie empirique puisque seul le réel survit à la réduction naturaliste.

Mais, « pour ce qu'elle signifie, le signe dont nous disons qu'elle est la signification est absolument indifférent. » . Le sens du représentant ne dépend donc pas du représentant lui-même, l'animation intentionnelle n'est pas conditionnée par la nature ou le type du représentant, le sens idéal ne dépend pas du support réel. Par conséquent, on peut dire que la fonction intentionnelle est une propriété affectée au support du seul fait de l'animation intentionnelle, en fonction du rôle qu'elle joue vis-à-vis de l'objet visé et des autres manières de le viser, mais indépendamment de la « réellité » du substrat de la représentation.

Cette phrase est donc une illustration avant la lettre des thèses du fonctionnalisme computationnel. Il existe bien sûr une différence de taille sur laquelle il nous faudra évidemment revenir : alors que le fonctionnalisme computationnel est syntaxique, c'est-à-dire que le sens du signe ou symbole est défini en fonction de sa combinatoire calculatoire avec les autres signes indépendamment de ce qu'il signifie sémantiquement, c'est-à-dire intentionnellement, le fonctionnalisme husserlien serait un fonctionnalisme sémantique où le sens du symbole serait défini en fonction de son rôle intentionnel, conformément à sa sémantique et non indépendamment d'elle.

Mais, et ce sera l'objet de notre prochaine section, il est possible sans doute de montrer que ces fonctionnalismes sont en fait identiques, puisque la seule manière de penser le fonctionnalisme sémantique des significations est d'en faire la logique. La logique étant l'outil mathématique permettant de penser la sémantique depuis la syntaxe, l'entreprise husserlienne d'élaborer une logique des significations n'est peut-être pas autre chose qu'une réduction du fonctionnalisme sémantique au fonctionnalisme syntaxique. Faudrait-il voir alors en Husserl le philosophe permettant de penser le cognitivisme computationnel et par un curieux retour des choses, voir en lui un chantre du naturalisme computationnel ? Ce serait voir en lui un contempteur du psychologisme davantage irrité par sa non scientificité que par son naturalisme et, selon nous, passer à coté d'un trait fondamental de la phénoménologie. Mais il ne faut pas nier une certaine pertinence à cette lecture dans la mesure où elle repose à notre avis sur d'authentiques ambiguïtés de la phénoménologie transcendantale qui ont été de nombreuses fois relevées, notamment le formalisme husserlien. Nous y reviendrons dans cette prochaine section.

Mais, auparavant, nous voudrions revenir sur le fonctionnalisme sémantique des contenus signitifs et des représentations qu'ils permettent de constituer. En effet, Husserl affirme pour les contenus intuitifs une thèse inverse, une thèse qui infirme toute attitude fonctionnaliste. En effet :

> « Il en est tout autrement dans le cas de *représentation purement intuitive*. Il y a ici un *rapport interne nécessaire entre la matière et le représentant*, rapport déterminé par le contenu spécifique de l'une et de l'autre. Seul peut servir de représentant intuitif d'un objet un contenu qui lui soit semblable ou identique. Pour parler un langage phénoménologique : le *quid* (als was) du contenu que nous appréhendons (son sens d'appréhension), nous ne sommes pas entièrement libres d'en décider ; et cela, non seulement pour des raisons empiriques – car toute appréhension, même l'appréhension significative, est empiriquement nécessaire – mais parce que le contenu à appréhender nous fixe des limites par une certaine sphère de ressemblances et de similitudes, donc par son contenu spécifique. » p. [92-93] RL VI

La signification ne se comporte donc pas comme le remplissement intuitif dans la mesure où le support réel du remplissement conditionne dans sa nature l'objectité intuitionnée. Faut-il voir dans cet extrait une régression de la phénoménologie à un sensualisme où l'on confond l'apparition et l'apparaissant, le rouge senti et le rouge perçu? Nous ne le croyons pas. Ce qu'il faut sans doute comprendre dans cette similitude entre le représentant et le représenté est que la visée intentionnelle n'est pas un fonctionnalisme sémantique dans la mesure où un tel fonctionnalisme ne sait pas rendre compte de la singularité de l'apparition en personne de l'objet intentionné, ne sait pas rendre compte de l'originarité de l'intuition.

Cela signifie que l'arbitraire du représentant réel ne saurait être absolu et être indépendant de la perception. Le fonctionnalisme sémantique comme logique des significations ne doit donc pas se confondre avec un système de significations, système au sens saussurien, où les significations seraient indépendantes parce que constituant un système de contenus ou de signifiés distinct du système des signifiants (plan de la substance et plan de la forme, comme on l'a vu dans le chapitre 4). Il faut donc résister à la tentation de sémantiser la phénoménologie en un système de signifiés. Il faudra bien plutôt phénoménologiser la sémantique dans la mesure où l'arbitraire du signifié est incapable de rendre compte de la donation de sens. Car si un contenu intuitif peut fonctionner comme un contenu signitif comme le rappelle Lévinas :

> « *Le rôle significatif ou intuitif du représentant, dépend donc de l'intention qui l'anime, du sens irréductible que lui prête l'intentionnalité.* [...] Quoique la structure interne de l'acte significatif se caractérise par l'indifférence du *représentant* à l'égard de l'objet que sa matière signifie, quoique d'autre part, il y ait dans l'acte intuitif, un rapport d'analogie entre l'un et l'autre, ce qui distingue ces deux actes c'est, en dernière analyse, le mode de liaison du *représentant* à la *matière*. Les *représentants*, même analogues à leurs objets, peuvent, d'après Husserl, fonctionner comme *contenus significatifs.* » (Lévinas, 1988)

il semble en revanche qu'un contenu signitif ne peut fonctionner comme un contenu intuitif. Par conséquent, il ne peut s'agir d'un retour à un sensualisme puisque ce qui est déterminant c'est « le mode de liaison du *représentant* à la *matière* », et non la nature intuitive ou signitive du représentant. Cela indique que les appellations « signitive » et « intuitive » ne sont pas des propriétés de la struture réelle du contenu primaire ou représentant, mais des noms donné au rôle joué par le représentant lors d'une animation intentionnelle : il fonctionne à vide ou à plein. Ceci correspondrait donc à un fonctionnalisme intentionnel. Mais, l'insistance sur l'analogie rappelle que l'arbitraire de la signification d'un représentant doit se plier au rapport unissant l'apparition de l'objet à l'objet apparaissant, et montre que la phénoménologie n'est ni un sensualisme, ni un fonctionnalisme.

Cette conception paraît être confirmée par les remarques de Husserl sur les sentiments dans la cinquième recherche, concernant le problème de savoir si ce sont des vécus intentionnels ou non. Husserl distingue la *sensation* de plaisir par exemple du *sentiment* de plaisir. Alors que le premier est un contenu réel dépourvu d'animation intentionnelle, le second est intentionnel dans la mesure où il porte sur un objet perçu comme plaisant :

> « Les sensations de plaisir et de douleur peuvent persister, alors que les caractères d'actes édifiés sur elles ont disparu. Lorsque les faits générateurs de plaisir ont passé à l'arrière-plan, quand ils ne sont plus aperçus avec leur coloration affective, que peut-être ils ne sont même plus du tout des objets intentionnels, l'excitation du plaisir peut cependant durer encore assez longtemps ; il peut alors arriver qu'elle soit elle-même ressentie comme agréable ; au lieu de

> figurer comme représentant d'une qualité agréable de l'objet, elle est désormais simplement rappportée au sujet qui la ressent, ou bien alors elle devient elle-même un objet représenté ou agréable. » p. [395] RL V

Cela signifie que quand l'animation intentionnelle délivrant un objet intentionnel comme plaisant ou agréable cesse, la sensation ne perd pas pour autant toute détermination: elle reste une sensation de plaisir. La nature de la sensation n'est donc pas indifférente à l'animation intentionnelle dont elle sera le support : une sensation de plaisir détermine l'intention à être une intention de plaisir.

Il existe une communauté de principe entre les sensations affectives et les vécus intentionnels (les actes) affectifs. Un fonctionnalisme au sens strict affirme une indépendance de la fonction vis-à-vis de la structure sous-jacente qui la réalise. Or Husserl considère la question de séparer radicalement les sensations affectives des actes du même nom en ne voyant dans ce cas qu'une équivoque verbale :

> « Le fait qu'elles sont, à la rigueur, des contenus figuratifs ou encore des objets d'intentions, mais non elles-mêmes des intentions, témoigne d'une différence descriptive si essentielle que nous ne pouvons songer sérieusement à soutenir qu'il y a ici l'unité d'une genre véritable. Sans doute parle-t-on également de « sentiments » dans les deux cas, dans les actes de plaisir dont nous avons parlé plus haut, comme dans ces sensations dont il est ici question. Mais cette circonstance ne saurait nous embarrasser, pas plus que nous ne nous laissons abuser, en ce qui concerne les sensations tactiles, par l'expression courante de sentir, employée au sens de toucher. » RL V

Ainsi, il ne faudrait pas se « laisser abuser » par une communauté de terme pour y voir une communauté générique. Cela voudrait dire qu'il n'existe pas de rapport entre le fait pour une sensation d'être une sensation de plaisir et le fait pour une sensation d'être animée intentionnellement en un acte de plaisir ; la notion de « sentiment » pour qualifier tant la sensation que l'acte reposant sur une pure équivoque verbale. Mais nous avons vu que cette independance sensations/actes n'est pas tenable et qu'il existe bien une dépendance entre la fonction intentionnelle animant la sensation et la nature de la sensation : il existe une dépendance du type intentionnel au type sensitif. Par conséquent, si le fonctionnalisme correspond à une identité type/occurrence, l'animation intentionnelle ne saurait être un fonctionnalisme et nous devons affirmer avec Husserl que :

> « Nous ne saurions donc non plus admettre qu'il n'y ait entre les tendances ou les désirs intentionnels et les non intentionnels aucun rapport de communauté générique, mais seulement un rapport d'équivoque. » p.[396] RL V

Il ne faut évidemment pas conclure qu'Husserl affirme une identité des types sensitifs aux types intentionnels. La principale originalité des Recherches Logiques et notamment de la V^{e} est d'affirmer que le rouge senti n'est pas le rouge perçu, que l'apparition n'est pas l'objet apparaissant. Par conséquent, il n'y a nulle correspondance bijective entre l'objet intentionnel et les données sensitives sur lesquelles sa donation s'appuie. La correspondance entre l'intention animatrice et la sensation animée n'est donc ni l'identité, ni la fonctionnalité.

Il n'en demeure pas moins que le fonctionnalisme sémantique est inscrit dans son principe dans la phénoménologie transcendantale, dans la mesure précise où elle est transcendantale d'ailleurs. Dépasser son transcendantalisme, c'est dépasser l'arbitraire de la signification et de l'idéalité pour la faire dialoguer avec l'instance singulière de la donation intuitive originaire, pour constituer et faire évoluer l'idéal dans un dialogue avec le *reell.*

3.4 Le positivisme des contenus réels

La discussion du fonctionnalisme présumé de la phénoménologie husserlienne renvoie à la question du statut des contenus réels. La principale contribution de la phénoménologie consiste en effet dans son étude de l'intentionnalité comme renvoi et projection du sujet vers une transcendance objective. La transcendance objective se constitue dans la conscience par l'animation intentionnelle des contenus immanents dits réels.

L'intentionnalité permet de dépasser l'opposition séculaire du sujet et de l'objet en montrant comment l'objet est construit, se constitue dans la conscience. La conscience se rapporte directement à l'objet, et dans la perception, mode privilégié de l'intentionnalité, l'objet perçu est donné directement, il est présent en chair et en os. L'aporie sujet/objet est dépassée en montrant que le sujet n'est qu'un sujet pour un objet, et que l'objet n'est objet que pour un sujet. En outre, il semble bien que l'intentionnalité soit en fait un dépassement du paradigme de la connaissance comme représentation.

En effet, dans le paradigme de la connaissance comme représentation, la représentation est ce qui permet à un sujet de se rapporter à un objet. Mais c'est aussi ce qui fait obstacle dans le rapport à l'objet, si bien que le sujet a davantage affaire à ses représentations qu'à l'objet de ses représentations, au représentant qu'au représenté, au contenant qu'au contenu.

La force de la phénoménologie est de réaffirmer avec vigueur que ce qui est connu n'est pas la représentation, mais l'objet ou l'état de chose visé par la représentation. En rappelant que la sensation est vécue et non perçue, et que l'objet est perçu et non vécu, Husserl affirme que la connaissance ne consiste pas à se rapporter à un décalque de la réalité. C'est pourquoi, contrairement à ce qui souvent affirmé, la phénoménologie ne peut pas être un solipsisme dans la mesure où ce que connaît le sujet est directement le monde, et non la copie représentative qu'il en possède. Le solipsisme de la phénoménologie désigne alors simplement le point de vue privilégié de la conscience que l'on adopte dans une problématique de la conscience.

Mais, et c'est là où l'on peut sans doute voir une contradiction, ou un inachèvement de la phénoménologie husserlienne, ce renvoi à l'extériorité repose sur les contenus sensitifs qui doivent s'animer intentionnellement pour que vivre les contenus sensitifs permette de percevoir les objets du monde. Cela signifie que l'extériorité ne peut s'atteindre que par la présence interne de contenus sensitifs. Or cela voudrait alors dire que les contenus sont des représentations matérielles (au sens de data hylétique) nécessaires pour que la conscience puisse se rapporter à ce qui n'est pas elle, que les parties réelles des vécus intentionnels sont nécessaires pour leur conférer l'intentionnalité qui leur associe des contenus non réels, l'essence (*Recherches Logiques*) ou noème (*Ideen*).

La nécessité de recourir à des parties *réelles* de la conscience pour que celle-ci puisse les animer intentionnellement est-elle une concession au positivisme ? S'agit-il d'un psychologisme larvé de la phénoménologie qui ne pourrait porter à bout sa critique de la représentation ? C'est en tout cas l'avis de nombreux commentateurs (Barbaras, 1994), et l'évidence textuelle permettant de soutenir ce point de vue est abondante.

Pourtant, même si l'on ne peut contester le bien fondé de cette critique, il nous semble qu'elle ne puisse être aussi radicale, et prétendre ainsi que le naturalisme psychologique chassé par la porte de l'intentionnalité revienne par la fenêtre des contenus réels. En effet, la rupture avec le psychologisme ne vient pas tant de la considération des parties non réelles des vécus que de la volonté d'étudier les vécus non comme des faits empiriques, mais comme des faits de la conscience. Et, ce que l'on considère comme l'essence même de la phénoménologie transcendantale, à savoir qu'il est possible d'abstraire du vécu des parties idéales non réelles, n'est alors que la conséquence du statut non empirique des vécus.

Par conséquent, les contenus réels ne sont pas plus naturels ou positifs que les contenus intentionnels, ils sont immanents au vécu alors que la partie idéale y est transcendante. C'est la raison pour laquelle l'immanence des parties réelles implique qu'elle partage la particularité du vécu et que seule une phénoménologie noétique comme par exemple celle des *Recherches Logiques* peut les étudier comme telles, alors que la transcendance des parties non réelles implique leur idéalité.

Mais, si l'on s'abstient de tout naturalisation empirique des vécus, fussent-ils non intentionnels, donc sensitifs, on peut se demander comment il est possible de les thématiser : vécus sans jamais être perçus, ne sont-ils pas alors un principe caché de l'intentionnalité dont l'évidence phénoménologique est toujours refusée et dont l'existence ne peut alors être que déduite ? Cela renouerait avec un thème authentiquement transcendantal où les conditions de l'expérience ne peuvent être que déduites puisque seule l'expérience peut être constatée avec évidence, de manière ostensive, alors que ses conditions ne peuvent être exhibées mais seulement déduites apagogiquement.

Mais la déduction transcendantale des conditions *a priori* de l'expérience est kantienne et non husserlienne : même si Kant, dans la déduction subjective aborde une phénoménologie de la pensée scientifique, cela ne saurait faire oublier qu'il s'agit là d'un exercice certes méritoire[8]mais non essentiel pour la philosophie critique. Cette dernière serait, d'après Kant lui-même, tout aussi complète et achevée si elle devait se borner à la déduction objective des catégories, *i.e.* des conditions *a priori*. Mais la déduction transcendantale n'est pas phénoménologique car elle ne repose que sur l'évidence médiate du raisonnement et non sur le voir immédiat de la conscience phénoménologique. La question est donc de savoir comment il est possible de thématiser les contenus réels alors que l'on ne peut que percevoir par leur intermédiaire sans que jamais ils ne puissent être perçus pour eux-mêmes, car, comme le répète inlassablement Husserl, ils sont vécus mais non perçus.

Il faut donc trouver quelle raison *phénoménologique* permet de les thématiser. Une première raison nous est fournie encore une fois par la discussion sur les sentiments, intentionnels ou non. On constate que quand des sentiments intentionnels (actes affectifs) perdent leur caractère intentionnel pour ne plus devenir que des vécus sensitifs (sensations affectives), ces derniers sont attribués au sujet : « au lieu de figurer comme représentant d'une qualité agréable de l'objet, elle est désormais simplement rappportée au sujet qui

8. « Car si l'on peut prouver QUE les catégories dont la raison doit se servir pour toute sa conaisance, ne peuvent être utilisées que par rapport aux objets de l'expérience (parce que c'est dans ce domaine seul qu'elles rendent possible la forme de la pensée), la réponse à la question de savoir *comment* elles la rendent possible, est assurément assez importante pour *achever* si possible cette déduction, mais elle n'est nullement *nécessaire*, elle est simplement *méritoire* en ce qui concerne la fin principale du système qui est de déterminer les limites de la raison pure. » (Kant, 1990, p. 17, note)

la ressent, ou bien alors elle devient elle-même un objet représenté ou agréable. » p. [395] RL V.

On sait que dans les *Recherches Logiques*, la cinquième notamment, Husserl refuse la notion de moi et confond le sujet avec le flux de la conscience. Le moi, le sujet, est donc la totalité constituée par l'ensemble des vécus. Pour notre propos, cela signifie qu'il n'existe pas de moi transcendantal sujet de tous les vécus, mais un moi empirique ou psychogique constitué dans les vécus. Or, la dernière indication de Husserl que nous avons relevée laisserait penser que les contenus primaires, hors d'une animation intentionnelle les éclatant vers un objet intentionnel, ne sont pas pour autant dépourvus de toute intentionnalité dans la mesure où ils participent à la construction du moi empirique, dans la mesure où on les « rapporte au sujet qui les ressent ». Nous pensons être autorisé à parler d'intentionnalité dans la mesure où, s'il est clair que l'intentionnalité des sensations ne saurait être confondue avec l'intentionnalité des vécus intentionnels, les sensations sont dépassées vers la constitution du sujet. Il s'agit bien de constitution puisque le sujet n'est pas premier, n'est pas une condition transcendantale de la conscience. La conscience comme flux des vécus est a-subjective.

C'est aussi pourquoi lorque la réflexion de la conscience amène sous son regard le vécu lui-même en une perception adéquate, les contenus réels peuvent être considérés en eux-mêmes et caractérisés comme des parties réelles, incluses dans l'immanence de la conscience, car constitutive du sujet comme flux. Le fait que le moi ne soit que l'unité du flux implique que la perception adéquate du vécu soit une perception du moi, car le vécu est perçu en tant que vécu appartenant au flux. C'est donc la raison pour laquelle les contenus réels, considérés en eux-mêmes, c'est-à-dire lors d'une réflexion, correspondent à une apparition du moi comme tel.

Il y a donc un parallèle entre notre démarche et celle de Husserl dans les *Leçons* sur le temps: pour considérer les contenus réels en eux-mêmes, il faut trouver une perception qui ne soit pas intentionnelle. Notre formulation est boîteuse, puisque toute perception étant perception de quelque chose, toute perception est intentionnelle. Il s'agit plutôt d'une conscience non intentionnelle visée dans une perception interne intentionnelle. Or quelle peut être un vécu authentiquement non intentionnel donné à la conscience comme tel? Ce ne peut être un contenu sensitif où comme on le constate bien à travers les formulations compliquées de Husserl, les contenus sensitifs ne se délivrent qu'à l'occasion de perception d'un objet (réel (*real* ou non) mais jamais pour eux-mêmes. Comme le remarque R. Barbaras:

> « Autrement dit, il est légitime de questionner la pertinence de la distinction entre le donné sensible immanent et l'aspect correspondant de l'objet, entre le rouge comme vécu et le rouge comme spatial. Que peut bien signifier en effet un rouge qui serait strictement vécu sans être perçu, c'est-à-dire sans être ce rouge là-bas, le rouge de cet objet . Cette distinction, sur laquelle Husserl a d'ailleurs longuement hésité, apparaît comme une concession à l'idée empiriste des sensations, contenus sensibles immanents. » (Barbaras, 1994, p.50)

C'est pourquoi Husserl ne peut les nommer que par les noms des objets perçus : le rouge senti est défini à partir du rouge perçu. Mais, le sentiment nous donne l'occasion d'un vécu non intentionnel donné comme tel à la conscience, qui n'est pas abstrait d'un autre vécu. Le problème avec les contenus sensitifs, c'est qu'ils sont à la fois réels et abstraits: ils sont abstraits des vécus concrets comme des parties réelles. Mais, pour étudier les contenus réels comme tels, il nous faut les définir comme réel et concret : il nous faut un vécu concret, c'est-à-dire complet, pour que la réellité puisse être caractérisée autrement que par abstraction depuis un vécu concret où l'intentionné sert à définir le support de l'intention. De la même manière, Husserl dans les *Leçons sur la conscience intime du temps* choisit le vécu du son pour étudier sa temporalité, puisque son caractère vécu concret est de coïncider avec cette temporalité, si bien que le constitué ne masque pas dans ce cas le constituant.

En conclusion, notre interprétation est la suivante. En tant que substrat de l'animation intentionnelle, les contenus réels correspondent en effet à la donnée préalable de données positives résultant de l'interaction sensible du corps avec son environnement. Mais, ce que notre discussion a permis de dégager est qu'il est impossible de thématiser en lui-même et pour lui-même le contenu sensible. En d'autres termes, le contenu primaire hylétique n'est pas un fait positif dans la mesure où il ne possède pas une positivité, une évidence positive, suffisante permettant de le considérer indépendamment du sens intentionnel qui l'anime.

Cela signifie que quand Husserl considère les contenus hylétiques en « eux-mêmes », il ne les considère pas comme des faits positifs mais comme le support d'une animation intentionnelle qui n'est pas celle dont on fait abstraction lorsque l'on dégage ces contenus primaires. Ainsi, quand dans la perception sensible, Husserl distingue les contenus sensibles vécus indépendamment des contenus objectifs perçus, quand il fait abstraction de l'intentionnalité qui les transcende vers une extériorité objective, il n'établit pas la positivité des contenus primaires mais permet de mettre en évidence *une autre intentionnalité* que l'intentionnalité transcendante masque : l'intentionnalité qui constitue le corps propre. Quand en effet les contenus ne sont pas renvoyés à l'objectivité transcendante dont ils sont le support intentionnel, ils ne sont pas considérés pour autant pour et en eux-mêmes, mais à travers une intentionnalité qui constitue le corps sentant, le corps vivant, le corps propre.

Il n'existe et ne peut exister de positivité *dans* la conscience dans la mesure où il ne peut exister de données immanentes que la conscience pourrait considérer en elles-mêmes. Il est de l'essence de la conscience d'animer intentionnellement tout ce qu'elle touche : autrement dit elle ne peut considérer ce que sont les choses indépendamment de ce qu'elles veulent dire, c'est-à-dire ce que sont les choses en elles-mêmes indépendamment de ce qu'elles sont pour la conscience. Le slogan phénoménologique du « retour aux choses elles-mêmes » ne signifie donc pas que l'on veut considérer les choses telles qu'elles sont *ontologiquement*, comme seul Dieu peut le faire, mais les choses telles qu'elles sont *phénoménologiquement*, c'est-à-dire telles qu'elles apparaissent en elles-mêmes à la conscience.

La conscience donne du sens à tout ce qu'elle considère : elle reconnaît comme étant quelque chose, comme ayant un sens, comme étant différente d'elle tout ce qui lui apparaît. La « prise de conscience » signifie qu'il existe un « choc », le choc de la conscience, par lequel la conscience « aperçoit » qu'elle contient quelque chose qu'elle n'est pas : exister pour la conscience est se projeter vers l'extériorité en « prenant conscience » du sens de ce qui se trouve « en elle » : cette dernière expression est impropre puisque la seule manière de caractériser *ce qui est dans la conscience* consiste à le rapporter, comme support de l'animation intentionnelle, à *ce qui n'est pas dans la conscience*.

Il ne peut donc exister d'intériorité de la conscience, ni d'intimité de la conscience au sens propre. La conscience est d'emblée une projection vers l'extériorité. Mais cette extériorité à la conscience peut être *interne* quand il s'agit de l'animation intentionnelle constituant le corps propre ou *externe* quand il s'agit de l'animation intentionnelle constituant les objets mondains. Le plus difficile à comprendre est que cette projection ne s'effectue pas à partir de quelque chose, comme le voudrait une compréhension naïve et ontologisante des choses, mais à partir de rien. Rappelons nous en effet de Descartes : une idée n'est une idée non pas parce qu'elle est, mais par ce qu'elle n'est pas : en elle-même un idée ne se distingue pas du *Cogito* dans lequel elle se confond avec toutes les autres idées ; mais par ce qu'elle représente, par donc ce qu'elle n'est pas, l'idée tire une identité qui la distingue de toutes les autres idées. L'idée est donc un pur renvoi, une pure projection sur une réalité, fût-elle seulement représentative, qui se distingue de la conscience. Si l'on résiste à la tentation de faire de la conscience une réalité mondaine, une « chose qui pense », et que l'on remarque, avec Husserl, que la conscience ne peut se distinguer des idées qui la peuplent, il en résulte que la conscience n'est rien par elle-même et qu'elle consiste toute entière dans sa sortie hors d'un « elle-même » alors qu'elle n'est rien par elle-même[9] .

Cette discussion permet de reformuler la problématique originale de l'intentionnalité husserlienne et de l'interpréter de manière totalement phénoménologique, c'est-à-dire sans des vestiges positivistes où des données auraient une signification par elles-mêmes en tant qu'hylétiques ou sensibles indépendamment d'un sens dont elles seraient le support. L'intentionnalité n'est donc pas le problème de savoir comment constituer ce qui est présent *à* l'esprit à partir de ce qui est présent *dans* l'esprit. Il s'agit plutôt de comprendre que la conscience n'a accès qu'au sens et que toute réalité se manifeste à elle comme une constitution du sens. C'est pourquoi considérer les matériaux de cette constitution en dehors de cette constitution est discutable : cela reviendrait à considérer que la structure des traces écrites sur le papier puisse être déterminée indépendamment du sens qu'elles possèdent en tant qu'écriture.

Si notre lecture peut être une interprétation de la phénoménologie husserlienne, il est clair cependant qu'elle n'est pas husserlienne. Elle constitue une réponse à des ambiguïtés qui, si elles n'ont pas été tranchées par Husserl lui-même, ne sauraient trouver au sein même de la phénoménologie transcendantale leur résolution.

Notamment, l'affirmation que la conscience ne possède pas d'intimité mais seulement une extérioritéentre en contradiction avec flagrante avec les analyses de Husserl concernant la temporalité puisque cette dernière relève exclusivement de l'intimité de la conscience. Cela signifie que notre interprétation qui refuse de voir dans les contenus immanents un vestige de positivisme et qui prône une extériorité de la conscience doit se confronter à l'analyse husserlienne de la temporalité.

C'est en effet dans les *Leçons sur la conscience intime du temps* que Husserl élucide les structures intentionnelles permettant de constituer les contenus immanents. C'est dans la temporalité que se détermine la structure de l'intériorité de la conscience. Or, c'est là, qu'à notre sens, gît le véritable vestige positiviste de Husserl : l'affirmation d'une intériorité de la conscience constituée à partir d'un flux *auto-constitué* d'impressions ponctuelles. Cette auto-constitution permet d'affirmer l'unité pour lui-même du flux de la conscience, d'affirmer comme un fait positif la conscience elle-même. Cette auto-constitution renoue pour ainsi dire avec Descartes dans la mesure où, comme lui, Husserl affirme que la conscience est une « chose qui pense », c'est-à-dire un flux auto-constitué qui constitue les vécus immanents de la conscience.

9. Nous reprenons cette question dans le chapitre suivant (section 7-3.1) où nous insistons sur le fait que la concience ou l'esprit est une projection sur l'extériorité qui ne possède pas d'intériorité.

L'enjeu est d'importance puisque c'est l'affirmation de l'intériorité de la conscience qui permet de considérer les lois propres de la conscience : qu'il existe des lois caractérisant la manière de penser et de vivre le sens indépendantes de la matière de ce qui est pensé. Si la conscience est en effet pure extériorité car projection à partir de rien vers ce qu'elle constitue à partir d'apparitions qu'elle anime, elle ne peut faire l'objet de lois qui vaudraient pour tout ce qu'elle vise. Au contraire, elle devrait emprunter la légalité à chaque fois régionale de ce qu'elle constitue. Cela reviendrait à affirmer qu'il ne peut exister de sphère du formel et que la pensée ne se partage qu'entre régions matérielles.

Il est d'autant plus nécessaire d'aborder la question du temps que le niveau des *Recherches logiques* et des *Ideen* est impropre à décider de la nature du fonctionnalisme husserlien. Nous venons de voir que la question du fonctionnalisme sémantique est pour le moins confuse et que même le problème du fonctionnalisme syntaxique, pour lequel, comme nous l'avons vu plus haut, la situation semble plus claire, ne peut être tranché sans faire appel à l'analyse de la temporalité.

3.5 Le fonctionnalisme syntaxique

La phénoménologie husserlienne possède une dimension fonctionnaliste dans la mesure où les contenus primaires ou hylétiques sont animés de fonctions intentionnelles dans le cadre des noèses pour viser l'objet à travers le noème, composante non réelle du vécu. La question que l'on peut se poser est de savoir si, comme le prétend Dreyfus, ce fonctionnalisme sémantique, peut se rapporter à un fonctionnalisme syntaxique, dans un sens computationnel. En d'autres termes, peut-on considérer Husserl comme un précurseur de la psychologie cognitive ?

Il faut reconnaître que tout semble aller dans ce sens si l'on s'en tient à une analyse rapide du « formalisme » husserlien. En effet, Husserl établit les lois d'une logique du sens [10] où le lecteur moderne reconnaît la théorie des systèmes formels dont Husserl exige, avant les théorèmes de Gödel, la complétude syntaxique [11].

La logique formelle est en effet l'étude des lois de la pensée : Husserl considère « la logique formelle en sa qualité de science des significations de pensée dans leur généralité formelle et des objets en général qui leur correspondent, signifiés dans leur généralité formelle. » *Ideen III* (Husserl, 1993, p.95). La logique est donc l'étude de la pensée en tant qu'un quelque chose en général est pensé, et considère donc les lois *formelles* dans la mesure où elle fait abstraction de ce qui est pensé pour s'intéresser exclusivement à la manière de le penser et aux lois caractérisant les manières de le penser.

Il n'est pas dans notre propos de rentrer dans les détails de la logique husserlienne (voir la première partie qui en retrace les principales propriétés pour établir des parallèles avec l'IA). Notons simplement le fait fondamental que la logique formelle husserlienne correspond à une logique du sens et non à une logique du vécu. Or, parler d'un éventuel formalisme computationnel de Husserl reviendrait à vouloir lui faire endosser le projet d'une algorithmique du vécu. Mais un tel projet est explicitement écarté par Husserl : le domaine formel ne peut s'annexer le domaine phénoménologique du vécu.

10. Voir la chapitre 3 pour des précisions sur la logique husserlienne.

11. Un système formel est syntaxiquement complet quand toute formule bien formée close du langage, ou sa négation, est un théorème du système : $\forall F$ on a $\vdash F$ ou $\vdash \neg F$.

En effet, une logique du sens n'implique pas que tout sens soit formel, mais que tout sens est soumis au formel. En particulier, le domaine matériel échappe au formel, et possède des lois que le formel ne peut appréhender. Il y a une science du contenu qui échappe à la science de la forme.

Or, la phénoménologie est une science matérielle, qui possède pour contenu les lois du vécu. Le problème du computationnalisme husserlien est donc celui de la relation au domaine formel qu'entretient la phénoménologie comme science matérielle de la conscience. Husserl se pose explicitement la question et envisage le statut mathématique de la phénoménologie à un double point de vue : le point de vue des mathématiques formelles, c'est-à-dire de la logique formelle, et le point de vue des mathématiques matérielles, qui ne sont pas des mathématiques de la pensée mais des mathématiques d'une région abstraite de la matière, comme la géométrie se définit comme les mathématiques de l'espace.

Pour Husserl, la question ne fait pas de doute : la phénoménologie est d'emblée non formelle, et une mathématisation formelle de la phénoménologie est à « éliminer d'emblée » : la seule mathématisation qu'il faut envisager avec soin est celle qui rapporterait une discipline matérielle à une autre discipline matérielle d'ordre mathématique. Par exemple, s'il est absurde d'envisager une formalisation au sens d'une axiomatiation de la phénoménologie, le formel ne pouvant jamais épuiser dans sa légalité logique la légalité propre au matériel, il faut considérer si le projet d'une mathématisation comme une *géométrisation* de la phénoménologie est possible ou non. Car la géométrie est une science éidétique matérielle et mathématique. Husserl s'exprime ainsi :

> « Partons de la distinction des essences et sciences d'essences en matérielles et formelles. Nous pouvons éliminer les disciplines formelles et avec elles tout l'ensemble des disciplines mathématiques formelles, puisque de toute évidence de la phénoménologique appartient aux sciences éidétiques matérielles. Si d'une façon générale l'analogie peut être un guide dans les questions de méthode, elle prendra toute sa force si nous nous bornons aux disciplines mathématiques matérielles, par exemple à la géométrie ; nous poserons donc la question précise : faut-il ou peut-on constituer une phénoménologie qui serait *une « géométrie » des vécus ? » Ideen I* (Husserl, 1950, p.229)

La question du formalisme husserlien paraît donc définitivement réglée par Husserl lui-même en ce qui concerne notre problème d'une algorithmisation du vécu : s'il y a bien soumission du matériel au formel dans la mesure où toutes les régions matérielles sont soumises aux lois formelles, *i.e.* les lois formelles y sont valables, il n'y a pas épuisement du matériel dans le formel : il y a des lois du matériel qui ne sont pas formelles. Puisque le phénoménologique est matériel, il ne peut donc être formel.

On peut se demander pourquoi tant de commentateurs ont néanmoins soutenu outre atlantique une version computationnelle de la phénoménologie husserlienne. C'est que, selon eux, la distinction formel / matériel ne tient pas : il ne peut y avoir d'opposition qu'entre les sciences formelles qui seraient analytiques *a priori*, et les sciences matérielles qui seraient synthétiques *a posteriori*. En effet, dans la mesure où il ne saurait y avoir de synthétique *a priori*, toute science mathématique, puisqu'elle est *a priori*, est obligatoirement formelle. Par conséquent, pour ces lecteurs, il est tout aussi légitime de s'interroger sur une géométrisation du vécu que sur sa formalisation, puisque toutes deux rapporteraient le phénoménologique à une mathématique formelle.

La lecture anglo-saxonne de la phénoménologie, y projetant son primat de l'analytique sur le synthétique en matière d'*a priori*, légitime par ce moyen la question d'une algorithmique du vécu. Mais, si l'on plaide en faveur d'un primat du synthétique en ce qui concerne les disciplines *a priori*, cette question se trouve encore légitimée. En effet, nous avons argumenté dans le chapitre précédent en faveur d'une conception synthétique de la logique formelle et de l'informatique. Dans ce cas, ce n'est donc pas la géométrie qui serait analytique comme la logique, mais la logique qui serait synthétique comme la géométrie. Il faut donc se demander si la région matérielle phénoménologique peut se rapporter à la région matérielle computationnelle.

A ce niveau, puisque l'on sort du cadre formel / matériel posé fixé par Husserl, il est impossible de s'en tenir aux déclarations de principes comme celle rappelée ci-dessus. Puisqu'il est possible de se poser le problème d'une algorithmique du vécu, il faut considérer si les lois s'appliquant aux vécus sont des lois de nature computationnelle ou non. Il faut donc s'interroger sur la nature du vécu et de sa légalité et envisager leur algorithmisation.

Pour cela, il faut s'engager dans la phénomélogie des régions obscures de la conscience, où le vécu se constitue comme tel. Il s'agit en effet de considérer ce qu'est le vécu, sa nature, sa constitution, alors que jusque là n'était considéré que la constitution d'objets transcendants grâce aux données immanentes que sont les vécus. Il ne s'agit plus d'une phénoménologie de la constitution de la transcendance, mais d'une phénoménologie de la constitution de l'immanence.

> « Au niveau de considération auquel nous nous limitons jusqu'à nouvel ordre, et qui nous dispense de descendre dans les profondeurs obscures de l'ultime conscience qui constitue toute temporalité du vécu, nous acceptons plutôt les vécus tels qu'il s'offrent à la réflexion immanente en tant que processus temporels unitaires ; » *Ideen I* (Husserl, 1950, p.288)

Dans les *Ideen* en effet, la constitution du vécu comme contenu immanent est mise hors circuit : elle est renvoyée aux *Leçons* sur le temps.

> « En ce sens, j'ai, consciemment et expressément, mis hors circuit dans mes *Ideen* les problèmes de la conscience immanente du temps, ou ce qui revient au même, ceux de la constitution de ces objets de la temporalité égologique et j'ai tenté d'esquisser et en partie aussi de mener à bien, de façon suivie, une importante problématique pour des descriptions possibles dans cette sphère. Dans cette sphère intervient alors nécessairement une distinction radicale, celle entre *data hylétiques* et *fonctions intentionnelles*. » *Logique Formelle et Logique Transcendantale* (Husserl, 1957, p. 379-380)

3.6 La constitution temporelle des vécus

Il est temps à présent de considérer l'ouvrage fondamental que constituent les *Leçons sur la conscience intime du temps* pour envisager si à leur niveau[12] les questions laissées en suspens plus haut trouve une réponse. Le vécu et la région phénoménologique qui le comprend se caractérisent par leur exclusion de toute dimension spatiale. En effet, l'espace constitue la transcendance des objets et implique leur perception par esquisses. C'est l'espace qui est le principe de l'autonomie des objets transcendants et de leur exclusion de la conscience :

> « L'esquisse est du vécu. Or le vécu n'est possible que comme vécu et non comme spatial. Ce qui est esquissé n'est possible par principe que comme spatial (il est précisément par essence spatial) et n'est pas possible come vécu ; Il est même particulièrement absurde de prendre l'esquisse de forme (celle par exemple d'un triangle) pour quelque chose de spatial, de possible dans l'espace ; ce faisant, on la confond avec la forme esquissée, c'est-à-dire avec la forme qui apparaît. » *Leçons sur le temps* (Husserl, 1964, p.75-76)

Le vécu est donc par essence temporel, et non spatial. Il possède donc une structure temporelle et ses lois seront celles de la concience du temps. Dans ses *Leçons sur la conscience intime du temps*, Husserl analyse la conscience en trois couches :

- une première couche, celle de l'objet transcendant, *e.g.* la maison que je suis en train de voir ; cet objet se constitue dans le flux des vécus de l'apparition de la maison, par lesquels elle apparaît, qu'Husserl appelle « esquisses » ; alors que l'objet transcendant est spatial, impliquant sa donation par esquisse, le vécu par lequel l'objet apparaît est non spatial, et de nature temporelle ;
- une seconde couche, celle de l'objet immanent, ou vécu ; non spatial, l'objet immanent possède une objectivité et individualité temporelles : c'est sa constitution temporelle qui lui confère son existence en tant qu'objet ; l'objet immanent se constitue dans le flux des impressions ponctuelles ;
- une troisième et dernière couche, celle du flux impressionnel des présents vivants ; non objective, cette couche s'auto-constitue, c'est-à-dire qu'elle n'est pas constituée par l'intermédiaire d'une autre couche à laquelle elle renverrait, mais qu'elle se constitue elle-même comme *un* flux ; totalité en tant qu'elle est une, non objective en tant qu'elle n'est pas constituée.

L'analyse de la conscience du temps révèle que le vécu, objet immanent constituant l'objet transcendant, explicitant la fondation de la transcendance dans l'immanence, est lui-même constitué :

> « Ces contenus immanents sont les vécus au sens habituel : les données sensibles (même si elles ne sont pas remarquées), par exemple un rouge, un bleu, etc.; puis les apparitions (apparition de maison, d'alentours.), qu'elles et leurs « objets » soient observés ou non. Ensuite les « actes » d'énoncer, désirer, vouloir, etc., et les modifications reproductrices qui leur appartiennent (imaginations, souvenirs). Tout cela est contenus de conscience, contenus de la conscience originaire qui constitue des objets temporels et qui n'est pas elle-même à son tour en ce sens contenu, objet dans le temps phénoménologique. » *Leçons sur le temps* (Husserl, 1964, p. 110)

Le problème est donc de comprendre comment les vécus peuvent être des objets, comment ils peuvent être des durées : n'étant pas spatiaux, leur unité ne peut être conférée

12. L'ouvrage central de Husserl sur le temps est *Leçons sur la conscience intime du temps* (Husserl, 1964). *Expérience et Jugement* (Husserl, 1991) contient également des analyses de la temporalité. Ces analyses ont suscitées d'abondants commentaires profonds et éclairants : (Ricoeur, 1985), (Desanti, 1992), (Granel, 1968), (Dastur, 1995).

par l'espace ; ils ne peuvent avoir une extension ni une individualité spatiale. Leur unité leur est conférée par le temps : ce sont des objets purement temporels ; un objet temporel est une durée.

Comment se constitue donc une durée dans la conscience : comment avoir une objectivité dans un temps qui par essence est évanescent, qui oublie tout ce qu'il a vécu en ignorant encore ce qu'il va vivre. Le temps comme forme de la succession ne peut constituer d'objectivité car il ne peut rassembler en une unité ce qui est éparpillé dans la succession. Nous avons comment, chez Kant, l'espace comme forme de la juxtaposition unifie dans une même présence ce qui se succède dans la conscience : les mathématiques permettent de synthétiser ce qui est juxtaposé dans l'espace.

Le problème est donc ici de savoir rassembler en une unité ce que la succession temporelle éparpille, *sans faire appel à un principe de juxtaposition spatiale*. Il faut un principe de synthèse temporelle respectant la forme du temps comme succession.

Il faut tout d'abord se prémunir des fausses solutions. Ainsi la solution brentanienne, qui voit dans l'imagination le principe de la synthèse temporelle : une durée se constitue parce que l'imagination reproduit les instants passés au moment présent, si bien qu'à chaque instant, on a à la fois l'instant courant plus les instants reproduits par l'imagination. Ces instants sont donc donnés ensemble à un même instant, échappant ainsi à la dispersion de la succession. Mais, comme le remarque Husserl, échapper à la dispersion de la succession n'est pas constituer une unité : il manque en effet un principe pour la synthèse des éléments ponctuels juxtaposés. En d'autres termes, il n'y a pas de conscience d'*une* durée, mais de plusieurs instants simultanés. Car le problème n'est pas que la succession soit donnée à un instant donné, ce que fait la reproduction imaginative, mais que la conscience de la succession soit donnée. La solution brentanienne ne résout rien faute d'avoir compris ce point.

Le coup de génie de Husserl est d'avoir compris que, pour avoir une authentique synthèse temporelle, et non un avatar de synthèse spatiale, il fallait non pas rendre présent le passé, mais rendre présent *la conscience du passé comme passé*. Le passé doit toujours resté passé : il ne doit pas être rendu simultané au présent, pour dépendre ensuite d'un principe de synthèse spatial qui de toute façon ne pourrait jamais rendre compte du caractère temporel de la durée, à savoir que le début est passé, inexorablement passé, de manière irreversible. En effet, dans une synthèse spatiale, l'objet constitué, résultant de la synthèse, ne possède pas de structure temporelle. Or, par définition, une durée, si !

Comment rendre présente la conscience du passé, sans rendre présent le passé ? Comment rendre présente la conscience du passé pour conserver à l'objet temporel constitué la structure temporelle qui par définition le caractérise ?

La réponse tient en deux mots : rétention et protention. Chaque instant vécu par la conscience possède une rétention du passé et une protention de l'avenir. La rétention est une structure intentionnelle par laquelle le passé est visé intentionnellement par le présent. Présent intentionnellement en tant que visé, le passé reste temporellement toujours passé. Le passé est présent en conscience, et seulement pour la conscience. Car il est bien passé, et pour toujours.

> « La rétention n'est pas une modification en qui les données impressionnelles demeureraient contenues réellement, en changeant seulement de forme : c'est au contraire une

intentionnalité, une intentionnalité de type spécifique. » *Leçons sur le temps* (Husserl, 1964, p.159)

Si la rétention est ainsi attachée au présent, c'est qu'elle est une modification de l'intuition du présent vivant : dès qu'il y a eu intuition d'un présent, cette intuition présente, actuelle, se modifie en une conscience actuelle de « tout juste passé ». Il n'y a donc pas reproduction du passé dans l'actualité du présent, mais permanence dans la conscience de ce qui a été donné il y a juste un instant dans l'actualité du présent sous la forme du « tout juste passé ».

Cette modification rétentionnelle est récursive : chaque présent étant rétentionnellement retenu dans les présents suivant, un présent donné retient les présents juste passés avec leurs propres rétentions. A un présent, il faut donc associer la rétention du « tout juste passé », mais aussi la rétentions des rétentions des présents « tout juste passés », etc.

La rétention et la protention, pour lesquelles les analyses sont identiques, avec évidemment, une conscience matériellement vide en ce qui concerne la protention, constituent le vécu comme durée. En effet, à un instant donné, il faut donc associer la rétention des instants passés et leurs propres rétentions. Par conséquent, à cet instant, la conscience vise intentionnellement tout un ensemble d'instants passés structurés entre eux par des liens rétentionnels. Si l'on reprend le célébrissime diagramme du temps que propose Husserl (voir la figure 6.1), si l'objet immanent correspond à la suite des instants retenus en tant qu'ils se rattachent à un même instant, il se schématise sur le diagramme en la ligne verticale associée à un instant présent (ligne S2 – O'), chaque point de cette ligne verticale indiquant la rétention d'une impression ponctuelle passée (lignes O'– O) :

« Chaque phase, en étant conscience rétentionnelle de celle qui la précède, enferme en elle-même toute la suite des rétentions écoulées en une chaîne d'intentions médiates : c'est précisément ainsi que se constituent les unités de durée qui sont rendues par les lignes verticales du diagramme du temps, et qui sont les objets des actes de rétrospection. » *Leçons sur le temps* (Husserl, 1964, p. 159)

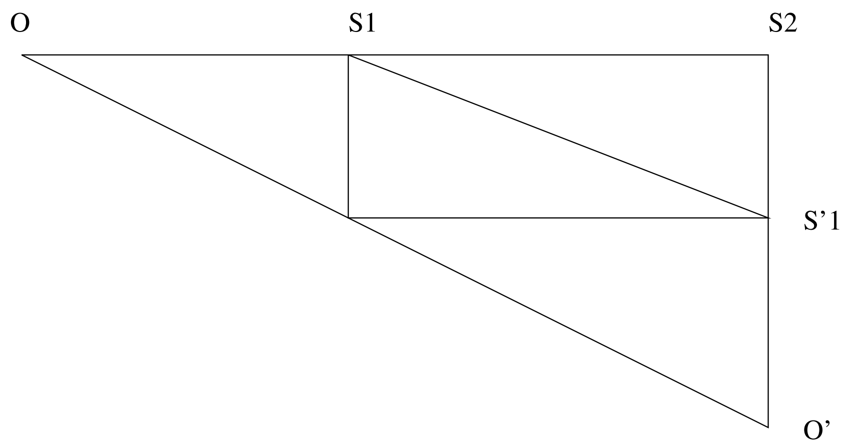


FIG. 6.1 –*Diagramme du temps*

L'individualité de l'objet immanent procède donc de la synthèse temporelle assurée par la rétention. Mais cette synthèse n'est que formelle, et elle doit se compléter d'une autre synthèse, celle du contenu. La synthèse du contenu est l'identité : des contenus sont fusionnés s'ils sont identiques. S'ils sont différents, ils ne peuvent être synthétisés matériellement, il faut faire appel à un principe formel permettant de les rendre commensurables l'un à l'autre et donc synthétisables. Ces principes sont bien sûr l'espace, non considéré ici par principe, et le temps.

> « La suite des appréhensions de chose ne se recouvre pas dans la seule mesure où elle contribue à se constituer une succession continue, mais aussi dans la mesure où elle constitue la même chose. Le premier recouvrement est celui de la similitude par essence, qui sert de lien ; le second, un recouvrement de l'identité, parce que, dans l'identification continue de la succession, nous avons conscience de quelque chose d'identique qui dure. » *Leçons sur le temps* (Husserl, 1964, p. 121)

Et encore :

> « La parfaite similitude, indifférenciable, de la matière du temps [principe d'identité pour la synthèse matérielle des contenus], et la continuité de la modification de la conscience qui pose celui-ci [principe rétentionnel et protentionnel pour la synthèse formelle de contenus distincts par leur position temporelle respective], sont par essence les fondements de cette unité par fusion de l'extension sans rupture du *do*, et c'est seulement ainsi que prend naissance une unité concrète. » *Leçons sur le temps* (Husserl, 1964)

Protention et rétention sont les principes intentionnels donnant au présent son extension temporelle sans confondre cette dernière avec une juxtaposition spatiale. La solution husserlienne permet de conserver à l'objet temporel sa structure temporelle, contrairement à la solution brentanienne. L'objet constitué temporellement par les rétentions possède un ordre temporel qui est reflété par l'emboîtement des rétentions.

Mais il faut lever ici une ambiguïté : la nature de l'intentionnalité de la rétention. Il faut en effet distinguer dans la rétention si le vécu est bien ce qui est retenu, objectivité de niveau deux, constituée, ou s'il s'agit d'une impression ponctuelle du présent vivant qui vient « tout juste de passer ». Il faut distinguer une double intentionnalité dans la rétention :

> « Chaque dégradé de conscience du type « rétention » possède une double intentionnalité : la première, celle qui sert à la constitution de l'objet immanent, du son ; c'est celle que nous nommons « souvenir primaire » du son (tout juste senti) ou, plus clairement, que nous nommons précisément rétention du son. L'autre est l'intentionnalité constitutive de l'unité de ce souvenir primaire dans le flux ; c'est-à-dire que la rétention, par cela même qu'elle est conscience d'« encore », conscience qui retient, rétention précisément, est du même coup rétention de la rétention écoulée du son : elle est, dans son dégradé continu dans le flux, rétention continue des phases qui ont précédé continûment. » p. 106

C'est-à-dire, soit la rétention retient le présent tout juste passé et seulement lui, en une espèce de rétention primitive, soit la rétention retient, par l'intermédiaire de la rétention primitive, tout le cortège des rétentions de chaque instant passé, accédant ainsi à la continuité des instants correspondant à la durée à constituer. Il y a donc la rétention primitive, la rétention du tout juste passé, et la rétention de rétention qui, accédant aux instants constitutifs de la durée en *une seule visée* rétentionnelle, vise intentionnellement l'objet immanent complet. Il faut insister sur le fait que la rétention de rétention permet d'accéder, en une seule visée de la conscience, à l'objet immanent. C'est en effet la conscience une qui confère à l'objet son unité, son individualité, son objectivité. Les choses se passent comme si, lors d'une rétention primitive, on s'aperçoit qu'il y a quelque chose de plus de donné avec le « tout juste passé ». En *plongeant* le regard sur ce « tout juste passé », on trouve que se profile derrière lui tout un cortège d'instants passés reliés entre eux par cette relation d'ordre qu'est le « tout juste passé ». C'est donc ainsi que l'objectivité temporelle se constitue sur le fond d'une continuité intentionnelle de rétentions. Il faut par conséquent opposer une intentionnalité rétentionnelle dont le corrélat est le vécu constitué appartenant au temps immanent à une intentionnalité rétentionnelle dont le corrélat n'est pas une unité intentionnelle constituée mais les instants ponctuels composant le flux originaire.

> « Nous avons dans le courant de la conscience, une double intentionnalité. Ou bien nous considérons le contenu du flux avec sa forme de flux. Nous considérons alors la suite des vécus originaires, qui est une suite de vécus intentionnels, conscience de... Ou bien nous tournons le regard vers les unités intentionnelles, vers ce dont on a, dans l'écoulement du flux, intentionnellement conscience comme possédant une unité : alors se tient là pour nous dans le temps objectif une objectivité, le champ temporel proprement dit par opposition au champ temporel du courant des vécus. » *Leçons sur le temps* (Husserl, 1964, p. 156)

Les *Leçons* montrent que les vécus immanents, substrats des fonctions intentionnelles décrites dans les *Ideen* sont des unités intentionnelles constituées. Cette constitution restait ininterrogée dans ce dernier ouvrage dans la mesure où il s'agissait d'élucider la constitution de la transcendance dans l'immanence, la transcendance objective du monde dans l'immanence subjective du vécu. En d'autres termes, les *Ideen* se tiennent entre les niveaux deux et trois que nous avons évoqués plus haut, tandis que les *Leçons* se tiennent entre les niveaux un et deux, dévoilant la constitution du vécu.

Or, ce que révèle cette constitution est que la structure du vécu n'est pas compositionnelle, syntaxique, mais différentielle. Les vécus originaires ne se composent pas selon un formalisme syntaxique imposé par la structure temporelle. Ce que montre Husserl, c'est que le temps est bien une structure formelle, mais elle est formelle au sens de Kant et non au sens du formalisme logique. En d'autres termes, le temps constitue un cadre formel au sein duquel les impressions originaires ponctuelles *s'intègrent* en unités immanentes : ce cadre formel est un continuum d'impressions ponctuelles qui s'intègrent à la manière du calcul différentiel pour constituer les vécus concrets, et non un langage formel dont les structures syntaxiques s'instancient dans les vécus concrets. Les analyses de la structure temporelle prouvent ainsi que la constitution du vécu à partir du flux originaire n'est pas à appréhender sur le modèle des langages formels selon le rapport type/instance, mais à considérer selon le modèle mathématique du calcul différentiel selon le rapport continu/intégration.

Mais si la conception du temps comme structure d'intégration permet d'échapper à la critique formaliste l'assimilant au fonctionnalisme computationnel, il n'en demeure pas moins que cette solution est tributaire d'une conception continue du temps comme totalité infinie en acte. Ce problème est celui bien connu de l'autoconstitution du flux des impressions originaires comme flux unitaire, comme totalité une et actuelle.

C'est sans doute là que gît une faiblesse potentielle de la solution husserlienne et c'est donc à ce problème qu'il faut nous atteler à présent.

4 Critique de la temporalité du vécu

En substance, notre critique consistera à montrer que la solution husserlienne ne parvient pas à penser la temporalité en tant que telle, mais à partir de son autre qu'est l'espace. Cela nous suggère, non pas qu'il faut dépasser la solution husserlienne pour une posture philosophique abordant le temps authentique, dans sa vérité, tel qu'il est en lui-même, sans le concevoir depuis le modèle de l'espace, mais qu'au contraire une telle posture n'existe pas. Temps et espace sont indissociables et ne peuvent être pensés indépendamment l'un de l'autre. Quand ils sont pensés pour eux-mêmes, ils sont en fait abstraits en des principes formels de juxtaposition pour l'espace et de succession pour le temps. Mais dès que l'un des principes est mis en œuvre, il fait appel à l'autre pour déployer son contenu. C'est ainsi que le principe temporel de succession ne peut se concrétiser que dans le cadre d'une totalité déjà constituée d'entités coprésentes parce que co-données pour l'application de la succession. Sinon, le temps comme succession doit faire place au temps comme création, où chaque instant est une unité incommensurable aux autres et sans relation avec elles. Unité absolue et indépendante, l'instant est créé constamment indépendamment de ce qui le suit et ce qui le précède. Mais dans ce cas, il devient impossible de retrouver une quelconque régularité caractérisant le temps, ce dernier n'étant plus une forme de donation imposant ces lois au donné.

Par ailleurs, le principe spatial de juxtaposition mobilise un principe de succession dans la mesure où l'espace doit être parcouru dans un certain ordre de succession si l'on veut y reconnaître des structures spatiales. En d'autres termes, le temps est un parcours qui comme tel mobilise un principe spatial, le lieu du parcours, et l'espace est une construction qui comme telle mobilise un principe temporel d'ordre de construction.

L'analyse temporelle du temps chez Husserl est d'autant plus intéressante qu'elle n'assimile pas d'emblée le temps à l'espace, comme nous avons pu le voir plus haut. En particulier, la synthèse du vécu à partir des impressions originaires ne repose pas sur une coprésence, une juxtaposition de ces impressions, mais sur la chaîne continue des rétentions.

En cela, Husserl dépasse Kant, et donne le principe d'une psychologie des vécus qui ne soit pas subordonnée à l'espace. Voyons plus avant ce qui oppose Husserl à Kant.

4.1 Kant et Husserl

Considérons en effet la déduction transcendantale subjective. Nous avons rappelé qu'elle se compose de trois étapes : la synthèse de l'appréhension, la synthèse de la reproduction, la synthèse de la recognition dans le concept. Ces trois synthèses s'impliquent dans un ordre régressif dans la mesure où l'objet construit est compris comme une coprésence dans l'espace des différents moments de l'objet. Cela signifie qu'il faut rassembler tous les moments perçus en une même présence pour y reconnaître celle d'un objet. Il faut donc une synthèse de l'appréhension qui appréhende à la fois, simultanément tout le divers de la sensation correspondant à l'objet. Mais ce divers est successivement donné à la conscience selon la forme *a priori* du temps. Par conséquent, il faut que la succession de la sensation dans le temps devienne une juxtaposition dans l'espace :

> « Cette synthèse consiste à poser la chose ou une coprésence dans l'espace à partir d'une succession dans le temps. La moindre perception étant l'addition d'une série de perceptions plus petites, *la constitution de la chose n'est possible que par celle du présent*. Dans la continuité du temps, l'esprit rassemble plusieurs instants pour constituer la présence à la conscience. La synthèse réunit le divers du contenu et le divers du temps – le divers du contenu ne peut être rassemblé qu'en rassemblant le divers du temps –, pour en faire une unité qui donnera le présent. » (Rivelaygue, 1992, p.112)

Pour rendre coprésent ce qui est successif, il faut donc une seconde synthèse qui *reproduise* ce qui est écoulé pour le rendre co-présent à ce qui est perçu dans l'instant présent. C'est la synthèse de la reproduction. Mais, cette reproduction des instants passés ne peut s'effectuer au petit bonheur : si ce qui est reproduit n'obéit pas à des règles, ce qui se retrouve coprésent au terme de la synthèse de reproduction peut varier, un même divers de sensation pourra donner lieu à différents objets ; ce qui est un objet à un moment donné ne le sera plus à un autre, ce qui est un objet pour l'un ne le sera pas pour l'autre. Par conséquent, pour qu'il y ait véritablement réalité objective, il est nécessaire que la reproduction s'effectue selon des règles, c'est-à-dire selon un concept : c'est la synthèse de la recognition dans le concept. Le concept donne la règle permettant de rendre coprésents les moments successivement perçus correspondant à la réalité de l'objet.

La faculté permettant la synthèse de la reproduction est l'imagination : l'imagination rend présent ce qui est passé. On est donc fort proche de la solution de Brentano si vivement critiquée par Husserl. En effet, pour Husserl, le passé n'est pas *reproduit*, mais *retenu*. Le passé n'est pas présent parce qu'il est reproduit, étant alors effectivement présent au même titre que la perception courante, mais il est présent par son absence, en tant qu'il n'est plus là, ainsi qu'en témoigne la rétention qui le vise.

Par conséquent, il est clair que Husserl ne spatialise pas le temps au sens où Kant l'effectue puisque la synthèse de l'appréhension ne repose pas sur une reproduction chez Husserl, au sens où le successif n'est pas compris comme une juxtaposition.

C'est pourquoi la non spatialité du vécu ne possède pas les mêmes conséquences dirimantes chez Husserl que chez Kant : alors que pour ce dernier la non spatialité du temps interdit la synthèse des phénomènes purement temporels que sont les phénomènes psychologiques [13], Husserl pose grâce à la rétention un procédé synthétique non spatial permettant d'appréhender un matériau immanent lui-aussi défini comme non spatial.

C'est ainsi que la constitution du vécu à partir des instants ponctuels ne lui conférera pas une *réalité objective* au sens kantien, puisqu'il ne saurait être spatial, mais une *essence*, au sens husserlien. Le vécu husserlien est bien une entité constituée, tout comme l'objet kantien. Mais l'originalité de Husserl est de proposer une constitution lui conférant une objectivité essentielle et non spatiale, confirmant de ce fait l'originalité (ainsi que son originarité) de la région conscience eu égard à la région spatiale mondaine. Husserl peut envisager une science de l'esprit parce qu'il reconnaît un type de science ignoré ou délaissé par Kant : une science descriptive d'essence. Le prix à payer est de l'opposer aux

13. Voir le chapitre précédent, section 5-2.

sciences de la nature puisque ces dernières négligent l'essence pour la construction spatiale d'objet, introduisant ainsi un dualisme épistémologique entre le monde et l'esprit. Le vécu se construit à partir d'un glissement continu du présent vers le passé et le vécu est cette mémorisation du passé dans le présent *via* la rétention. L'unicité de l'instant contient l'unité du vécu grâce à la structure rétentionnelle qui inscrit dans son actualité le souvenir primaire des instants tout juste passés. La synthèse kantienne fonde l'unité de l'objet sur la reproduction des différents moments de la perception sensible pour les inscrire dans un simultanéité spatiale : l'unité de l'objet procède de la pluralité des moments successifs reproduits dans la présence. La synthèse husserlienne, en revanche, fonde l'unité de l'objet sur la rétention dans le présent ponctuel du présent vivant des différents moments constituant l'objet : l'unité de l'objet procède de l'unicité de l'instant ponctuel qui contient en lui le cortège de rétentions constituant l'objectivité temporelle de l'objet. Comme le souligne Rivelaygue :

> « Il faudra attendre Husserl pour tirer les conclusions de ce qui est impliqué dans ce texte : les phénomènes intérieurs, du fait de leur appréhension dans le temps seul, n'ont pas une nature mais une essence; ... puisque dans le sens interne on ne perçoit pas successivement la même chose, ou des faces diverses de la même chose, mais que la succession implique alors l'altérité de ce qui précède par rapport à ce qui succède. Le concept d'un phénomène du sens interne doit pouvoir être tiré d'une unique apparition de ce phénomène, puisque sa répétition ne rendrait pas davantage possible une synthèse: ce concept est donc trouvé par abstraction et non par construction, par intuition d'essence et non par synthèse. » in (Rivelaygue, 1992, p.314)

Il peut donc sembler à première vue que Husserl parvient, dans les *Leçons sur le temps* à élaborer une authentique phénoménologie de la conscience du temps, authentique car ne s'appuyant pas sur une conceptualité héritée de l'espace.

4.2 La finitude rétentionnelle

Néanmoins la constitution du vécu mobilise un principe spatial dans la mesure où elle repose sur la notion du flux unitaire impressionnel de la conscience : la constitution du vécu ne se comprend que si le flux se constitue lui-même comme unité, dans laquelle les impressions sont en quelque sorte présentes ensemble pour constituer des structures rétentionnelles. Cela implique qu'il y a une totalité actuelle comme présence ensemble, et que si Husserl a pu échapper à l'interdit kantien de l'espace, c'est qu'il l'a repoussé à un niveau ultime, celui de l'autoconstitution de la conscience comme unité et totalité.

En effet, Husserl s'interroge en dernier lieu sur le matériau permettant la constitution du vécu : le flux des impressions originaires. Comment se dernier est-il constitué ? En tant que structure unitaire, car il faut qu'il soit un pour que les instants aient quelque chose à faire ensemble et que la rétention puisse retenir des instants passés qui aient quelque chose en commun avec l'instant présent, notamment le fait d'appartenir au même flux, il est nécessairement constitué. Mais comment est-il constitué ? Faut-il envisager une régression à l'infini de constitution ? Cela n'est pas raisonnable. Il faut donc poser que le flux impressionnel *s'autoconstitue* en une unité qui est ici totalité puisqu'il s'agit de l'unité d'une multiplicité.

Cette totalité est comprise comme un infini actuel au sein duquel le divers est simultané sinon temporellement du moins logiquement. Les instants sont tous présents en droit, c'est ce que signifie l'autoconstitution, pour que la rétention puisse y installer son intentionnalité.

C'est ce qui apparaît avec précision avec la thèse sous-jacente aux analyses de Husserl de l'infinité de la rétention. Les rétentions et protentions sont de fait finies mais *en droit infinies*. Nous avons vu, avec la double intentionnalité de la rétention, que la rétention peut viser, dans une intentionnalité dite horizontale, le flux des impressions lui-même, dans sa totalité. A chaque instant, la totalité du flux est donnée par cette intentionnalité. La synthèse rétentionnelle devient en fait une synthèse définie sur la coprésence des instants impressionnels dans l'unité d'un flux unitaire : la structure rétentionnelle du vécu s'enlève en quelque sorte sur un fond continu dont le principe de continuité est la rétention. La rétention devient le nom temporel de la contiguïté spatiale, si l'on comprend par ce terme un principe d'infini actuel donnant de manière unitaire et totalitaire une multiplicité.

Si Husserl rompt avec la synthèse de la reproduction pour la constitution du vécu, c'est qu'en fait il mobilise ailleurs un principe de coprésence. Or, une conséquence est inéluctable dans l'approche husserlienne : le flux comme totalité, c'est-à-dire l'infinité de la rétention est nécessaire dans l'analyse que propose les *Leçons*. En effet, si la rétention n'est pas infinie, si elle est finie, cela signifie que le vécu peut être entièrement déconnecté des autres vécus : ces vécus possèdent des finitudes rétentionnelles et protentionnelle telles qu'ils ne sont pas reliés aux impressions constituant le premier vécu. Si cela est le cas, cela implique qu'il n'y a plus d'unité de la conscience. La conscience est éclatée, ce n'est plus qu'une rhapsodie d'impressions et de vécus constitués localement, mais sans unité globale.

L'infinité en droit de la rétention est donc nécessaire pour qu'il puisse y avoir auto-constitution et unité du flux de la conscience. Sans une telle structure rétentionnelle, on assiste à une dissolution de la conscience qui se retrouve atomisée en impressions ponctuelles sans aucune solidarité les unes avec les autres. Or, il faut bien dégager un *mode d'être ensemble* spécifique aux impressions ponctuelles qui puisse justifier qu'on les rassemble en un même ensemble dont la structure est celle du flux évanouissant entrelacé de liens rétentionnels.

Or, on vient de le voir, Husserl adopte comme mode d'être ensemble celui de l'espace géométrique où chaque point, indépendant des autres en lui-même, se retrouve solidaire des autres de par sa position dans l'espace. Cette solidarité spatiale s'exprime par la structure différentielle dont se retrouve doté l'espace géométrique. Parallèlement, l'espace des impressions ponctuelles (auto-)constituant la conscience possède une structure rétentionnelle qui est donc à la temporalité ce que la structure différentielle est à la spatialité.

Il faut donc choisir entre la disparition de la conscience et le fait de penser sa temporalité selon une conception spatiale. Devant une telle alternative, Husserl ne peut hésiter : refusant le sacrifice de la conscience, Husserl opte pour cette auto-constitution du flux impressionnel qui correspond, sans doute, au véritable positivisme de Husserl.

En déclarant le flux impressionnel auto-constitué, Husserl reconnaît qu'il existe en toute indépendance, pour lui-même, puisque, auto-constitué, il n'est constitué par personne et donc pour personne, pour personne d'autre que lui-même. Le flux impressionnel est donc une substance, un fait positif donc l'existence s'impose de soi et non pour quelqu'un. L'auto-constitution du flux en fait par conséquent une *chose* : la conscience est une chose qui pense car elle est un flux auto-constitué dans lequel se constituent les pensées.

Il y a donc à ce niveau le passage d'une phénoménologie de la conscience à une ontologie de la conscience : ce passage est imposé par la menace d'une disparition phénoménologique de la conscience dans la mesure où, en l'absence d'un flux unitaire que nulle nécessité phénoménologique ne justifie, la conscience comme donnée phénoménologique disparaît. L'irruption de l'ontologie est donc commandée par cette démarche transcendantale qui consiste à poser (ontologie) le flux auto-constitué à partir du fait phénoménologique de la conscience. L'auto-constitution est donc le prérequis ontologique transcendantalement déduit de la conscience comme fait phénoménologique.

Mais, en fait la dérive ontologique et positiviste de l'auto-constitution est la conséquence fort logique de la notion d'impression ponctuelle. Le fait même de comprendre la conscience comme étant constituée à partir de points, points d'actualité et de présence, la fait procéder du fait positif qu'est alors ce point, préjugé sur la nature de la conscience. En voulant considérer la conscience pour elle-même, Husserl importe des considérations ponctuelles qui reflètent un préjugé positiviste : il y est contraint dans la mesure où il est impossible de considérer la conscience pour elle-même puisque par essence, nous l'avons vu plus haut, elle est pour l'altérité, ce qui n'est pas elle. Découvrant une intimité qui ne peut être que vide, Husserl se voit donc obligé de la meubler de ses préjugés positivistes, à savoir sa compréhension mathématique de l'espace comme solidarité différentielle d'un ensemble de points.

Par conséquent, si Husserl reconnaît bien que la rétention est en fait finie [14], ses analyses impliquent néanmoins qu'elle est en principe infinie pour assurer l'unité de la conscience. Mais si la manière d'obtenir cette unité s'avère discutable, introduisant d'une part un préjugé spatial dans la description de la conscience et d'autre part un positivisme de l'impression ponctuelle, il en résulte que la nécessité d'affirmer l'infinitude rétentionnelle n'est plus légitime et qu'il faut rendre par conséquent justice au fait que la rétention est finie. Mais si l'on affirme la finitude rétentionnelle [15] il faut alors rendre compte de l'unité de la conscience.

4.3 Dépasser les impressions ponctuelles

Mais si l'unité ne peut se faire via l'unité globale des impressions en flux, et qu'il y a néanmoins unité, c'est que l'unité procède d'autre chose que du flux autoconstitué comme un.

La notion d'impression ponctuelle procède d'une conception du temps comme présence ponctuelle, instantanée, unique et indépendante dans sa ponctualité. L'auto-constitution de la collection d'impressions en flux unitaire implique que tout ce qui découle du flux comme structures temporelles du vécu interne à la conscience se construit à partir de la pure présence actuelle de l'impression ponctuelle. N'est présent à la conscience que ce qui est vécu, l'a été, ou va l'être. Tout sens constitué par la conscience doit par conséquent provenir d'une présence dans l'intimité même de la conscience. Autrement dit, il est im-

14. Notamment, à la fin du paragraphe 11 des *Leçons* : « Le champ temporel originaire est manifestement limité, exactement comme la perception ».

15. La finitude rétentionnelle est un thème phénoménologique qui a été introduit par J. Derrida dans son introduction à *L'origine de la géométrie* de Husserl (Husserl, 1968) et qui a été repris par la suite par B. Stiegler (Stiegler, 1996). Nous reprenons ce thème dans le chapitre suivant à propos de la question de la mémorisation pour la constitution du sens.

possible pour la conscience de constituer un sens qu'elle n'aurait pas vécu en son sein par le moyen d'une présence révélée et distillée point par point en impressions ponctuelles. Conséquences fondamentales qui découlent toutes de la conception de conscience comme flux ponctuel.

Notre discussion précédente a montré que la notion d'impression ponctuelle relève d'un préjugé positiviste : l'impression est comprise comme un fait positif. Mais, et c'est là que notre analyse de ce que doit être une « bonne » phénoménologie (cf. 6-1) intervient, de même que notre discussion sur les contenus hylétiques (cf. 6-3.4) : un fait positif ne peut être une donnée phénoménologique ; en particulier il ne peut posséder d'évidence phénoménologique. L'impression ponctuelle n'est pas une donnée phénoménologique, mais un fait positif.

Mais si l'impression ponctuelle est un fait positif, donc sans valeur phénoménologique, il n'en demeure pas moins que la conscience, le corps propre, la perception sensible, etc., sont d'authentiques données phénoménologiques comme en témoignent avec brio les analyses phénoménologiques husserliennes. Si la conscience est une donnée phénoménologique, c'est-à-dire qu'elle existe comme telle, il faut donc qu'il y ait une unité de la conscience sans les impressions ponctuelles qui nous servent de liant grâce à la rétention.

Ce que nous a appris nos analyses phénoménologiques précédentes est qu'il n'existe pas d'intériorité ou d'intimité de la conscience. C'est donc que l'unité ne peut provenir que de l'extériorité de la conscience : il existe dans le monde connu par l'esprit un principe d'unité qui relie et solidarise ce qui semble être ses matériaux de constitution.

Quel peut bien être ce principe ? Ce principe caractérise d'une part *l'extériorité de* la conscience, et d'autre part *l'extériorité pour* la conscience. En effet, la conscience est par essence une constante projection vers une extériorité qu'elle reconnaît par le fait que cette extériorité possède un sens qui s'impose à la conscience. L'extériorité, c'est le fait qu'il existe un sens autonome pour la conscience, autrement dit des objets, qui lui sont donnés. Les objets s'imposent à elle dans le fait qu'elle se voit obligée de les constituer selon des modalités sur lesquelles elle ne peut imposer son arbitraire : leur objectivité réside dans le fait qu'ils sont ce qu'ils sont indépendamment de la conscience. Mais, du fait même qu'ils lui sont donnés, et que la conscience n'y accède que par la manière dont ils lui sont donnés, les objets sont *pour* la conscience, et même *de* la conscience. Ils possèdent un sens que seule la conscience peut leur concéder : ils sont investis d'esprit dans la mesure où leur seule manière d'être un objet est de parler à la conscience pour qu'elle les constitue tels qu'ils sont.

Ces objets investis d'esprit sont donc des objets (extériorité *de* la conscience) et sont investis d'esprit (extériorité *pour* la conscience). En d'autres termes, ce sont des signes : ils n'existent qu'en tant qu'ils parlent à la conscience, c'est-à-dire qu'en tant qu'ils ont un sens, et qu'en tant qu'ils ne sont pas dans la conscience. Le signe est un objet dont l'essence est d'avoir du sens : son essence réside dans ce qu'il n'est pas, dans ce à quoi il renvoie, et non dans ce qu'il est. Transcendant la conscience par son objectivité, le signe lui est homogène parce qu'il lui parle.

Le signe serait donc un principe *sui generis* du sens, une dimension originaire du sens dans la mesure où seul le signe permet de rendre compte du fait que si la conscience ne trouve pas en elle de quoi constituer le sens, c'est qu'elle le trouve hors d'elle. Le signe doit donc être le principe synthétique permettant à la conscience d'unir en une même signification les différents moments qui la composent. Mais plusieurs questions restent néanmoins en suspens :

Une entité extérieur permettant de constituer le sens est nécessairement un signe dans la mesure où, puisqu'on ne peut présupposer de structures internes à la conscience prescrivant à la consitution du sens, la conscience doit pouvoir « lire » sur ce qu'elle rencontre les prescriptions de la constitution ; autrement dit, la conscience lit sur le signe comment le lire comme signe.

Il nous reste par conséquent à décrire comment peut se concevoir une telle constitution du sens à partir des signes, c'est-à-dire une constitution « sémiotique ».

5 Conclusion

On comprend l'enjeu d'une telle question pour notre problématique. En effet, nous avons vu dans le chapitre précédent que l'informatique constitue, en tant qu'elle est une science expérimentale, une objectivité scientifique autonome, celle du calcul. L'informatique est une physique du signe ininterprété, en tant qu'elle l'enrôle dans des processus effectifs temporels. Nous venons de voir que, dans des modalités qu'il nous faut maintenant aborder, le signe interprété est constitutif des connaissances qu'une conscience peut élaborer. En d'autres termes, la conscience doit s'appuyer sur une extériorité sémiotique pour constituer le sens et ses connaissances.

La question qui se pose à présent est donc de savoir si le support informatique, mécanique symbolique de signes ininterprétés, peut s'investir de sens pour prescrire la constitution de connaissance. Autrement dit, le signe ininterprété de l'informatique, le calcul sans signification de l'informatique, peut-il s'investir de sens pour devenir le signe sémiotique rencontré au terme de ce chapitre et participer ainsi à la constitution de nouvelles connaissances ?

Si la réponse est positive, on obtient une caractérisation philosophique du constat que nous avions effectué dans l'introduction, à savoir que l'informatique permet de constituer des connaissances dans un domaine alors que ce dernier est de nature non computationnel, c'est-à-dire que les lois du sens dans ce domaine ne se formalisent pas de manière computationnelle. Si la réponse est positive, il devient possible d'assigner à l'IA l'objectif de formuler les règles selon lesquelles le signe ininterprété de l'informatique, le signe dynamique enrôlé dans le calcul, peut devenir un signe sémiotique prescrivant à la conscience la constitution du sens.

Chapitre VII

Sens et herméneutique matérielle

Nous avons mobilisé la phénoménologie husserlienne comme théorie de la connaissance pour étudier le problème de savoir si la constitution des connaissances pouvait reposer sur une médiation technique. Nous sommes arrivé à la conclusion que la constitution phénoménologique des connaissances repose sur une médiation sémiotique.

En voulant être une théorie de la connaissance, la phénoménologie doit se confronter au problème de l'intentionnalité selon lequel le sens caractérise les entités non par ce qu'elles sont mais par ce qu'elles signifient, par ce qu'elles veulent dire. Il faut expliquer comment des entités peuvent se voir attribuer un sens qu'elles ne possèdent pas par elle-même mais seulement via une conscience qui les anime d'une intention. La phénoménologie husserlienne aborde ce délicat problème mais bute sur le fait qu'elle se voit obligée de poser une positivité fondamentale au sein de la conscience, à savoir un flux d'impressions ponctuelles auto-constitué. La question du sens se ramène sur une question ontologique où le temps compris comme intériorité de la conscience consiste en un présent vivant ponctuellement renouvelé.

Mais si l'on veut éviter les difficultés que posent de telles conséquences ontologiques, on est amené à constater que l'unité phénoménologique de la conscience que l'auto-constitution de la concience devait permettre d'expliquer ne peut procéder d'une intériorité de la conscience mais bien plutôt de son extériorité sémiotique. Rencontrant des objets investis d'esprit, c'est-à-dire d'un sens qu'elle leur attribue, la conscience constitue le sens de ces objets selon des procédures qui s'imposent à elle si bien que le sens constitué est bien plus le sens imposé par l'objet à la conscience que le sens conféré arbitrairement à l'objet par la conscience.

Il en ressort que ces objets rencontrés sont des signes et donc qu'une phénoménologie du sens est une sémiotique. Cette sémiotique doit être phénoménologique dans la mesure où le signe n'est pas une entité positive, mais un objet investi d'esprit prescrivant son sens à la conscience qui le constitue. La phénoménologie du sens comme sémiotique phénoménologique doit donc être comprise sur le modèle de la lecture où le sens est reconnu et déchiffré plutôt qu'il n'est librement attribué à un véhicule matériel indifférent à sa signification. Ce qu'a permis de montrer notre discussion des contenus primaires de la phénoménologie husserlienne, c'est que toute matière devait être comprise en phénoménologie non comme une positivité indépendante mais à partir du sens qu'il anime et dont elle commande l'animation.

L'objectif de ce chapitre est d'expliciter ce projet d'une phénoménologie sémiotique en précisant comment peut intervenir l'extériorité sémiotique dans la constitution des connaissances. Notamment, il s'agit de préciser la nature de l'extériorité sémiotique : nous reprendrons pour cela les hypothèses élaborées par B. Stiegler (Stiegler, 1994; Stiegler, 1996) selon lesquelles l'extériorité sémiotique consiste dans l'objet technique : c'est l'objet technique qui représente le principe de la constitution des connaissances. Comme nous allons l'argumenter, l'objet est technique dans la mesure où son usage est prescrit par sa structure : la conscience « lit » sur l'objet la manière d'en faire usage, l'objet impose au sujet la manière de l'appréhender. L'hypothèse qui est la nôtre ici est que la conscience puise dans l'objet technique et sa structure les principes de la constitution des connaissances. C'est ainsi que nous voulons mettre en perspective le constat que nous avions effectué dans notre introduction et dans notre première partie : l'outil technique en général, et informatique en particulier, intervient de manière essentielle comme condition de possibilité dans la constitution des connaissances ; cette perspective, prise généralement dans le rapport du sens à la technique, est ce que nous appelons l'« herméneutique matérielle » ; prise particulièrement, en restreignant l'outil technique à l'outil informatique, cette perspective est celle de l'artéfacture, qui correspond ainsi à une reformulation du programme technologique de l'IA dans le cadre de l'herméneutique matérielle. Nous donnons pour conclure un exemple des problématiques dégagées dans le cadre de l'artéfacture en évoquant le projet HOSPITEXTE de station de lecture professionnelle du dossier patient que nous avons initié dans le cadre de nos activités professionnelles.

Pour argumenter ces différents points, nous adoptons le plan suivant. Dans un premier temps, nous reprenons brièvement la problématique de la constitution en reppelant que le problème de la constitution d'un sens et le problème de la constitution de l'idéal (section 7-1). Dans un second temps (section 7-2), nous argumentons l'hypothèse selon laquelle toute constitution du sens comme idéal repose sur une médiation spatiale et matérielle. Pour cela, nous argumentons que la possibilité du sens comme idéal repose sur la possibilité d'une mémorisation (section 7-2.1), que la possibilité d'une mémorisation revient à reposer la question déjà rencontrée du temps (section 7-2.2) et de l'espace (section 7-2.3), et que cette question trouve une réponse dans l'extériorité matérielle (section 7-2.4).

Dans un troisième temps (section 7-3), nous précisons ce qu'il faut entendre par extériorité matérielle en tant qu'elle peut répondre au problème de mémoriser en vue de la constitution du sens. Nous proposons alors de considérer le problème de la constitution du sens à partir de l'extériorité matérielle comme un problème d'inscription matérielle du sens, où la matérialité extérieur devient une matérialité signifiante sur le modèle de l'objet technique. Pour cela, nous formulons (section 7-3.1) une caractérisation de la conscience adaptée aux considérations de ce chapitre pour montrer comme la technique (section 7-3.2) permet à une telle conscience de constituer des connaissances et nous terminons en prenant comme cas exemplaire de la constitution technique des connaissances le cas de l'écriture (section 7-3.3). L'affirmation que la technique permet de constituer des connaissances correspond au fait que l'outil technique prescrit la constitution. Pour motiver cette affirmation, nous distinguons les différentes possibilités d'inscription selon la nature du support, pour distinguer le type de lecture qu'ils impliquent et le type de conaissances qu'ils permettent de constituer. Nous en venons alors à distinguer d'une part la raison orale dont le support d'inscription est le corps propre et d'autre part la raison graphique dont le support d'inscription consiste dans les objets matériels mondains. Cette distinction permet à la fois de rendre compte d'une distinction réelle et de justifier le bien fondé de l'hypothèse de la médiation technique.

Dans un quatrième temps (section 7-4), nous suggérons que le fonctionnement sémiotique de la signification tel qu'il est mis en œuvre dans une médiation sémiotique et technique renvoie à un schéma de type herméneutique que nous appliquons à notre propos pour préciser l'opposition entre raison orale et raison graphique.

Dans un cinquième temps (section 7-5), nous rassemblons les résultats proposés dans ce chapitre pour reformuler le projet de l'IA comme artéfacture.

1 Sens et idéalité

Nous avons plusieurs fois souligné, dans notre introduction et dans notre première partie, que le problème soulevé par la pratique de l'IA comme technique informatique consistait dans le fait que les outils formels et computationnels mobilisés permettent de constituer des connaissances dans des domaines non formels. Le problème est par conséquent de comprendre le rôle joué par l'outil informatique dans la constitution des connaissances : nous avons vu que l'informatique en tant que science expérimentale permet de construire des connaissances sur des objets empiriques particuliers, les calculs. Ces objets n'ont en soi pas de rapports particuliers avec les domaines non formels. Nous avons vu par ailleurs que toute constitution du sens, dans une perspective phénoménologique, repose sur la médiation d'une extériorité sémiotique. L'hypothèse qu'il faut donc creuser est que l'informatique fournit l'extériorité sémiotique permettant la constitution des connaissances dans le domaine. Mais, puisqu'il ne s'agit pas d'une construction scientifique au sens kantien dans la mesure où les connaissances ainsi constituées ne sont pas des connaissances scientifiques portant sur les calculs, il en résulte qu'il faut envisager un cadre plus large que la construction scientifique pour cette constitution des connaissances. Il est donc nécessaire, pour comprendre comment l'informatique peut être l'extériorité sémiotique permettant la constitution de connaissances, d'adopter une démarche où l'on précise la nature de cette extériorité pour y reconnaître l'informatique, dans la mesure où la non formalité du domaine, étant irréductible à la formalité de l'informatique, interdit de rapporter la constitution des connaissances en IA à la construction informatique d'objets computationnels.

Pour préciser ce que peut être l'extériorité sémiotique, nous reprenons la constitution du sens à partir du problème de l'idéalité. En effet quand on parle de sens, que faut-il comprendre par là ? Il y a sens quand quelque chose est compris, interprété, non en fonction de ce qu'il est, mais en fonction de ce qu'il n'est pas. Contrairement à l'être qui se donne lui-même comme ce qu'il est, le sens se donne comme ce qu'il n'est pas. Par exemple, une trace dans le sol renvoie au passage d'un animal, c'est là son sens ; le signal acoustique qu'est un signifiant linguistique renvoie à un signifié. Mais en tant qu'ils sont, ces traces ou signifiants linguistiques s'analysent en fonction de leur forme, de leurs constituants physiques, si bien qu'il n'est pas possible de retrouver par l'analyse de ce qu'ils sont la compréhension en termes de ce qu'ils ne sont pas. Par exemple, la trace dans le sol est causée dit-on par un animal : mais elle pourrait être causée par tout autre sorte de processus physique ; c'est notre compréhension du monde qui rapporte de manière privilégiée la trace à l'animal qui vient de passer. La trace n'est pas une notion reposant sur la causalité physique dans la mesure où cette dernière ne permet pas de la caractériser en tant que trace.

Ces exemples suggèrent que le sens fonctionne comme une adresse et un renvoi: il y a du sens si quelque chose veut dire quelque chose, s'il s'adresse à nous comme signifiant quelque chose et renvoyant par conséquent à une signification. Le sens fonctionne donc comme un renvoi. Mais il ne s'agit pas d'un renvoi de type quelconque, comme la corrélation entre un fait singulier et un autre fait singulier. Dans ce cas, en effet, on ne parle pas de sens. Il n'y a de sens que si le renvoi rapporte le fait à un concept qui en donne la compréhension. Cela signifie que le sens rapporte le fait empirique à un concept idéal qui en donne le principe. Par exemple, le signifiant linguistique se comprend à partir du signifié qu'il exprime, la trace à partir du rapport unissant un animal aux traces qu'il laisse dans le sol[1].

Le sens est ainsi le rapport d'un fait empirique à son type idéal: le fait n'est pas compris comme le fait singulier qu'il est, mais comme l'occurrence d'un type qu'il n'est pas. Tant que le fait n'est pas rapporté à un type, il n'y a pas de compréhension, le fait ne fait pas sens.

Le sens correspond au dépassement de la pure présence du fait : il n'y a sens pour un organisme que si ce dernier dépasse la simple facticité de ce qui lui est donnée dans sa pure présence pour faire déborder l'actualité du fait vers l'inactualité du passé révolu et du futur à venir. Le fait singulier fait sens s'il est plongé dans un horizon temporel qui dépasse l'actualité instantanée de sa donation. En d'autres termes, il faut que le fait singulier et son actualité soit compris à partir d'un type conceptuel dont la temporalité est une présence dilatée pour prendre en compte toutes les occurrences passées et celles à venir : le type est omnitemporel[2]. Si de l'encre noire sur la papier est comprise comme étant la représentation graphique du chiffre deux, le type graphique « deux » prend en charge toutes les traces passées et futures.

Mais dès lors se pose le problème de comprendre comment plusieurs faits distincts peuvent être des occurrences d'un même type idéal. Quel peut être le principe du passage de l'occurrence au type, quelle peut être la genèse du type à partir du fait ? Comment un fait peut-il devenir une occurrence ? Notre hypothèse est que la genèse de l'idéal procède de la matérialité spatiale.

2 Sens et matérialité

2.1 Sens, idéalité, mémoire

Le problème du sens est le problème de l'idéal, comme nous venons de le voir. Ne fait sens que ce qui est interprété à l'aune d'un concept qui le subsume comme son occurrence. Ne fait sens que ce qui dépasse la dimension factuelle, ce qui dépasse la présence événementielle du fait. En d'autres termes, il n'a y de sens que s'il existe une instance interprétative qui dépasse la présence pour viser l'absence comme présence possible et constituer ainsi la signification de la présence actuelle. Or, par définition, cette instance est la conscience, ce pour quoi il y a sens, ce avec quoi il y a sens.

1. Il s'agit par conséquent de distinctions analogues à celles qu'établit Husserl dans la première *Recherche logique* (Husserl, 1959) entre expression et indication (voir plus haut, chapitre 6).

2. C'est un thème explicité par Husserl dans *Expérience et Jugement* (Husserl, 1991) selon lequel le mode temporel de l'idéalité n'est pas l'éternité, mais l'omnitemporalité : la différence est que l'omnitemporalité n'est pas originaire comme l'éternité (comme par exemple l'éternité de Dieu) mais constituée à partir des modes de la temporalité.

Le sens n'est en effet sens que pour une conscience, un esprit pour lequel il y a sens. Le sens est une adresse : il s'adresse à quelqu'un qui le comprend en le subsumant sous l'idéalité. C'est l'esprit qui interprète le fait singulier comme comme occurrence d'un type idéal. Par conséquent c'est au niveau d'une conscience, dans les termes de l'esprit qu'il faut poser le problème de la genèse de l'idéal. Par quel principe un fait singulier donné à la conscience est-il compris par elle comme un type ? C'est donc par une phénoménologie de la conscience et de l'esprit qu'il faut en passer pour résoudre ce problème.

La phénoménologie husserlienne a en effet montré (cf. chapitre 6) que le principal problème d'une phénoménologie de la connaissance est de rendre compte de son caractère idéal. L'idéalité de la connaissance se traduit par le fait qu'elle possède une identité à soi qui permet de reconnaître que différents actes de la conscience relèvent d'une même connaissance. Dans les termes husserliens, le problème est de savoir comment les données matérielles de la conscience, les faits singuliers qu'elle rencontre dans sa vie sensitive, peuvent-elles s'animer d'un sens qui leur confère intelligibilité et compréhensibilité ? Comment différentes données hylétiques peuvent-elles être comprises comme identiques, ou comme participant à la constitution d'un même type ? La phénoménologie de la conscience doit donc comprendre comment l'idéalité peut définir le même et l'identique à travers la différence.

Par ailleurs, ce que nous ont permis de comprendre les analyses kantiennes et husserliennes de la constitution des concepts, c'est que, d'un point de vue phénoménologique, la genèse de l'idéalité passe par le problème de la mémoire : en effet, pour reconnaître parmi des réalisations matérielles *individuelles* les occurrences d'un même type idéal, il faut qu'il y ait *mémorisation* des occurrences passées pour que leur confrontation aux occurrences actuelles permette la constitution du type comme principe commun idéal auquel elles participent. Il ne peut y avoir de perception de l'identité entre deux faits singuliers que s'il est possible de les comparer et de les opposer.

Or, dans la conscience, les faits de la conscience s'ordonnent selon l'ordre du temps, dans un flux temporel sans qu'un ordre spatial n'apporte une juxtaposition nécessaire au dégagement du type à partir de faits reconnnus identiques grâce à leur juxtaposition. Dans la conscience, les faits ne sont pas simultanés, mais nécessairement ordonnés dans le flux temporel, si bien que quand un fait est présent à la conscience, les autres sont perdus dans le passé ou à venir dans le futur. Pourtant, il faut bien être capable de rendre présent simultanément ces faits de la conscience pour qu'un type idéal puisse émerger. Il faut donc être capable de confronter un fait passé et un fait présent pour rendre possible cette émergence.

Le seul moyen de comparer un fait passé et un fait présent pour constater leur identité et donc leur participation en tant qu'occurrences à un même type, est de garder présent à la conscience le fait passé pour qu'il puisse être comparé aux faits actuels et futurs. Or, garder présent à l'esprit, c'est ne pas oublier, c'est en garder la mémoire. Le problème de la genèse de l'idéal, c'est-à-dire de la genèse du sens, est par conséquent le problème de la temporalité de la conscience et de la mémoire.

2.2 Mémorisation et temporalité

2.2.1 Phénoménologie et temporalité

On a vu dans le chapitre précédent la structure de la temporalité telle qu'elle est formulée dans la phénoménologie husserlienne, notamment dans les *Leçons sur la conscience intime du temps*. On se souvient que, pour Husserl, la conscience consiste en un flux temporel d'impressions ponctuelles. En elle-même, chaque impression ponctuelle est une donnée singulière, unique et isolée, indépendante des autres impressions composant le flux. Dans un tel flux, les notions de présent, passé et futur ne sont pas données, mais doivent être construites à partir du flux. En effet, le flux se compose comme un amas d'impressions ponctuelles instantanées qui, *a priori*, n'ont rien à faire les unes avec les autres. Autrement dit, il n'est pas possible, avec ce seul flux comme structure temporelle, d'avoir des relations de « arrivé avant », « survenu après », c'est-à-dire de passé et de futur. L'ordre du flux n'est pas donné, mais à constituer.

Il est ainsi impossible de comprendre comment, au niveau du flux, quelque chose peut être *présent* à la conscience. Une impression ponctuelle est instantanée, elle possède une actualité quand elle survient, mais elle n'est pas présente au sens où les autres impressions ponctuelles sont nécessairement « avant » ou « après ». Ces autres impressions ponctuelles possèdent leurs propres instants d'actualité, qui restent incommensurables entre eux.

L'un des principaux résultats de des analyses husserliennes est que, pour qu'il y ait un présent qui soit autre chose qu'une impression évanescente, isolée des autres impressions, pour qu'il puisse y avoir la présence de quelque chose, présence au sens temporel et propre du terme car distinguée d'un passé et d'un avenir, il est nécessaire qu'il existe un principe mettant en relation et en rapport les impressions ponctuelles du flux de la conscience pour que s'établisse une structure temporelle entre ces impressions. Ce principe est la rétention, par laquelle une impression ponctuelle « retient » dans son actualité l'actualité d'autres impressions ponctuelles comme étant « tout juste passées ». L'impression ponctuelle peut alors devenir un présent, un « grand présent », par opposition à la présence de l'impression ponctuelle isolée, c'est-à-dire son actualité. C'est donc ainsi que le principe de la rétention, avec son corrélat vers le futur qu'est la protention, permet de constituer à partir du flux des impressions ponctuelles les vécus, présents, passés et futurs. La rétention permet de constituer l'avant et l'après du vécu de la conscience.

Mais, pour qu'il y ait confrontation des vécus constitués, passés et présents, c'est-à-dire pour qu'il y ait confrontation des souvenirs (souvenirs secondaires dans la terminologie husserlienne par opposition au souvenir primaire qu'est la rétention) et des vécus présents, il faut que la rétention étende sa retenue du passé aussi loin qu'il est nécessaire : pour qu'il y ait rétention secondaire, il faut que la rétention soit en droit infinie. En effet, pour qu'il y ait souvenir, il faut qu'il y ait une relation avec le présent, avec l'instant présent. Or cette relation ne peut se construire qu'à l'aide de la structure rétentionnelle. Cette structure doit donc étendre ses liens aussi loin qu'il est nécessaire pour qu'il puisse y avoir des souvenirs : la rétention doit donc être infinie. En effet, si de fait on ne se souvient pas de tout, il n'existe pas de borne fixée *a priori* au delà de laquelle le souvenir est impossible. Si on ne se souvient pas de tout, on peut se souvenir aussi loin qu'il est possible : souvenir de l'enfance notamment. Puisque de fait, le souvenir peut reculer aussi loin que possible dans le passé, et que de droit le souvenir repose en sa possibilité sur la rétention, il faut que, de droit, la rétention soit infinie.

Nous avons vu dans le chapitre précédent que l'infinitude de la rétention était une manière pour Husserl de considérer le flux des impressions primaires comme autoconstitué comme totalité et unité, dans la mesure où il faut bien que le flux existe comme flux : en effet, collection d'impressions isolées, instantanées et indépendantes, le flux doit réunir des impressions qui n'ont rien à faire les unes avec les autres. Il n'existe aucune solidarité des impressions entre elles justifiant le fait que l'on puisse les considérer dans un même flux dans lequel on reconnaît la conscience. Pour que soit constitué le fait qu'une impression ait affaire avec une autre impression ponctuelle dans une relation de rétention, il faut que le flux s'autoconstitue comme un espace où les impressions acquièrent une solidarité les unes avec les autres simplement du fait d'appartenir au même flux, tout comme les points de l'espace géométrique acquièrent une solidarité non par eux-mêmes mais par le fait d'appartenir à une même totalité qui leur donne une valeur du fait de leur seule appartenance à la totalité.

Ce faisant, Husserl se donne une totalité actuelle, une infinitude ininterrogée qui revient au fait que le temps reste toujours pensé sur le modèle de l'espace malgré l'effort phénoménologique considérable consacré à élaborer la synthèse temporelle non selon une juxtaposition tenant lieu de principe spatial mais une succession, tenant lieu de principe temporel. En fait, la succession est une notion spatiale, et, Kant comme Husserl, ont davantage pensé le temps comme un *espace* unidimensionnel totalement ordonné rassemblé en une totalité par les infinis de la rétention et de la protention, que dans son authenticité.

2.2.2 Possibilité de la mémoire

Mais, dans l'optique du présent chapitre, une autre argumentation s'impose pour arriver à la même conclusion de l'impossibilité d'une infinitude rétentionnelle. En effet, le chapitre précédent discutait l'analyse husserlienne dans la mesure où elle pensait le temps sous l'égide de l'espace. Ici, nous voulons montrer que l'analyse husserlienne se heurte au problème que l'infinitude rétentionnelle nie toute possibilité du souvenir et de la mémoire, car elle rend ces notions inutiles.

En effet, si la rétention est actuellement infinie, cela signifie que chaque impression ponctuelle du flux de la conscience se retrouve porter dans son cortège rétentionnel la totalité du flux passé. La mémoire comme réactivation d'un passé englouti n'a donc plus lieu d'être: il n'y a jamais rien d'oublié. Affirmer l'infinitude rétentionnelle, c'est abolir la différence opposant le souvenir primaire au souvenir secondaire. Or, cette différence est constitutive du dispositif husserlien d'analyse de la temporalité: en effet, cette différence consiste dans le fait que le souvenir primaire ou rétention est en fait une *intentionnalité perceptive* du tout juste passé et non une *remémoration* ; Husserl considère la rétention du côté de la perception, la perception du passé *comme tel*, et non du côté du souvenir qui ne perçoit pas le passé comme tel mais le réactive dans un présent, qui le rend à nouveau présent. Comme le souligne avec force Husserl, la rétention est une perception du passé qui ne le rend pas présent, qui l'appréhende comme passé sans le rendre présent, alors que le souvenir rend présent le passé sans l'appréhender comme passé. Le souvenir primaire *vit encore dans le présent le passé* alors que le souvenir secondaire *fait revivre le passé dans le présent*.

Si la rétention étend sa portée à l'infini, la réactivation devient inutile: se souvenir veut dès lors dire retenir le passé dans le présent. A partir du moment où la rétention est infinie, rien n'est jamais oublié, tout reste en quelque sorte accessible depuis le présent, et par conséquent la réactivation du passé dans le présent est dès lors inutile.

Mais, on comprend bien la nécessité de postuler l'infinité de la rétention : outre son rôle dans l'autoconstitution du flux de la conscience, il faut bien que ce qui a été vécu laisse une trace dans le présent pour qu'il puisse être réactivé. En effet, sans la toile de fond que constitue le flux comme unité, plus rien ne relie le passé au présent, et par conséquent le passé tombe dans le néant.

On arrive ainsi au paradoxe que, impliquée par la nécessité de la mémoire, l'infinitude rétentionnelle en supprime la possibilité dans la mesure où tout est retenu dans la présence de la rétention. Puisqu'il paraît difficile de récuser cette évidence phénoménologique qu'est la mémoire, il faut donc renoncer à prôner l'infinitude de la rétention. Mais si on le fait, on perd alors la possibilité de rendre présent le passé, et donc la possibilité d'une genèse de l'idéalité.

Si on comprend bien pourquoi Husserl devait affirmer l'infinité de la rétention dans ses analyses, il faut se demander pourquoi il avait conservé simultanément le principe du souvenir secondaire et de la mémoire, et il faut découvrir quelle nécessité phénoménologique l'y forçait. C'est en effet en mettant en évidence une telle nécessité qu'il sera possible de savoir comment d'une part affirmer la finitude de la rétention et d'autre part de conserver la possibilité du souvenir et de la mémoire, nécessaire à la genèse du sens.

2.2.3 Souvenirs primaire et secondaire

Il semble en effet contradictoire de vouloir affirmer avec Husserl l'infinité de la rétention et l'opposition entre souvenir primaire et souvenir secondaire. Or, ce qui semble légitimer cette dernière opposition, c'est bien que, dans le présent vivant de la conscience, le passé tout juste passé est encore présent en tant que passé, alors que dans le cas du souvenir secondaire, le souvenir est revécu dans le présent vivant : la reproduction n'est pas production. En effet, le présent de la reproduction ne contient pas dans sa vivacité, dans l'authenticité de la présence, ce qui est passé, contrairement au souvenir primaire.

C'est donc que l'opposition entre rétention et souvenir repose sur l'opposition entre production et re-production. Il y a une nécessaire distorsion entre la reproduction du souvenir secondaire et la présence productive du présent vivant dans le souvenir primaire. Cela signifie par exemple qu'entre entendre une mélodie et s'en souvenir, la différence n'est pas d'intensité : la re-production ne peut être qu'une déformation, qu'une re-formulation de la mélodie.

On peut s'interroger alors sur la frontière démarquant le primaire du secondaire, la rétention du souvenir. A partir de quel moment, la présence du passé comme passé se mue en présence du passé comme nouvelle présence, comme re-présentation ? En d'autres termes, il faut disposer d'un critère permettant de discriminer entre la rétention primaire et le souvenir. Par définition même de la rétention, le présent est une impression ponctuelle qui est retenue dans le présent quand elle s'enfonce dans le passé. Cela signifie que l'individualité de l'impression ponctuelle, ce qu'elle est et ce qu'elle contient, ne dépend pas de ce qui est passé : au contraire, c'est le passé qui procède du présent et qui est retenu. Dans l'analyse husserlienne de la conscience du temps, il n'y a pas de place pour une rétroaction du passé sur le présent : le présent est une réceptivité pure et absolue entièrement constituée par l'impression ponctuelle. Il ne peut y avoir de rétroaction du passé sur la détermination du contenu de l'impression ponctuelle. Chaque « grand maintenant », c'est-à-dire la fenêtre temporelle constituée d'un présent vivant ponctuel muni de son cortège de rétentions et protentions, est une durée dont le contenu est donné à la conscience et non constitué par elle, entièrement original et autodéterminé.

Par conséquent, il est exclu que le souvenir d'une mélodie puisse altérer l'entente actuelle de cette mélodie, pour reprendre les exemples chers à Husserl, et que le souvenir d'un son puisse participer à la détermination et la constitution du son actuellement entendu. La distinction permettant de discriminer le souvenir primaire du secondaire est que le souvenir comme vécu est une re-production qui peut donc être modifié par le passé, alors que le présent n'est pas constitué par le passé mais par l'actualité de l'impression ponctuelle et de ses rétentions. C'est donc là la nécessité phénoménologique que Husserl a retenu pour mener ses analyses et conserver une notion de souvenir malgré l'infinitude de la rétention qu'il affirmait simultanément.

2.3 Temps et espace

Mais, si l'on se tourne vers l'expérience phénoménologique de la conscience, il apparaît que le présent n'est pas une donnée brute et originale. C'est-à-dire que l'entente actuelle d'une mélodie, phénoménologiquement, dépend des ententes passées. Comme le souligne B. Stiegler (Stiegler, 1996), les nouveaux supports d'enregistrement permettent de souligner particulièrement ce fait : à partir du moment où l'on peut disposer de l'enregistrement *exact* d'une mélodie, et que l'on peut par conséquent avoir deux fois de suite la même audition d'une mélodie, le fait que néanmoins on perçoive différemment la mélodie d'une fois sur l'autre ne peut qu'être imputé à une rétroaction du passé sur le présent, de la première écoute sur la seconde.

Le présent est donc constitué par le passé, et ne consiste pas simplement dans la réception passive d'un contenu indépendant de la mémoire de la conscience qui le reçoit. Cela signifie que phénoménologiquement parlant, il est impossible de revivre deux fois la même chose, d'appréhender deux fois le même contenu, de reconnaître deux fois le même présent. Autrement dit, on ne se baigne jamais deux fois dans le flux de la conscience.

Pour Husserl, le problème est que le flux temporel emporte tout sur son passage et qu'il faut expliquer comment une durée ou vécu peut se constituer comme unité à partir d'une série d'impressions ponctuelles indépendantes les unes des autres. Sa solution, on l'a vu, et de considérer que le flux s'autoconstitue comme une unité totale, une co-appartenance des impressions ponctuelles leur permettant de participer à une même durée. Il en découle que, les impressions étant distinctes puisque ponctuelles et néanmoins comparables puisque appartenant à un même flux autoconstitué, la donation d'une impression ponctuelle ne peut pas être influencée par d'autres impressions ponctuelles, fussent-elles reliées par une relation temporelle de passé avec elle.

Le constat « phénoménologique » selon lequel souvenir primaire et souvenir secondaire s'opposent selon leur originarité est donc commandé par le préjugé selon lequel la conscience est un flux autoconstitué d'impressions ponctuelles. C'est encore une fois le préjugé de l'espace pour penser le temps qui fait apparaître une indépendance de l'impression ponctuelle à l'instar du point dans l'espace géométrique. La conception du temps comme *continuum* conduit Husserl à prôner des évidences phénoménologiques qui ne sont valables que depuis ce préjugé : en effet, si l'on s'abstrait d'un tel préjugé, le présent est toujours plein du passé, et la donation n'est jamais ponctuelle et indépendante de celles tout juste passées. Il semble donc que l'origine des difficultés rencontrées par les analyses husserliennes réside dans le fait d'avoir considéré le flux de la conscience davantage comme une continuité mathématique dont la rétention constitue la structure différentielle que comme un flux héraclitéen.

La volonté de distinguer un souvenir primaire ou rétention du souvenir secondaire, c'est-à-dire du souvenir dans son sens habituel, provient de la nécessité de disposer d'un principe de contiguïté entre les impressions ponctuelles. La rétention est un artefact du préjugé ensembliste du flux de la conscience. La finesse phénoménologique de Husserl lui a permis de voir que le souvenir secondaire ne pouvait être assimilé au souvenir primaire et qu'il fallait conserver un clivage net entre les deux. Mais, la rigueur phénoménologique conduisait au paradoxe puisque l'infinitude rétentionnelle rend inutile la notion même de souvenir.

Mais si la rigueur phénoménologique justifie la notion de souvenir secondaire, en revanche celle de rétention paraît bien moins assurée. Au contraire, il apparaît que chaque instant présent se constitue depuis les instants passés : le passé constitue le présent au lieu d'être institué par lui. Mais se passer de la rétention, c'est s'exposer aux difficultés que Husserl voulait éviter en élaborant ce concept : par conséquent, soit il faut prôner une infinité de la rétention qui rend possible la mémoire tout en la rendant inutile, soit il faut affirmer la finitude de la rétention, et donc affirmer l'utilité de la mémoire sans en donner le principe de sa possibilité.

En outre, cela signifie qu'il est impossible d'avoir conscience de l'identique ou du même si l'on veut faire reposer cette conscience sur la comparaison du passé et du présent, puisque la possibilité de la mémoire n'est plus assurée et que, quand bien même elle le serait, le présent étant à chaque fois différent du passé puisque constitué par lui, la considération synoptique du passé et du présent ne pourrait conduire qu'à constater les différences et non un identité génitrice de l'idéalité. En effet, si mémoire il y a, si répétition grâce à la mémoire il y a, cette répétition instaurera des différences et non des identités. Ce qu'il faut donc, c'est expliquer d'une part la possibilité de la mémoire, c'est-à-dire de la répétition, et d'autre part de la répétition comme identité et non comme différence, pour qu'il y ait sens.

Or, il y a sens : la temporalité de la conscience est donc phénoménologiquement incapable de rendre compte de la conscience d'identité pourtant attestée phénoménologiquement. En effet, il faut rendre compte de ces deux faits phénoménologiques qui, à la lumière des présentes analyses, peuvent paraître incompatibles entre eux : d'une part il est impossible de réunir des instants vécus en une même appréhension (possibilité de la répétition) pour y constater des identités (possibilité d'une répétition comme identité), d'autre part il est phénoménologiquement avéré qu'il peut exister une telle conscience d'identité puisqu'il y a sens.

Comment se sortir d'une telle difficulté ? Il faut ici se souvenir que l'impossibilité d'avoir deux fois la même conscience repose sur une analyse *temporelle* de la conscience. Cela signifie que si, d'un point de vue temporel, il est impossible d'avoir deux fois (répétition) la même conscience (identité), c'est que cette conscience d'identité procède d'un principe qui n'est pas le temps, un principe qui assure la permanence d'un contenu (répétition) et la possibilité de sa réappropriation à l'identique (répétition comme identique) à un moment ultérieur.

Notre hypothèse est que ce principe est l'espace, principe que Husserl avait mobilisé, de manière problématique à notre sens, pour poser l'autoconstitution du flux impressionnel de la conscience. Si l'on refuse de mobiliser l'espace au sein même de la conscience temporelle, c'est-à-dire au sein de l'intériorité de la conscience, c'est qu'on le renvoie à son extériorité. L'espace correspond par conséquent à la permanence dans l'extériorité de contenus que la conscience peut ressaisir à l'identique, ou du moins d'une manière autorisant en principe sinon en fait une ressaisie à l'identique.

2.4 Matière et espace

Mais l'espace en lui-même n'est rien : il n'est qu'un cadre formel dans lequel la matérialité des choses apparaît. En effet, le principe spatial n'est pas un contenu : si c'est l'espace qui rend compte de la permanence d'un contenu qui dure au milieu du flux de la conscience, ce n'est pas l'espace qui peut être le principe de ce contenu. Cette constatation a pour conséquence que si, phénoménologiquement, il existe une conscience d'identité, et qu'une telle conscience doit nécessairement reposer sur un principe spatial, et que ce principe spatial pour être effectif doit reposer sur un contenu, il doit nécessairement exister un contenu extérieur à la conscience, et ce contenu est ce qu'il faut appeler la matérialité, la matière[3].

L'espace et la matérialité qui s'y inscrit apporte donc la solution à notre problème de la possibilité de la mémoire et de la genèse du sens. En effet, la matière peut se comprendre comme un contenu permanent. A ce titre, elle remplit une double fonction. La première est qu'elle conserve le passé dans le présent en lui donnant un support permanent, si bien que le passé est constamment répété dans le présent. La seconde est que la répétition est toujours une répétition identique à soi, puisqu'il s'agit d'une permanence. Par conséquent, la constitution ne dépend plus de l'intériorité de la conscience, mais de l'extériorité spatiale et matérielle. On aurait ainsi une genèse matérielle du sens.

Mais la question se pose de savoir ce qu'il faut entendre par matière dans l'espace. Il ne s'agit pas en effet du concept de matière au sens tel que la physique moderne permet de le construire. En effet, l'espace est compris ici, comme nos discussions de Kant et Husserl l'ont montré, comme étant un principe de simultanéité et de juxtaposition, c'est-à-dire un principe de coprésence. La matière désigne donc ici ce qui est présent, ce qui est juxtaposé dans l'espace. Il est donc nécessaire de préciser cette notion fort vague de matière pour déterminer ce qu'il faut comprendre par génèse matérielle du sens.

3 Sens et matérialité

3.1 Conscience, intériorité, extériorité

Avant de tirer toutes les conséquences de l'introduction d'un principe matériel dans la constitution du sens, il est nécessaire de revoir notre conception de temps et de la

3. Ainsi que la célèbre « réfutation de l'idéalisme » (Kant, 1980) kantienne le montre.

conscience comme flux temporel. En effet, nos analyses précédentes se sont fondées sur la conception husserlienne de la conscience pour en souligner les difficultés et leur nécessaire dépassement par l'introduction de la médiation spatiale. Il sera impossible de tirer toute l'intelligibilité de cette introduction si nous gardons des caractérisations qui y sont par principe opposées.

Le temps husserlien est un temps constitué dans l'intimité de la conscience comprise à son origine ultime comme une continuité d'impressions ponctuelles, totalité close dotée d'une structure différentielle grâce à la rétention. Le temps husserlien se trouve ainsi compris sous le primat de la présence, de l'impression ponctuelle actuelle, qui commande par son cortège de rétention et de protention toute la structure de la conscience.

Cette construction est impropre, comme on l'a vu, à expliciter la genèse du sens dans la mesure où elle ne rend pas compte de ses caractéristiques essentielles. Le sens est en effet une adresse, avons-nous dit, un message adressé à une conscience en attente. Cela nous montre que la conscience ne doit pas être comprise sur le modèle de la présence, mais sur celui de l'attente, de l'attente vigilante au sens où on « s'attend à quelque chose ».

Si la conscience est attente, cela signifie qu'elle est toute orientée sur le futur et non plus sur le présent, et ouverte sur l'extériorité dont elle attend le message[4]. Fait donc sens ce qui répond à une attente, ce qui correspond à une possibilité projetée par la conscience vers l'avenir. La conscience n'est donc pas structurée comme un flux impressionnel mais selon les structures de l'attente, c'est-à-dire les structures du sens. La conscience est donatrice de sens pas tellement dans la mesure où elle anime des données matérielles mais dans la mesure où elle reconnaît, grâce à sa structure, ce qui lui est adressé.

Vigilante, l'attente n'est pas passive : la conscience n'attend pas passivement qu'un message lui soit adressé par l'extériorité. Au contraire, conformément à ses attentes, la conscience se projette dans le monde et agit de manière à construire les réponses qu'elle attend. Si bien que l'extériorité devient pour elle le moyen de réaliser ses attentes et le lieu d'accomplir les actions transformant le monde pour les remplir. La conscience possède dès lors des connaissances sur le monde qui structurent ses attentes, et donc se comportent comme des *actions possibles* : telle action permettant d'atteindre tel but, d'assouvir telle attente.

Mais il existe toujours un décalage entre le monde, l'extériorité, et les attentes : ouverture et projection vers l'extériorité, la conscience n'embrasse pas la totalité du monde et est constamment débordée par l'infinité du monde. Par conséquent, les réponses suscitées par les connaissances comme actions possibles ne sont jamais pleinement conformes aux attentes si bien qu'elles en subissent une refonte permanente, ébauchant ainsi un cycle très classique de révision des connaissances.

Mais le plus important est sans doute de noter que la structure de la conscience, c'est-à-dire ses connaissances, ne peut être *dans* la conscience puisque celle-ci est toute entière tournée vers l'attente. La conscience possède sa structure d'attente en dehors d'elle-même, dans son extériorité, c'est-à-dire dans la matière spatiale. La technique montre quotidiennement que les connaissances sont externalisables, comme nous allons brièvement le rappeler.

4. Cette caractérisation de la conscience n'est pas sans rappeler l'analytique existentiale du *Dasein* qu'effectue Heidegger dans *Être et Temps* (Heidegger, 1986). Il n'est pas dans notre propos d'aborder ici la philosophie heideggerienne et notre discussion doit être considérée ici davantage comme une conséquence de ce que nous avons dit que comme une allusion à de grandes philosophies.

Le point le plus important est que les connaissances externalisées sont bien des connaissances dans la mesure où les outils, les machines, etc., renvoient à des actions possibles dont ils sont le principe. Mais le point fondamental est que les connaissances externalisées permettent, grâce à cette externalisation, c'est-à-dire cette spatialisation, de constituer de nouvelles connaissances, c'est-à-dire de nouvelles actions possibles. Ce que montre la technique, c'est que les connaissances peuvent être externalisées, ce qu'ont montré nos analyses phénoménologiques, c'est qu'elles le doivent. Le fait même que l'externalisation permettent la genèse de nouvelles connaissances illustre bien le fait que l'extériorité est au principe de la connaissance.

Toute connaissance ne pourrait donc être qu'externe. Une telle affirmation semble aller à l'encontre de l'évidence commune selon laquelle nous avons des représentations internes ou mentales, des connaissances intériorisées. Notre hypothèse, concernant ce type de connaissances, est que ces connaissances sont en fait externalisées, mais externalisées sur un support tellement proche de la conscience que l'on se méprend sur leur caractéristique. Ce support est le corps propre. Ce qui fait que les connaissances externalisées sur le corps propre sont dites habituellement internes, c'est que ce support est interne par rapport aux supports externes que sont les objets extra-corporels. Mais ce support « interne » est extérieur à la conscience. Nous y revenons dans les sections suivantes en nous essayant à déterminer ce qu'un tel support interne peut impliquer quant aux connaissances qu'il « supporte » par rapport aux supports « externes », c'est-à-dire extra-corporels.

En affirmant que la conscience ne possède pas d'intériorité, nous affirmons en quelque sorte que la conscience doit puiser dans son extériorité les principes de la constitution qu'elle effectue du sens. Il nous semble que la technique permet d'illustrer ce fait et qu'il est possible de considérer que l'extériorité sémiotique de la conscience lui permettant de constituer le sens est une extériorité technique. Pour cela, nous allons étendre le sens du mot technique pour qu'il signifie ici toute extériorité spatiale et matérielle dont la structure prescrit l'usage ou l'interprétation sans néanmoins les déterminer. C'est un élargissement de ce qui est habituellement compris comme technique dans la mesure où les objets techniques ne sont pas nécessairement construits (au sens d'une production industrielle ou artisanale) et qu'ils ne procèdent pas obligatoirement d'une intentionnalité préalable qui leur assigne une fonction ou signification en vue de laquelle ils sont construits. Au contraire, avec notre définition, une telle intentionnalité est davantage une conséquence de l'objet technique que sa cause (finale). Cela signifie également que ce nous appelons technique est dans ces conditions originaire et non constitué : la conscience rencontre des objets qui, en tant qu'ils sont techniques, lui prescrit un usage ou une signification à partir desquels elle constitue ses connaissances. Mais ces objets n'ont pas obligatoirement d'origine assignable dans la mesure où le seul moyen de les aborder et de les caractériser est de les considérer comme « prescripteurs de sens » dans qu'il soit possible de remonter à une instance originaire qui leur aurait donné ce sens. D'une certaine manière, la technique correspond à ce qui renvoie à un sens sans renvoyer à une origine du sens, à une intentionnalité sans auteur[5]. En d'autres termes, la catégorie « technique » embrasse par conséquent la notion habituelle de technique augmentée de la notion d'intentionnalité originaire sans origine.

Dans les sections suivantes, nous tentons de préciser ces considérations en motivant notre définition d'objet technique par quelques exemples.

5. Cette caractérisation de la notion de « technique » est très proche de ce qu'en propose B. Stiegler dans (Stiegler, 1995).

3.2 Matière et technique

Si l'on convient en effet que la connaissance peut se définir comme une action possible[6], c'est-à-dire le fait de savoir qu'une action peut être entreprise, et comment, l'outil se définit comme un support matériel qui mémorise dans sa structure technique l'action possible dans la mesure où cette structure commande, programme, l'utilisation de l'outil. La structure technique de l'outil détermine toutes les actions possibles que l'on peut entreprendre avec cet objet et c'est en cela qu'il mémorise une connaissance : l'outil opérationnalise le rapport au monde qu'est la connaissance comme action possible dans le monde et l'opérationnalise de manière à mémoriser ce rapport.

Non seulement le marteau permet d'enfoncer des clous, mais beaucoup d'autres objets permettent de le faire, mais sa structure de marteau mémorise la manière de prendre comme un objet pour frapper un clou ; le manche du marteau et sa tête massive et lourde prescrivent la manière de le saisir et de le manipuler. C'est en cela que l'outil mémorise une action possible, la connaissance pratique correspondant au fait d'enfoncer les clous, et n'est pas qu'un moyen parmi d'autres d'opérationnaliser cette connaissance : c'est le moyen privilégié d'enfoncer des clous et il est même conçu pour représenter dans sa structure le meilleur moyen d'enfoncer des clous.

C'est pourquoi il faut bien distinguer l'*outil* du *moyen* : le moyen opérationnalise une connaissance, mais d'une manière spontanée et contingente ; en d'autres termes, le moyen n'établit pas un rapport constant entre l'individu et son action dans le monde : le moyen peut être utilisé à d'autres fins sans qu'on y voie une déviation de son usage. En revanche, un outil établit un rapport constant dans la mesure où il prescrit, commande, programme, l'action qu'il permet d'opérationnaliser. Ainsi, utiliser un marteau pour bloquer une roue dentée avec son manche est une déviation de l'usage « normal » de l'outil ; mais utiliser un caillou pour ouvrir des noix n'est pas plus prescrit par la structure du caillou que son utilisation dans le ballast de la SNCF. Le caillou est un moyen, le marteau est un outil. Le moyen *autorise* l'opérationnalisation de la connaissance, l'outil la *commande*. En d'autres termes, le moyen ne mémorise pas un usage particulier et peut être utilisé à de multiples fins qui paraissent toutes aussi pertinentes les unes que les autres. L'outil mémorise un usage particulier, un type d'action, puisque sa structure prescrit son utilisation. C'est donc que sa structure matérielle est l'inscription de la connaissance comme rapport au monde, comme action dans le monde, comme action possible dans le monde.

En permettant de mémoriser l'action possible, l'outil est un support matériel externe de mémorisation. Il correspond à une externalisation de la mémoire[7] : au lieu de garder la mémoire interne psychologique de la manière d'enfoncer les clous et de rechercher à

6. Une telle caractérisation peut paraître bien réductrice. Mais il faut considérer que l'action peut être prise en un sens physique ou matériel, et dans un sens « spirituel » ou mental : une action correspond à toute entreprise effectuée dans le temps du monde ou de l'esprit. Le fait de considérer une action « possible » renvoie au fait que l'on s'intéresse à l'action comme à une idéalité, dans sa possibilité future : on dépasse ainsi le niveau de fait individuel, de l'action comme événement pour le niveau du type idéal de l'action qui embrasse les actions passées et à venir. Comme idéalité, comme possibilité, l'action est dès lors une connaissance, le savoir comme possibilité d'agir dans le monde ou dans l'esprit, le savoir comme action mondaine ou comme pensée. Une telle caractérisation n'est pas sans rappeler le pragmaticisme de Peirce (Peirce, 1960).

7. Ce fait a été particulièrement souligné par Leroi-Gourhan dans (Leroi-Gourhan, 1964) et repris par Stiegler (Stiegler, 1994).

chaque fois qu'il en est besoin un moyen permettant, autorisant, cet usage, l'outil épargne cette recherche de la conformité entre le moyen et l'usage voulu en mémorisant dans sa structure matérielle cette adéquation. L'outil n'est pas un moyen neutre, mais un moyen dédié, dévolu à un usage particulier dont il mémorise l'opérationnalisation puisqu'il prescrit la manière de s'en servir. L'outil décharge donc la mémoire psychologique en devenant une mémoire matérielle externe dont la structure programme son utilisation.

On connaît l'exemple tiré du monde industriel de la presse à tôle métallique : un ouvrier commande l'abaissement et le relèvement du lourd panneau assurant la presse des tôles ; la commande s'effectue au moyen d'un bouton : l'ouvrier met en place la tôle et appuie sur le bouton pour abaisser la presse. Il se trouvait que régulièrement des mains étaient écrasées : une main restait pour la mise en place de la tôle alors que l'autre actionnait l'abaissement de la presse. Il fallait donc que l'ouvrier se souvienne systématiquement d'enlever les deux mains pour abaisser la presse. Plutôt que de se fier à une mémoire humaine dont la principale qualité n'est pas d'être systématique, on entreprit de la décharger en modifiant la structure technique de l'outil : désormais, l'abaissement ne pouvait être commandé que par deux boutons situés sur la presse elle-même ; l'ouvrier devait donc faire usage de ses deux mains pour abaisser la presse si bien qu'il ne demeurait aucun risque qu'une main libre ne traîne sous la presse. Le dispositif technique des boutons de commande prescrit par sa structure un usage dont l'ouvrier n'a pas à se souvenir : retirer les mains avant l'abaissement.

L'outil mémorise de manière externe une connaissance dont l'homme n'a plus à assurer une mémorisation interne. En d'autres termes, la connaissance n'a plus de représentation mentale ou psychologique, mais seulement une représentation matérielle. Cette décharge psychologie est possible dans la mesure où, répétons-le, l'outil possède lui-même les conditions de son utilisation et de sa réappropriation individuelle par l'utilisateur. Par conséquent, la dépsychologisation de la connaissance par sa mémorisation matérielle externe permet qu'un autre que le concepteur de l'outil accède à la connaissance matérialisée par l'outil. L'outil permet de transmettre la connaissance de sujet en sujet puisqu'il prescrit par sa propre structure la méthode permettant de retrouver ce qu'un précédent concepteur a voulu opérationnaliser dans l'outil. En d'autres termes, l'outil *objective* la connaissance.

L'outil, en mémorisant la connaissance sur un support matériel accessible à tous, permet l'objectivation de la connaissance et la reconnaissance de son indépendance à toute représentation mentale ou psychologique. Sur un autre registre, le triangle mathématique n'est pas le triangle psychologique ou représenté mentalement puisque les propriétés du triangle ne dépendent pas de ma représentation mentale mais du triangle mathématique tel qu'il est décrit dans les textes mathématiques et des règles de réécritures autorisées. L'écriture mathématique est un outil, constitué de ses symboles et règles de manipulation, qui prescrit la manière de manipuler les objets mathématiques (en l'occurrence, des symboles d'objets) qui sont indépendantes de la psychologie particulière et individuelle des mathématiciens. C'est en cela qu'elle possède cette objectivité qui étonne tant et qu'elle peut transcender psychologies et cultures. C'est que l'écriture autorise une matérialisation de la connaissance qui décharge l'esprit de ses règles de manipulation : en déchargeant la mémoire de chacun, la matérialisation rend la connaissance accessible à tous.

Mais la mémorisation externe de la connaissance ne permet pas seulement son objectivation, mais également sa genèse. L'objectivation est déjà, dans une certaine mesure, une genèse de la connaissance puisqu'elle permet de constituer la connaissance comme un fait idéal partagé par tous transcendant le fait psychologique individuel. La matérialisation dans la structure des outils techniques de la connaissance est donc une genèse *épistémologique* de la connaissance puisqu'elle en assure l'objectivation, c'est-à-dire son existence en tant que connaissance et non en tant que fait psychologique. Mais la matérialisation est également une genèse *phénoménologique* de la connaissance dans la mesure où elle permet l'émergence de *nouvelles* connaissances. L'objectivation de la connaissance sur un support matériel externe s'accompagne nécessairement d'une spatialisation de la connaissance, puisque tout matière est par principe spatiale. Or, la spatialisation permet la juxtaposition d'éléments qui, pris dans le flux temporel de la pensée discursive, restent séparés par la succession temporelle sans jamais pouvoir être confrontés lors d'une juxtaposition spatiale. La confrontation spatiale permet seule d'établir des relations d'identités et de différence entre les éléments juxtaposés dans le mesure où elle permet de les maintenir ensemble sous le regard et de les saisir en un même regard *synoptique*[8].

La juxtaposition et les relations d'identités et de différences qu'elle autorise permettent de reconnaître que plusieurs éléments étant identiques, ils constituent des occurrences d'un même type. Ces éléments peuvent être des objets statiques, mais aussi des enchaînements temporels successifs identiques. C'est ce que permet l'informatique : en élaborant une géométrie du temps, c'est-à-dire en élaborant une mécanisation du calcul, l'informatique effectue quant au temps ce que la géométrie a réalisé quant à l'espace : la systématisation de rapports permanents et idéaux entre des objets géométriques spatiaux pour la géométrie et des objets algorithmiques temporels pour l'informatique. En matérialisant le calcul, en extériorisant la succession temporelle, l'informatique permet d'objectiver les lois du calcul et de les soumettre à l'investigation expérimentale, c'est-à-dire à la découverte de nouvelles connaissances informatiques.

On rencontre un exemple (Eisenstein, 1991; Martin, 1996) de ce processus quand on considère l'évolution des supports de l'écriture, quand on passe successivement du *volumen* au *codex* puis à l'imprimé. Le *volumen* antique, repris par les Romains aux Grecs, consiste en deux rouleaux parallèles où est enroulé le parchemin support de la matière écrite. De la main gauche, on dévide le premier rouleau, de la main droite, on enroule, au milieu défile le texte qui est lu. Avec un tel système, il est facile de constater qu'il est impossible d'accéder à une information contenue dans le texte *directement* : il faut pour cela dérouler tout ce qu'il y a avant. La lecture d'un *volumen* s'effectue donc sans accès à ce qui va être lu ni à ce qui vient d'être lu : enroulé sur l'un des rouleaux, le texte n'appartenant pas à la fenêtre de lecture est inaccessible, sinon par le rythme linéaire de la lecture.

L'invention du *codex* modifia considérablement les choses dans le mesure où il se composait de feuilles reliées ensemble par un bord, comme nos livres modernes. Cela signifie qu'il était possible d'accéder directement à une page et à son contenu sans passer par les pages précédentes. C'est ainsi qu'en faisant une table des matières et des index, il était possible d'enrichir les possibilités d'accès au contenu. La modification du support, ici le passage du *volumen* au *codex* permettait de concevoir de nouvelles structures, de nouvelles informations sur le texte, c'est-à-dire de nouveaux textes, à savoir la table des matière et les index. L'imprimé vint quant à lui donner la possibilité de reproduire à l'identique ce que le manuscrit déformait lors de la recopie. Au lieu d'une déformation progressive de l'original,

8. Formule pléonastique à proprement parler, puisqu'un regard est toujours syn-optique, c'est-à-dire « voir ensemble ». Cette particularité est fondamentale et a été soulignée par tous les historiens des techniques (Leroi-Gourhan, 1964) et de l'écriture (Bottéro, 1987; Detienne, 1981), comme par les anthropologues (Goody, 1979; Goody, 1985). Nous y revenons dans les sections suivantes.

l'imprimé assurait la répétition à l'identique et la constitution d'un type. Par exemple, il devenait facile de faire la différence entre une coquille et un changement de contenu et de distinguer différentes versions dans un document. Alors que pour un manuscrit, il était impossible de distinguer l'erreur de la correction.

La construction de type à partir d'occurrences permet la construction de connaissances. C'est la répétition de l'action à l'identique, comme une seconde occurrence, qui permet de poser le principe d'une action future à réaliser, mais qui ne l'est pas encore, c'est-à-dire d'une action possible idéale, ou encore une connaissance. C'est le fait de dessiner plusieurs fois le même triangle sur le papier qui fera dépasser les différences individuelles des dessins empiriques pour dégager l'identité de la méthode adoptée pour construire le triangle. Si bien que la méthode de construction est ce qui détient l'objectivité géométrique et non son résultat empirique. Et c'est parce que cette méthode vaut pour tous les triangles passés et *futurs* qu'elle est idéale et constitue une connaissance.

La spatialisation est par conséquent la condition transcendantale de la genèse phénoménologique de la connaissance. C'est parce que la matérialisation, en tant que spatialisation, étale dans l'espace ce qu'il y a à voir que le regard de la connaissance peut s'exercer.

3.3 Matière et écriture

L'exemple de la technique et des outils permet de constater d'une part que les connaissances s'externalisent et que d'autre part elles puisent dans cette externalisation le principe de leur genèse. Il s'agit à présent de montrer qu'en fait toute connaissance est externe. D'un certain point de vue, c'est évident d'après ce qui a été dit : l'extériorité et la matérialité qu'elle manifeste sont les conditions nécessaires à la constitution de sens, c'est-à-dire à la genèse de l'idéal. Mais, d'un autre point vue, cela semble heurter les évidences les mieux ancrées quant aux caractères internes des pensées et des connaissances. C'est pourquoi il faut reprendre la question de la mémoire externe pour comprendre en quoi toute connaissance est bien externe.

3.3.1 Sens, matière et écriture

Comme nous l'avons dit, la matérialité de la mémoire externe permet de constituer la conscience de l'identique, phénoménologiquement impossible sinon. En effet, la spatialité de la matière permet par la juxtaposition de transformer la répétition différentielle imposée par la succession temporelle en une répétition identitaire constituant les éléments répétés en occurrence d'un type idéal.

La construction phénoménologique du sens grâce à la médiation de l'espace s'établit quand la matière devient une *marque* permettant d'accéder à un sens dont elle est la manifestation matérielle. Dès que l'élément matériel n'est plus considéré pour ce qu'il est, mais pour ce qu'il n'est pas, il devient intentionnel et par conséquent le lieu du sens : il adopte la nature d'un signe. Le signe, par sa présence, manifeste ce qui est absent et oublié. C'est pourquoi il est nécessaire que le signe soit matériel et intentionnel : seule la matière est présence car elle est permanente ; mais la présence ne suffit pas, il faut en outre que la matière soit un support animé intentionnellement. Si l'intentionnalité est la marque de l'esprit, cela signifie que la matière doit *être investie d'esprit* pour qu'elle puisse constituer le sens. La constitution du sens, la genèse de l'idéal, consistent dans l'investissement de la matière d'esprit.

En d'autres termes, c'est la matière et son extériorité spatiale qui assurent la rétention nécessaire à la constitution du sens, comme Husserl l'a montré. La rétention, comme présence dans le présent du passé comme passé, ne peut se construire sur un flux ponctuel sinon par préjugé spatial: la présence du passé comme passé ne peut se comprendre que comme *permanence* et non comme *rétention.* Au lieu de faire reposer la constitution du sens sur le présent qui retient, solution imposée par la notion d'un flux qui emporte inexorablement tout sur son passage, il faut au contraire la faire procéder d'un passé qui dure car permanent. C'est la raison pour laquelle, puisque l'intériorité de la conscience est soumise à la temporalité et à son caractère évanescent, fût-elle commandée par l'avenir ou le présent, il est nécessaire de faire appel à un principe externe qu'est la matière spatiale.

Il ne suffit cependant pas que la matière soit permanente pour qu'elle conserve son sens ou l'acquiert. En effet, il est tout à fait possible d'oublier ce que la marque matériel signifie : l'esprit ne sait plus interpréter la marque matérielle comme un élément intentionnel et non pas simplement comme une chose physique, une singularité sans signification. Il faut donc que ce soit la marque matérielle elle-même qui possède dans sa structure les éléments permettant de retrouver le sens qui l'anime intentionnellement. Si seul l'esprit peut investir d'esprit la matière qu'il rencontre, seule la matière peut contenir dans sa structure la manière dont elle doit être investie puisque la finitude rétentionnelle interdit que l'esprit en garde la trace interne : la trace comme subsistance du passé est nécessairement externe. La marque doit commander son interprétation tout comme l'outil prescrit son utilisation.

Néanmoins, la marque ne peut se suffire à elle-même pour qu'il y ait sens. En d'autres termes, il ne peut y avoir du sens que pour une conscience ou un esprit. Si donc la matière est une médiation indispensable à la genèse du sens, il ne peut s'agir en aucune façon d'une réduction matérialiste du sens. Car il manque à la matière comme présence ou permanence une dimension essentielle du sens ou de l'idéalité, à savoir de valoir pour l'expérience future. Il ne peut y avoir de sens que pour la conscience car seule la conscience possède une temporalité orientée sur le futur comme projet, comme action possible. Le sens n'est pas tant la signification donnée à l'expérience actuelle que la compréhension du monde en vue de l'action future. La marque ne devient proprement une marque ou signe que dans la mesure où sa réapparition est comprise comme une nouvelle occurrence. La répétition identitaire somme les occurrences passées en un type mais dispose également des occurrences futures. C'est en effet en cela que le type est idéal.

La marque matérielle fonctionne donc comme un signe prescrivant son interprétation. Le signe est en quelque sorte l'outil du sens. La sémiotique comme théorie du sens nous apprend que les signes fonctionnent en constituant entre eux un système : la prescription du sens commandé par le signe provient de sa position dans le système. Le signe ne peut donc recevoir n'importe quel sens, ce dernier n'est pas arbitraire. Non pas que la structure matérielle, ses propriétés physiques déterminent sa valeur, mais que le sens donné doit, pour être pertinent, être compatible avec le sens des occurrences passées. En d'autres termes, la structure technique du signe qui prescrit son interprétation n'est pas dans le signe lui-même mais dans le système d'écriture qui l'enrôle et l'investit d'une valeur, c'est-à-dire lui donne une position dans le système.

Toutes les marques matérielles fonctionnent comme les signes d'une écriture[9] et tout ce qui a du sens repose sur la mobilisation de marques matérielles dans l'écriture. En d'autres termes, il faut considérer l'hypothèse que le rapport au sens à travers les supports techniques doit se comprendre à partir d'une analogie avec le phénomène de l'écriture.

3.3.2 Sens interne et inscription corporelle

Mais on peut s'interroger, à propos du langage dit intérieur, sur la matérialité qui lui serait constitutive et qui pourrait constituer une écriture. En effet, nul support physique externe ne semble prêter sa permanence spatiale expliquant ainsi le fait phénoménologique de la conscience faisant que nous sommes capables de penser deux fois la même chose. Mais, selon le même raisonnement que celui nous ayant conduit à sortir de la conscience, il faut également considérer à quelle matérialité ce langage de la conscience peut renvoyer s'il n'est pas externe, au sens d'une extériorité perceptive.

On se souvient en effet comment Husserl disqualifiait le signe matériel comme *indiquant* la signification, que le signe soit acoustique ou écrit, au profit du vécu de la conscience *exprimant* sa signification. Il faut en effet dissocier l'indication de l'expression dans la mesure où, dans la vie psychique solitaire, on rencontre bien l'expression et la signification, mais pas l'indication :

> « Mais *les expressions* déploient aussi leur fonction de signification *dans la vie psychique solitaire* où elles *ne figurent plus comme indices* » RL. I p. [24]

C'est qu'en effet, l'indication se présente fondamentalement comme « le fait que des *objets* ou des *états de choses* quelconques de la réalité desquels quelqu'un a une connaissance *actuelle*, lui indiquent la *réalité de certains autres objets ou états de choses*, en ce sens que la *conviction de l'existence des uns est vécue par lui comme motif entraînant la conviction ou la présomption de l'existence des autres.* » RL. I p. [25]

L'indication est la relation unissant des existences actuelles. Cette relation se rencontre dans les situations de communication, où un signal phonique est émis par un locuteur à destination d'un auditeur. En effet, « on s'aperçoit aussitôt que toutes les expressions fonctionnent dans le discours *communicatif* comme *indices*. Pour l'auditeur, elles font fonction des signe des « pensées » de celui qui lui parle, c'est-à-dire de ses vécus psychiques donateurs de sens, ainsi que des autres vécus psychiques faisant partie de l'intention de communication » p. [33]. Il pourrait sembler que dans la vie psychique solitaire, il y ait une relation semblable d'indication entre le mot évoqué et la signification à laquelle il renvoie ou qu'il désigne. Mais

> « cette désignation n'est pas l'acte d'indiquer au sens où nous en avons traité. L'existence du signe ne motive pas l'existence, ou plus précisément, notre conviction de l'existence de la

9. Un phénomène aussi remarquable que l'écriture n'a pu manqué d'attirer une bibliographie abondante et érudite. Les aspects historiques, anthropologiques et philosophiques sont essentiels pour en entreprendre la compréhension. (Bottéro, 1987) décrit les aléas de la naissance de l'écriture tandis que (Goody, 1979) et (Leroi-Gourhan, 1964) abordent les aspects anthropologiques. Sur le plan philosophique, toute l'œuvre (en particulier (Derrida, 1968)) de J. Derrida est une réflexion sur le phénomène de l'écriture. Sur le plan linguistique, (Auroux, 1995) est une très utile mise au point sur l'impact de l'écriture et l'imprimerie sur la grammatisation comme entreprise technologique. Enfin, sur l'écriture et la technique, il faut consulter (Stiegler, 1994; Stiegler, 1996), déjà cités.

signification. Ce qui doit nous servir d'indice (signe distinctif) doit être perçu par nous comme *existant*. Cela vaut bien pour les expressions dans le discours communicatif, mais non pour les expressions dans le discours solitaire. C'est qu'ici nous nous contentons normalement de mots représentés au lieu de mots réels. Dans notre imagination, un signe verbal parlé ou imprimé est évoqué ; en réalité, il n'« existe » pas du tout. » RL. I p. [35]

Si bien qu'il faut bien distinguer l'expression comme étant le fait que des contenus sont investis d'une signification dans la conscience alors que l'indication est la relation unissant un signe matériel actuel, comme un signal phonique, aux vécus actuels de celui qui émet ce signal. C'est ainsi que le signal sonore engendrera un contenu intuitif perceptif sensible dans la conscience de l'auditeur qu'il devra investir de sens, par des actes donateurs de sens :

« la représentation intuitive dans laquelle se constitue le phénomène-mot physique subit une modification phénoménale essentielle quand son objet prend la valeur d'une *expression*. Pendant que ce qui en elle constitue le phénomène de l'objet demeure inchangé, le caractère intentionnel du vécu se modifie. Ainsi se constitue, sans qu'il doive surgir une quelconque intuition le remplissant ou l'illustrant, un acte de signification qui trouve son appui dans le contenu intuitif de la représentation du mot, mais qui est essentiellement différent de l'intention intuitive dirigée sur le mot lui-même. » RL. I p. [41]

Et c'est ainsi que l'on retrouve le problème du fonctionnalisme sémantique : quel rapport existe-t-il entre le contenu hylétique, contenu intuitif résultant de l'audition de la face physique (phonique) du mot et son investissement de signification par les actes donateurs de sens, qui permettent de viser à travers lui un objet intentionnel ? En quoi l'animation intentionnelle à laquelle donne lieu toute expression dépend-elle de la nature matérielle de l'expression ?

Le contenu positif qu'est le contenu hylétique apparaît à la conscience comme un produit du corps sensible dans son ouverture au monde réel. Si bien que le contenu hylétique apparaît comme étant le support matériel d'inscription de la signification. En remplaçant la notion husserlienne d'expression par celle d'inscription, on voit comment le sens pour l'esprit, sens intentionnel visant un objet, son objet en tant qu'il est visé dans l'acte donateur de signification, investit d'esprit, anime intentionnellement un contenu matériel tout comme l'esprit investit d'esprit les objets matériels et physiques que sont les traces écrites. Les contenus hylétiques sont ainsi l'écriture dans laquelle l'esprit investit la matière de son animation intentionnelle.

C'est donc ainsi qu'il faut réinterpréter à notre sens la dualité matière hylétique *versus* animation intentionnelle de la phénoménologie husserlienne. Les données hylétiques sont matérielles, et à ce titre *ne peuvent être immanentes*, contrairement à ce qu'affirme Husserl. Les données hylétiques ne sont pas immanentes mais transcendantes dans la mesure où elles appartiennent à la corporéité. Ces données sensibles correspondent à la matérialité spatiale que la conscience possède comme première structure de sens. C'est le corps qui conserve l'information perçue, la mémorise, et c'est donc le corps qui assure la conservation spatiale des données de la conscience pour qu'elle puisse se les réapproprier à tout instant pour actualiser les réponses données par l'extériorité à ses attentes.

Le corps propre serait par conséquent le principe matériel et spatial de corrélation entre les vécus de la mémoire. La mémoire aurait pour support d'inscription, comme condition d'accès, la spatialité corporelle. De la même manière que la perception externe consiste en une réception passive du réel, le corps propre serait ainsi le principe d'une réception passive de l'interne. La conscience verrait ainsi défiler le train de pensées dont elle serait en quelque sorte le spectateur non pas involontaire, mais impuissant à contrôler l'orientation. Et de la même manière que la volonté consciente peut entreprendre une action externe, la volonté peut entreprendre une action interne, ce qui s'appelle penser.

On peut donc reprendre ici la discussion menée dans le chapitre précédent à propos de l'éventuel fonctionnalisme sémantique husserlien. Indépendamment du fait de savoir si Husserl penche ou non pour un tel fonctionnalisme, il faut ici s'en démarquer : le contenu hylétique, en tant que marque matérielle corporelle investie d'esprit par l'animation intentionnelle, prescrit cette animation. Non pas que le sens procède empiriquement ou naturellement de la matière hylétique ou sensation : le sens en procède à la manière dont le sens d'un texte procède de la matière écrite du texte. Le contenu hylétique prescrit son interprétation comme une écriture commande la sienne. C'est à notre sens trancher entre l'empirisme naturalisant le sens et perdant alors son idéalité, comme Husserl l'a définitivement montré, et l'immatérialité du sens se heurtant à la finitude rétentionnelle.

On sait par exemple (Auroux, 1995) que les grands corpus de textes sacrés sont mémorisés par les religieux grâce à un apprentissage par cœur reposant sur les rythmes verbomoteurs : les rythmes corporels de récitation, balancements, prosodies, etc., soutiennent la mémorisation et la rendent possible.

Le langage intérieur est ainsi une intériorisation du langage oral qui substitue à l'évanescence du support phonique la permanence du support corporel. Mais le support corporel est une matière vivante, et à ce titre possède une permanence soumise aux changements et maturations inhérentes à la matière vivante. Cela signifie que la matière vivante est un support à la durée de vie courte qui, bien qu'autorisant une grande plasticité de l'inscription, la soumet à un devenir biologique qui compromet par là-même la permanence nécessaire à la mémoire. C'est pourquoi le corps doit s'armer de prothèse mémorielle dont la matérialité est moins sujette au devenir.

La plasticité vivante du corps implique également une plasticité de la constitution du vécu : le rôle constitutif du corps propre n'est pas invariable ni univoque. En effet, phénoménologiquement, le corps est le lieu de l'ouverture et de communication au monde : instance matérielle par laquelle la conscience se projette en dehors d'elle, instance par laquelle elle se constitue comme structure projective d'un renvoi à l'extériorité qui la transcende, le corps instaure une constitution du vécu par essence contextuelle. La constitution du vécu que permet l'inscription corporelle est contextuelle dans la mesure où le corps est par définition la relation au contexte. Cela explique dans une certaine mesure pourquoi la mémoire orale, l'inscription corporelle est à la fois si évanescente et si souple : capable de mémoriser tout contexte, le corps oublie les contextes passés pour appréhender les nouveaux.

3.3.3 Matérialité corporelle et matérialité inanimée

Il faut en effet radicalement distinguer la spatialité de l'inscription corporelle, support *vivant*, de l'inscription matérielle extracorporelle, support *inerte*. En effet, dans la mesure où la matérialité du support est constitutive du sens qui l'anime, sans que ce dernier puisse s'y réduire, dans la mesure où le rapport du sens à la matière n'est pas celui d'une fonction intentionnelle indifférente à son véhicule matériel, les différents types matériels des supports doivent entraîner des différences constatables au niveau du sens transmis. Or, on retrouve bien de telles différences entre la matérialité corporelle et la matérialité extracorporelle, à savoir celles opposant entre autres la *raison graphique* et la *raison orale*.

L'invention de ce qu'on appelle « écriture », c'est-à-dire pour l'inscription sur des supports matériels extracorporels inertes, par opposition à l'inscription sur des supports matériels vécus comme le corps propre, a en effet permis le surgissement ce qu'on peut appeler avec Goody (Goody, 1979) la « raison graphique ». Cette raison repose sur le fait que désormais s'adresse à la conscience non pas le support unilinéaire de la matière sonore du langage oral ou le support plastique de la matière corporelle de la pensée, mais le support bidimensionnel et synoptique de la feuille de papier, ou du parchemin, ou du papyrus, ou de la tablette d'argile. Comme le notent les traducteurs de Goody, l'essentiel, dans l'écriture, n'est pas

> « seulement que l'écriture rende visible, « matérialise », l'articulation inhérente à la langue, c'est aussi que la projection graphique permette d'agencer les significations tout autrement, dans un espace bidimensionnel. La notion saussurienne de linéarité du signifiant est une métaphore graphique dont il faut tirer toutes les conséquences. D'une part, la ligne écrite ne « représente » que fort mal la chaîne temporelle selon laquelle s'ordonnent nécessairement les éléments parlés ; le texte ne connaît pas cette contrainte : qui « lit un livre comme il écoute un discours »? D'autre part, la linéarité graphique est double, autant verticale et horizontale. Et c'est pourquoi Goody centre délibérément son analyse des processus écrits de la connaissance sur ces techniques purement graphiques et totalement dissociées de l'énonciation orale que sont la liste et le tableau. Il ne s'agit pas de simples modes de présentation du savoir, mais bien de matrices formelles qui en déterminent partiellement le contenu. Cette mise en ordre a sa dynamique propre : dans le tableau, chaque élément se voit assigner une place et une seule et il ne doit pas y avoir de case vide. La symétrie impose ses propres effets de pensée : entre les termes mis en colonne, la relation tend à n'être que de contradiction ou d'équivalence. Il y a une raison ou une logique (la Logique?) graphiques. » *in* (Goody, 1979, p. 11)

Spatialisant le langage, l'écriture institue des rapports qui échappent par nature à la temporalité linéaire du discours : elle offre à la pensée un espace synoptique où la conscience puise la matière lui permettant de constituer des significations qui lui resteraient inaccessibles sinon. Cette spatialisation synoptique permet en particulier la grammatisation du langage[10] puisqu'elle permet de juxtaposer dans une présence simultanée sur le support des unités langagières que l'ordre du discours sépare inexorablement. Sans écriture, la notion même de paradigme est inaccessible puisque par définition elle regroupe les voisins *in absentia* : seul le syntagme est accessible dans le discours parlé puisqu'il regroupe les voisins *in praesentia* d'une unité linguistique. Par ailleurs la raison graphique permet de rencontrer des problèmes qui sont inintelligibles dans une culture orale, comme par exemple les problèmes de classification. La classification consiste à catégoriser les éléments d'une liste, c'est-à-dire des éléments prélevés sur la langue parlée. Comme le note Goody,

> « un problème du genre : la tomate est-elle un fruit ou un légume? ne rime absolument à rien dans un contexte oral ; il est même d'un intérêt douteux pour la plupart d'entre nous, mais il peut se révéler décisif quant aux progrès de nos connaissances systématiques concernant la classification et l'évolution des espèces naturelles. C'est ce genre de problèmes qu'engendrent les listes écrites. » (Goody, 1979, p. 187)

10. Comme le montre Auroux (Auroux, 1995).

La spatialisation qu'autorise l'écriture est également la condition de possibilité de la pensée scientifique, dans la mesure où elle permet de constituer non seulement une raison « graghique », mais ce qu'il faudrait peut-être appeler une « raison formelle », c'est-à-dire que la matière signifiante devenant un objet physique stable et permanent pouvant être soumise à des transformations matérielles de tout ordre, elle peut en particulier être soumis à des procédures mécaniques, physiques, de transformations des symboles. L'écriture scientifique se manifeste par le fait qu'elle tend, à défaut d'y réussir, à formaliser ses contenus en imposant une transformation purement mécanique, algorithmique, des signes écrits. Celle-ci tend à devenir une écriture automatique qui s'autointerprète sans que la conscience n'ait à intervenir. Cela aboutit en fait à un monstre phénoménologique : une constitution sans conscience [11] . On sait que, depuis l'échec du programme de Hilbert [12], le formalisme n'est pas une description adéquate de l'activité mathématique. Il n'en demeure pas moins que le formalisme traduit la tendance naturelle de la pensée scientifique et que, si elle en constitue un horizon inatteignable, elle n'en est pas moins constitutive de son activité.

Il faut donc opposer nettement l'écriture à la parole comme deux modes radicalement distincts d'inscription. Anthropologiquement, le premier type de matérialité autre que le corps qui s'est prêté à l'inscription du sens fut de nature idéographique et pictographique et non phonétique. L'inscription matérielle inerte ne fut pas mise en place pour fixer la parole, mais pour inaugurer une nouvelle rationalité. L'inscription et la fixation dans la matière de symboles, ne redoublent pas le langage : l'idéographie ou la pictographie ne se parlent pas. Ce ne sont pas des écritures phonétiques, des écritures renvoyant à la parole, mais des « mythogrammes » qui renvoient à un ensemble d'images ou contenus réveillés par les multiples associations constituant le mythogramme comme signe ou inscription.

> « Tout d'abord le pictogramme, outre l'objet premier qu'il « dépeint », peut se rapporter à d'autres réalités, rattachées à ce même objet par des procédures mentales plus ou moins fondées en réalité , ou carrément conventionnelles, et parfaitement connues et utilisées dans les représentations de l'art : le pictogramme de la montagne évoquera aussi les pays étrangers qui, en Mésopotamie, sont délimités à l'est et au nord par des chaînes montagneuses ; l'épi rappellera tout le travail agricole. » (Bottéro, 1987, p. 100)

Ce n'est que par la suite que l'écriture se fait unilinéaire et phonétique. Il ne s'agit plus pour elle de renvoyer à des contenus vécus directement, mais de passer par la médiation du langage. Le langage est en effet un support bien plus précis et efficace de réactivation du vécus et ce qu'une pictographie ou idéographie réussissait imparfaitement, l'écriture phonétique l'atteint pleinement.

> « Pour qu'elle devint capable, non seulement de fixer les souvenirs, mais de communiquer la pensée, non seulement de remémorer du connu mais d'enseigner de l'inconnu, il suffisait de raccrocher l'écriture à ce qui constitue l'instrument le plus parfait de l'analyse et de la communication du réel par l'homme: la langue parlée. » (Bottéro, 1987, p. 103)

11. Ce monstre phénoménologique est sans doute une voie d'approfondissement pour les sciences cognitives puisqu'il consiste à envisager la question du sens en se démarquant des difficultés inhérentes à la notion de sujet. (Dupuy, 1994) insiste sur ce point et souligne ainsi le parallèle à première vue surprenant entre cybernétique et psychanalyse lacanienne.

12. Cf. notre introduction.

Il est frappant de constater comment l'écriture scientifique rompt avec l'écriture phonétique dans la mesure où, ne se parlant pas, elle exprime une autre rationalité que celle qui se fond sur la discursivité de la parole : jouant pleinement les jeux synoptiques de l'écriture sur le plan bi-dimensionnel de la feuille de papier, l'écriture scientifique permet de constituer un sens inaccessible à la pensée ressortissant seulement à la raison orale.

L'écriture scientifique renoue dans une certaine mesure avec les mythogrammes de la pensée appartenant aux civilisations pré-alphabétiques. Dans une certaine mesure seulement, puisque, quand les pictogrammes sont devenus idéogrammes, ce fut pour désigner une galaxie de notions en rapports étroits entre elles, galaxie relevant davantage de l'association que du rapport conceptuel, au lieu de seulement désigner un objet simple à l'instar de la pictographie (Bottéro, 1987). L'idéographie non scientifique est donc floue et imprécise, et en cela elle est une écriture imparfaite. En revanche, l'idéographie scientifique s'est constituée en visant un idéal d'exactitude, de précision, et d'univocité[13]. Il est en effet frappant de constater comment l'idéographie scientifique s'est constituée comme une écriture pleinement graphique et pleinement scripturaire : en rupture avec l'oral (donc pleinement graphique) pour assurer l'univocité du sens constitué mais avec toute la précision qu'une écriture alphabétique pouvait apporter. Mais au lieu d'avoir un alphabet des phonèmes, on avait un alpabet des pensées.

La raison orale s'oppose à la raison graphique de manière profonde. La différence est double : au niveau de la modalité sensorielle privilégiée, et au niveau du support d'inscription. En faisant appel à la vue plutôt qu'à l'ouïe, la raison graphique donne à voir des rapports spatiaux inaccessibles à la raison orale qui reste soumise au déroulement irréversible du flux temporel. L'enregistrement sonore de la voix ne permet pas de comprendre ou de saisir la même chose que sa transcription graphique. Comme on l'a rappelé, la spatialisation permet de constituer de nouveau rapport dans la mesure où elle délinéarise le message parlé. Mais la principale différence provient de ce que la raison graphique construit sa rationalité à partir d'un support externe matériel permanent : elle ne s'en remet pas au souvenir d'avoir vu, mais confie la charge de cette mémorisation au support externe. La raison orale ne s'appuie pas sur une telle médiation de la matérialité spatiale si bien que les seules traces à partir desquelles elle peut se constituer sont les traces sensibles dont le corps est le support d'inscription.

Il faut donc distinguer la raison graphique et la raison orale d'une part du fait que la vue offre et permet de saisir d'autres rapports que le son, et d'autre part parce que la raison graphique repose sur le fait que ce qu'elle voit est une matérialité permanente, qui dure, alors que la raison orale se heurte à l'évanescence du support phonique.

Tant la raison graphique que la raison orale passent par la médiation du corps pour se constituer : les données hylétiques de la vision et de l'audition. Mais les données hylétiques de la vision ne constituent pas le support de la mémorisation : cette dernière repose sur une inscription matérielle externe permanente. La permanence du support matériel externe fait en revanche défaut à la raison orale si bien que ce sont les données hylétiques de l'audition qui doivent constituer le support de la mémorisation. Par conséquent, c'est le corps qui constitue le support d'inscription de la raison orale, alors que c'est le support extérieur permanent qui constitue le support d'inscription de la raison graphique.

Le corps comme support matériel d'inscription est un corps vivant, un corps vécu. Il

13. Le projet de la *Begriffsschrift* frégéenne (Frege, 1971a) est une claire illustration de ce fait.

constitue une personne, un sujet, quand s'y ajoute la conscience qui l'investit d'esprit. Mais cela implique le caractère personnel et privé des inscriptions personnelles qui dès lors ne sont plus partageables ou communicables. Ne véhiculant du sens que pour la conscience les animant intentionnellement, les marques corporelles ou données hylétiques sont inaccessibles en tant que telles à autrui. Pour partager le sens, il faut donc le communiquer : pour connaître le sens pour autrui, dans la mesure où il ne peut accéder à ma mémoire corporelle, il faut que je la déchiffre pour lui et lui en communique le résultat via la matérialité du langage et les significations de la langue. Face à cette communication, autrui se saisit de cette matérialité pour constituer son propre sens et la mémoriser dans sa matérialité corporelle.

En d'autres termes, la raison orale repose sur une communication de sujet à sujet : ce qui n'a un sens que pour un sujet personnel se transmet via une matérialité qui doit se constituer comme un sens pour un autre sujet personnel. La communication orale est donc une intersubjectivité[14].

Il en est tout autrement pour la raison graphique. La raison graphique repose *in fine* sur l'inscription du sens dans un support matériel externe et permanent, inaltérable. En tant qu'objet matériel externe permanent, le support de la mémorisation de la raison graphique s'offre à tous. Chacun accède à ce support et peut l'investir d'une signification. L'investissement de la signification n'est pas quelconque, mais commandé par la matérialité du support : l'inscription commande son sens. Si bien que chacun constitue une signification identique à partir d'une même inscription. Si bien que contrairement à la raison orale, la raison graphique communique le support d'inscription lui-même et, par là-même, l'inscription, dans la mesure où il est accessible à tous, et chacun utilise cette matérialité pour constituer le même sens. Dans la raison orale, le support d'inscription est corporel donc privé, et ce qui est communiqué, transmis, c'est le sens et non son inscription.

La communication dans la raison graphique repose par conséquent sur un partage de la mémoire externe : il ne s'agit pas d'une communication du sens constitué, mais d'une communication du principe matériel de sa constitution. Il ne s'agit donc pas d'une communication directe de sujet à sujet, mais d'une communication indirect reposant sur la médiation d'un objet externe partagé. C'est pourquoi la raison graphique est une communication reposant sur une *objectivité partagée* alors que la raison orale est une *communication partagée*. Dans la raison graphique, les hommes se parlent à partir des objets qu'ils désignent et se transmettent, dans la raison orale, ils se parlent directement.

Il est facile de comprendre, dans ces conditions, que la raison graphique donne le principe de la possibilité du sens univoque et nécessaire de la science dans la mesure où se transmettent des mémorisations externes imposant la manière dont elles doivent être lues (comme un outil prescrit la manière de l'utiliser) et prescrivant ainsi la constitution univoque du sens. Dans la mesure où c'est la mémoire externe qui prescrit son sens, ce sens n'est plus dépendant de qui le constitue ni des circonstances et du contexte dans lequel il le fait. Le sens devient alors *objectif* puisqu'il est partagé par tous. Mais dans la mesure même où il est partagé par tous, où chacun peut constater cette objectivité, le sens se montre sous la figure exemplaire de l'idéalité.

14. L'intersubjectivité est un problème immense, particulièrement dans la tradition phénoménologique (voir notamment les *Méditations cartésiennes* (Husserl, 1928) et les *Ideen II* (Husserl, 1982). Il n'est pas question pour nous ici d'aborder ce problème autrement que pour les besoins de notre discussion, à savoir l'opposition de la raison orale à la raison graphique, du corps propre, privé, à l'objet extracorporel, public.

La mémorisation externe permettant la constitution d'un sens unique et univoque, il semble s'autonomiser par rapport à ses conditions matérielles de constitution. Procédant de la matière, l'univocité de la constitution objective du sens à partir des mémoires externes l'en éloigne dans la mesure où le sens étant le même dans tous les cas, il n'est réductible à aucun. C'est sans doute la raison pour laquelle le sens idéal est considéré comme transcendant la matière et y est irréductible.

Il en est autrement pour la raison orale puisqu'elle consiste dans la communication d'un sens constitué et non du principe matériel de sa constitution. Cela signifie que l'inscription corporelle est particulière à chaque fois et qu'elle ne peut donc assurer l'universalité et l'objectivité du sens communiqué. L'inscription corporelle mémorisant le message reçu à partir de son sens transmis dépend des conditions contextuelles et circonstancielles accompagnant l'actualisation du sens transmis.

L'opposition entre raison graphique et raison orale semble être en mesure d'expliquer la différence de statut entre les connaissances scientifiques, exprimées à travers les langages formels (logique et mathématiques) de la science et les connaissances exprimées en langue naturelle, opposition que nous avons appelée différence phénoménologique (Bachimont, 1994). D'une certaine manière, l'IA, en tant que projet technologique, n'existe que pour cette dissymétrie : c'est parce qu'il est difficile d'opérationnaliser les connaissances exprimées en langue qu'un tel projet suffit à constituer une problématique autonome et une discipline scientifique (au sens large) et technologique. La raison graphique permet en effet la constitution du sens scientifique, impossible dans le cadre de la seule raison orale. Par ailleurs, les connaissances phénoménologiques correspondent à des connaissances exprimées en langue naturelle, c'est-à-dire à des connaissances dont l'écriture reflète la parole et la raison orale. Les connaissances scientifiques s'écrivent mais ne se parlent pas, les connaissances phénoménologiques se parlent et leur écriture ne fait que transcrire leur expression orale.

En résumé, nos propositions sont les suivantes. La raison orale est une constitution intersubjective du sens, alors que la raison graphique en est la constitution objective. La différence phénoménologique qui en surgit oppose donc l'intersubjectivité à l'objectivité : ce dont on *parle* les uns avec les autres n'est pas identique à ce que l'on constitue quand on écrit dans la matière inanimée et permanente. L'objet du discours est un accord intersubjectif qui est contextuel et consensuel ; l'objet de l'écrit est un accord objectif qui est nécessaire et universel car imposé par l'ordre même de ce qui est écrit, dans sa matérialité scripturaire. L'accord entre les sujets que permet la raison graphique procède de l'objectivité du sens qu'elle constitue. L'accord entre les sujets que permet la raison orale procède de l'intersubjectivité du sens qu'elle communique. Dans la raison orale, l'objectivité procède de l'intersubjectivité, dans la raison graphique, l'intersubjectivité procède de l'objectivité.

4 L'herméneutique du sens

Nous venons de voir que l'extériorité sémiotique permettant de constituer les connaissances pouvait se comprendre sur le modèle de l'objet technique, et plus précisément encore selon le modèle de l'écriture : l'objet, comme une trace écrite, prescrit par sa structure matérielle la manière de le « lire ». Mais si l'objet prescrit sa lecture, il ne la détermine pas.

Autrement dit, plusieurs lectures sont possibles, même si l'objet technique, en tant qu'il est technique, prescrit à l'avance que certaines lectures sont pertinentes et d'autres non.

Mais, si l'on poursuit l'analogie de l'écriture et de la lecture, la question se pose si la pluralité des lectures possibles ne correspond pas au fait qu'aucune lecture ne permet d'épuiser le sens de l'objet, qu'aucune lecture ne puisse être définitive. En d'autres termes, il n'existe pas quelque chose comme *la* lecture d'une trace écrite, ni quelque chose comme *le* bon usage d'un objet technique, même s'il existe de mauvaises lectures et de mauvais usages.

D'une certaine manière, concevoir l'extériorité sémiotique sur le modèle de l'écriture, c'est-à-dire du signe inscrit, introduit le thème que la matérialité du support et sa structure permet constamment une relance interprétative : dans la mesure où le sens est prescrit et non déterminé, il est toujours possible de reprendre à nouveaux frais la question de l'interprétation, de la compréhension, de la lecture. Il faut donc pratiquer un matérialisme non réductionniste dans la mesure où la matérialité signifiante du signe est le principe de la constitution du sens sans pouvoir en être l'auteur : condition de possibilité du sens, elle ne peut s'y égaler.

Il est tentant, arrivé à ce point de notre discussion, d'adopter le vocabulaire de la sémiotique de Peirce pour approfondir la notion de relance interprétative que nous venons d'évoquer. La sémiotique de Peirce[15] est une phénoménologie centrée non comme Husserl sur les significations, mais sur les signes eux-mêmes, tels qu'une intelligence scientifique les manipule. La phénoménologie de Peirce est une phanéroscopie qui s'attache à entreprendre une théorie descriptive de l'esprit :

> «La phanéroscopie est la description du phaneron; par phaneron, j'entends la totalité collective de tout ce qui, de quelque manière et en quelque sens que ce soit, est présent à l'esprit, sans considérer aucunement si cela correspond à quelque chose de réel ou non. Si vous demandez : présent quand et à l'esprit de qui, je réponds que je laisse ces questions sans réponse, n'ayant jamais eu le moindre doute que ces traits du phaneron que j'ai trouvés dans mon esprit soient présents de tout temps et à tous les esprits. La science de la phanéroscopie telle que je l'ai développée jusqu'ici s'occupe des éléments formels du phaneron.» (Peirce, 1960, 1.284)

Et encore :

> «Il apparaîtra clairement de ce qui a été dit que la phanéroscopie ne se rapporte pas du tout à la question de savoir dans quelle mesure les phanerons qu'elle étudie correspondent à des réalités. Elle s'abstient religieusement de toute spéculation concernant les relations que pourraient entretenir ses catégories avec les faits physiologiques, cérébraux et autres. Elle n'entreprend pas, mais évite au contraire avec soin, de donner des explications hypothétiques de quelque sorte que ce soit. Elle scrute simplement les apparences directes et essaie de combiner la précision du détail avec la généralisation la plus large possible.» (Peirce, 1960, 1.287)

Malgré cette communauté d'intérêt pour une théorie descriptive de l'esprit, la phanéroscopie de Peirce s'oppose radicalement à la phénoménologie husserlienne dans la mesure où elle récuse fondamentalement la méthode intuitive. Pour lui, toute investigation sur le

15. voir notre note bibliographique du chapitre 5.

phanéron doit être menée à partir des faits externes, et non à partir des faits internes. Au lieu d'effectuer une suspension du monde externe, Peirce suspend le monde interne pour ne lui reconnaître *in fine* aucune prétention à être le lieu de la certitude absolue :

> s «Il apparaît donc, qu'il n'y a aucune raison de supposer une faculté d'introspection ; et que, par conséquent, la seule façon d'étudier une question psychologique, c'est par inférence à partir de faits externes.» (Peirce, 1984, p.189)

L'étude du phanéron sans recourir à l'introspection conduit Peirce à définir toute pensée comme signe et toute opération de l'esprit comme inférence : signe et inférence possèdent cette communicabilité propre à tout procédé discursif qui fait défaut à l'intuition husserlienne. Alors que la phénoménologie husserlienne est une phénoménologie de l'intuition et de la signification, la phanéroscopie est une phénoménologie de l'inférence et du signe. A ce titre, elle sera une sémiotique [16]. Ainsi, toute connaissance, en tant que connaissance et seulement en tant que cela, est un signe.

Un signe ou representamen renvoie à un objet grâce à un autre signe, son interprétant, qui interprète le representamen dans la mesure où il indique en quel sens il faut le prendre pour déterminer quel est son objet. Le signe et son interprétant renvoient au même objet dans la mesure où l'interprétant représente le representamen dans sa relation à son objet. C'est même la raison pour laquelle il lui est possible de l'interpréter, de l'expliciter en «un signe plus développé» :

> « Un signe, ou representamen, est quelque chose qui tient lieu pour quelqu'un de quelque chose sous quelque rapport ou à quelque titre. Il s'adresse à quelqu'un, c'est-à-dire crée dans l'esprit de cette personne un signe équivalent ou peut-être un signe plus développé. Ce signe qu'il crée, je l'appelle l'interprétant du premier signe. Ce signe tient lieu de quelque chose : de son objet. Il tient lieu de cet objet, non sous tous rapports, mais par référence à une sorte d'idée que j'ai appelée quelque fois le fondement du representamen. » (Peirce, 1960, 2.228)

L'interprétant est la médiation nécessaire pour comprendre le signe en tant que signe, il lui confère la tiercéité [17] qui le distingue d'une simple qualité et d'une «brutale» interaction en un premier objet appelé «signe» et un second appelé «objet du signe» :

> « Il est important de comprendre ce que j'entends par sémiosis. Toute action dynamique, ou action de la force brutale, physique ou psychique, ou bien s'exerce entre deux sujets ou bien est en tout cas la résultante d'actions entre paires. Mais par «sémiosis», j'entends, au

17. Pour Peirce, la tiércéité constitue avec la priméité et la secondéité les trois catégories fondamentales de l'Etre : «Le premier est ce dont l'être est simplement en soi ; il ne renvoie à rien et n'est impliqué par rien. Le second est ce qui est ce qu'il est en vertu de quelque chose, par rapport à quoi il est second. Le troisième est ce qui est ce qu'il est par les choses entre lesquelles il établit un lien et qu'il met en relation.» (1.356). Le troisième ou tiercéité est la catégorie de la connaissance et de l'intelligence. La connaissance est en effet une médiation entre un sujet et un objet, elle est troisième pour un premier vis-à-vis d'un second. La connaissance se reconnaît quand il faut postuler un moyen terme entre un premier terme et un second, et ce moyen terme est la preuve qu'il y a intelligence et esprit : «Mais maintenant, lorsqu'un biologiste examine au microscope un animalcule et se demande si les mouvements de cet animalcule sont ceux d'un être intelligent – fût-ce du degré le plus bas de l'intelligence –, le test qu'il utilisait quand j'allais à l'école et qu'il utilise toujours, je suppose, consiste à s'assurer qu'un événement A produit un second événement B comme moyen de produire un événement C.» (5.473).

contraire, une action ou influence qui est ou implique la coopération de trois sujets, tels qu'un signe, son objet et son interprétant, cette influence tri-relative n'étant en aucune façon réductible à des actions entre paires.» (Peirce, 1960, 5.484)

Contrairement à Saussure, un signe n'est pas signe en vertu d'une relation dyadique signe - signifié, mais en vertu d'une relation triadique signe - objet - interprétant. Ce qu'il faut noter, c'est que les trois instances nécessaires pour faire un signe, c'est-à-dire une médiation, sont elles-mêmes des signes. Que le representamen soit un signe, voilà qui semble facile à accorder. Que l'objet du signe en soit un également, cela semble en revanche plus difficile à comprendre. Pourtant Peirce affirme :

«Tout signe est mis pour un objet indépendant de lui-même ; mais il ne peut être signe de cet objet que dans la mesure où cet objet a lui-même la nature d'un signe, de la pensée. Car le signe n'affecte pas l'objet, mais en est affecté, de sorte que l'objet doit être capable de communiquer la pensée, c'est-à-dire qu'il doit avoir la nature de la pensée ou d'un signe. Toute pensée est signe.» (Peirce, 1960, 1.538)

Enfin, l'interprétant est un signe. Peirce note que l'interprétant immédiat correspond à ce que l'on appelle habituellement signification. De manière générale, l'interprétant est «l'effet signifié» du signe ou representamen. Cet effet signifié peut être affectif (les sentiments éveillés par l'écoute d'une symphonie), énergétique (action pour obéir à un ordre) ou logique (le concept). Contrairement aux deux premiers, le dernier correspond à l'interprétant «authentique» d'un signe ou concept. Or, cet interprétant logique est un signe : pensée suscitée par le signe pour le comprendre, il ne peut être que signe car toute pensée est signe. Mais si l'interprétant étant lui-même un signe, il doit posséder son propre interprétant, lui aussi un signe possédant son interprétant, etc. Il existe donc un renvoi infini d'interprétant en interprétant :

«[En bref, un signe est] tout ce qui détermine quelque chose d'autre (son interprétant) à renvoyer à un objet auquel lui-même renvoie (son objet) de la même manière, l'interprétant devenant à son tour un signe et ainsi de suite ad infinitum.» (2.203)

Il existe donc une interprétatabilité infinie du signe, son interprétation n'est jamais terminée. C'est pourquoi l'on peut parler avec Eco (Eco, 1985) de sémiosis illimitée, d'interprétation infinie du signe. Dans cette sémiosis, le sens, la signification, se manifeste comme étant la chaîne illimitée des réécritures des signes, mais n'est pas thématisé comme tel.

4.1 Sémiotique et signification

On pourrait croire que le sens est purement éliminé, abandonné comme une vieille fantasmagorie subjective désormais inutile. La sémiotique serait une syntaxe à laquelle on aurait supprimé sa sémantique. S'il est bien vrai que le signifié n'est pas un concept

17. La sémiotique est la théorie générale des signes. La contribution de Peirce à cette discipline est fondamentale. On pourra consulter (Peirce, 1978) et (Everaert-Desmedt, 1990) en introduction. La sémiotique de Peirce a été abondamment étudiée, notamment par Eco qui lui a consacré des pages importantes dans (Eco, 1985), (Eco, 1988a) et (Eco, 1992).

théorique de la sémiotique peircienne, il n'en demeure pas moins que la signification en reste un thème central. Celle-ci n'est pas un retour à l'attitude analytique où le signe est dépossédé de la sémantique car c'est le monde de l'expérimentation qui devient la seule source de vérité et de signification. Si la sémiotique ne se concentre que sur les signes, c'est parce que ces derniers s'inscrivent dans un cadre sémantique.

> «La formalisation du signe n'est donc pas synonyme chez Ockham, pas plus que chez Peirce d'une réduction de la sémantique à la syntaxe : elle veut uniquement dire une mise à l'écart et chez Peirce en tout cas, seulement provisoire de la signification du signe, bref de ce que le signe prétendrait vouloir dire sur le monde ou sur nous, indépendamment de sa pure et simple mise en œuvre comme signe.» (Tiercelin, 1993, p.98)

Cela signifie que la signification ne doit pas être recherchée dans un monde indépendant du signe, mais dans le fonctionnement même du signe comme signe. Or, le fonctionnement du signe consiste dans la réécriture interprétative de signes en signes. Par conséquent, la signification d'un signe se manifeste dans les interprétations qu'il autorise, les réécritures qui l'explicitent: «le sens d'un signe est le signe dans lequel il doit être traduit» (4.32), c'est-à-dire «la traduction d'un signe en un autre système de signes» (4.127).

Le signe et son interprétant ne sont pas des entités syntaxiques dans la mesure où l'interprétant indique à une «intelligence scientifique» [18] ou un «quasi-esprit» comment il faut entendre le signe et l'articuler à son objet. C'est donc la signification du signe qui commande son interprétation même si elle ne se manifeste qu'à travers la réécriture du signe en interprétant. La signification est donc l'instance régulatrice de la réécriture. Elle n'est pas accessible en tant que telle, mais seulement à travers les réécritures. Comme dit Deledalle,

> «Un signe a une signification qui ne peut être saisie que par le moyen de la signification d'un autre signe qui est le signe interprétant. Ce dernier a une signification comme le signe primitif, mais on ne saisit la signification du signe primitif que parce qu'on saisit la signification du signe interprétant.» (Peirce, 1978, p.222)

La signification instaure un horizon interprétatif dans lequel s'inscrivent les réécritures du signe en ses interprétants. Cette conception sémiotique de la signification lui confère alors trois caractéristiques principales étroitement liées les unes aux autres : le renvoi à l'infini de l'interprétation explicitante du sens, la familiarité du sens, son inaccessibilité. Le renvoi à l'infini de l'interprétation découle de la sémiosis illimitée. En effet, la signification ne peut être atteinte que de façon médiate, via l'interprétation. Le renvoi du signe à son signifié est donc impossible directement, il ne peut exister pour Peirce de sémantique des significations. Il faut pour cela introduire un troisième terme ; mais ce troisième ne met pas un terme au parcours qui va du signe à son sens dans la mesure où il reconduit pour lui-même la médiation qu'il apporte au premier signe.

18. L'intelligence scientifique est un concept typiquement peircien qui renvoie à la manipulation rationnelle et raisonnée des signes. L'intelligence sicentifique est une idéalisation correspondant à la communauté idéale savante : « L'« intelligence scientifique » apparaît alors comme une idéalisation de la communauté réelle des chercheurs, un agent collectif purement rationnel engagé dans la poursuite illimitée de la vérité, qui tend à être en fait le véritable « sujet », quoique ni personnel, ni individuel, du savoir. » (Chauviré, 1995, p. 53).

En d'autres termes, nous comprenons le signe : il nous est familier, proche de nous. Mais, quand il faut accéder effectivement au sens du signe, quand il faut dépasser le sentiment de compréhension qui gouverne notre utilisation du signe, nous nous apercevons qu'en guise de signification, nous n'avons alors que d'autres signes interprétants, explicitant la compréhension du signe étudié. Au sentiment de compréhension motivant un accès direct au sens, à la familiarité du sens, se substitue alors un désaisissement où le sens semble résister à toutes les interprétations et explicitations proposées, où il ne se laisse pas réduire à une expression particulière pour laquelle on pourrait dire: «voilà ce que veut dire ce signe». Pour plagier Saint Augustin, il en est du signe et du sens comme du temps:

> «Qu'est-ce donc que le temps [sens] ? Si personne ne me le demande, je le sais ; mais si on me le demande et que je veuille l'expliquer, je ne le sais plus». *Confessions*, chap IX, (Saint Augustin, 1964)

Le sens et la signification constituent cette familiarité par laquelle tout signe se laisse comprendre. Cependant, cette signification est inaccessible et inassignable : évidente dans les réécritures, elle se dérobe dès que l'on veut l'atteindre pour elle-même. Car pour Peirce, si la pensée ne se réduit pas à ses expressions, elle n'existe pas sans elles:

> «Une pensée identique peut être véhiculée par l'anglais, l'allemand, le grec ou le gallois ; en des diagrammes, en des équations ou en des graphes : ce ne sont tous qu'autant de peaux d'oignon, ses accidents inessentiels. Cependant, que la pensée doive avoir quelque expression possible pour quelque interprète possible, c'est l'être même de son être...» (4.6)

Cela s'oppose au motif husserlien qui, au moyen de l'intuition évidente, peut envisager un univers de signification indépendamment de l'expression:

> « Il y a, par conséquent, d'innombrables significations qui, au sens relatif habituel du mot, sont des significations simplement possibles, car elles ne parviennent jamais à l'expression et, en raison des limites de la faculté humaine de connaître, elles ne pourront jamais parvenir à l'expression. » RL. I p. [105] (Husserl, 1959)

Cette opposition dans la conception des rapports expression/ signification entre sémantique phénoménologique et sémiotique phanéroscopique nous amène à opposer ce que nous appellerons signifié transcendantal et signifié herméneutique. On peut dire que la recherche du sens et de la signification consiste dans la recherche des conditions d'utilisation des signes, ces conditions constituant le sens du signe. A partir du fait de l'usage du signe, on en recherche le droit en explicitant les règles conditionnant cet usage. Une telle recherche est classiquement le motif transcendantal. Le signifié du signe peut donc être qualifié de signifié transcendantal dans la mesure où il conditionne a priori l'usage du signe. Une grammaire pure des significations considère que le signifié transcendantal est accessible en tant que tel (sous le regard de l'intuition) et qu'il est possible d'en faire la grammaire indépendamment des signes qui les expriment, indépendamment des expressions qui ne constituent qu'un relais secondaire du sens. Une grammaire spéculative au sens de Peirce, en se reportant sur les signes, s'interdit comme nous venons de le voir l'accès direct au signifié transcendantal. Le transcendantal est invisible et inaccessible chez Peirce, alors que pour Husserl il est possible de l'exhiber. Le transcendantal devient donc la condition, la règle invisible conditionnant l'usage du signe. Son invisibilité consiste dans le fait que l'on ne peut l'envisager que par la médiation d'un interprétant qui est du niveau du signe et non du niveau du signifié. Par conséquent, si la grammaire du signe reflète la structure du sens, ce dernier ne se livre qu'à travers l'infini de l'interprétation. Cela implique qu'en passant de la sémantique à la sémiotique, on passe d'un signifié transcendantal à un autre type de signifié que nous appelons signifié herméneutique.

4.2 Sémiotique et herméneutique

Car en effet, sémiosis illimitée, familiarité et désaississement, tout cela évoque l'herméneutique[19]. L'herméneutique peut se caractériser comme l'entreprise consistant à élaborer des réponses à une question, dans la mesure où la question transcende constamment le niveau où se définit la réponse. Cet écart permanent entre la question et sa réponse, le défaut nécessaire et inéluctable de la réponse, font que la question doit toujours être reprise à nouveau frais. C'est l'infini de l'interprétation comme réponse à une question. Cette caractérisation repose sur le fait que comprendre quelque chose, c'est comprendre à quelle question ce quelque chose doit répondre. Si l'on suit Gadamer (1960), aborder un texte, pour nous, un signe, c'est retrouver la question à laquelle il répond :

> «La reconstitution de la question dont un texte donné représente une réponse ne peut, cependant être comprise comme réalisable sur la seule base de la méthodologie historique. [...] La chose transmise qui s'adresse à nous – texte, œuvre, vestige – pose elle-même une question, et par là met notre esprit dans une situation d'ouverture. Pour répondre à la question qui nous est posée, il faut que nous, à qui elle est posée, commencions nous-mêmes à questionner. Nous cherchons à reconstituer la question à laquelle répondrait la chose transmise. Mais nous ne pourrons absolument pas le faire sans dépasser par notre question l'horizon historique ainsi dessiné. La reconstitution de la question à laquelle le test doit répondre est elle-même comprise dans une interrogation plus vaste grâce à laquelle nous cherchons la réponse à la question qui nous est posée par la tradition historique. Une question, en tant qu'elle est reconstituée, ne peut jamais rester comprise dans un horizon initial. Car l'horizon historique dessiné dans la reconstitution n'est pas véritablement un horizon qui circonscrit. Il est à son tour compris dans l'horizon qui nous englobe, nous qui questionnons et sommes interpelés par la voix du passé.» (Gadamer, 1960, p.221)

Retrouver la question, c'est être capable de la comprendre, de comprendre en quoi le signe en est une réponse possible. Il faut donc une fusion d'horizon entre l'horizon dans

19. L'herméneutique est la théorie de l'interprétation des textes. Antant dire que la bibiographie est nourrie. (Grondin, 1993) propose une introduction utile aux problématiques herméneutiques ainsi qu'aux différentes traditions qui l'animent. C'est principalement à Dilthey (Dilthey, 1947) que l'on doit l'adoption des théories herméneutiques de l'interprétation comme paradigme d'étude des sciences humaines. Mais c'est avec Heidegger que l'herméneutique devient philosophique et existentiale dans *Etre et Temps* (Heidegger, 1986). L'herméneutique philosophique est essentiellement représentée aujourd'hui par Paul Ricœur et Hans-Georg Gadamer. (Ricoeur, 1969) et (Ricoeur, 1986) contiennent des études essentielles pour comprendre le projet herméneutique et sa mise en œuvre philosophique tandis que (Ricoeur, 1965) le confronte au paradigme psychanalytique et que (Ricoeur, 1985) fait une synthèse magistrale des problématiques phénoménologiques de l'interprétation des textes. (Gadamer, 1960) s'inscrit dans la filiation heidegerrienne pour l'élargir vers une réflexion approfondie sur le rôle de la tradition dans l'activité interprétative. Enfin, notons les tentatives récentes d'étendre le paradigme interprétatif vers les sciences formelles, rompant son application exclusive aux sciences humaines avec (Salanskis, 1991). Sur le plan linguistique, (Rastier *et al.*, 1994) milite pour une herméneutique matérielle dans la mesure où l'interprétation des textes doit s'ouvrir sur la dimension philologique.

lequel le donné à interpréter fut produit comme réponse à une question et l'horizon dans lequel nous réactualisons la question pour entreprendre notre compréhension du donné. Il faut donc que la question, pour susciter cette fusion d'horizon, transcende les réponses, celle du texte et la nôtre. Toute compréhension est réponse à une question, et toute question transcende ses réponses. C'est pourquoi comprendre, interpréter, ce n'est pas revivre ou retrouver le sens précis du texte à partir d'une question dont il épuise les possibilités, c'est innover par une nouvelle compréhension de la question pour y inscrire la réponse déjà donnée dans le texte. Cela signifie que la question n'est pas une essence, (un signifié transcendantal), qui définirait le tout du texte. C'est plutôt une interrogation, un mystère, qui se voit attribuer une essence dans les réponses qu'on lui donne, mais qui en propre n'en a aucune. Que les réponses données soient néanmoins commensurables, et que la fusion soit toujours possible provient du fait que la question est une familiarité entretenue avec un problème, un mystère qui, dès qu'on le considère, se voit attribuer une essence par le discours qui l'aborde. Mais, ces essences n'épuisent pas le mystère, elles n'en sont pas l'essence :

> «Dans cette lecture, ce qui est articulé, verbalisé au second moment du processus herméneutique n'est pas l'essence de ce qui a été projeté, mais un déplacement dans l'énigme.» (Salanskis, 1991, p.7)

Ces essences constituent des explicitations de la question, s'explicitent et s'interprètent les unes les autres avec comme horizon d'interprétation et de compréhension la question. C'est pourquoi, plutôt que de parler de question, nous préférons avec J.-M. Salanskis la notion de tenant de question, qui marque bien le caractère de retrait et de mystère de ce qui pose la question sans être pour autant un quelque chose. Cette perspective herméneutique rapportée à notre problème, cela signifie que les signes se déploient conformément à une question : ils s'interprètent les uns les autres, chacun donnant sa version du sens de l'autre. Ils constituent un parcours interprétatif, où chacun tente d'élaborer dans quelle mesure le signe qu'il explicite répond au tenant de question constituant l'horizon d'interprétation. L'interprétant interprète le signe à l'aune de sa réponse au tenant de question. Ce dernier constitue donc un contexte herméneutique d'explicitation des signes.

Il faut donc distinguer plusieurs types de plurivocité d'un signe. Un premier niveau correspond à la sémiosis illimitée dans le cadre d'un même contexte herméneutique. Un second niveau correspond aux différentes sémiosis qu'un signe peut susciter selon le contexte herméneutique choisi. Alors que la première constitue davantage une indétermination qu'une équivoque, la seconde constitue une fracture profonde dans la mesure où choisir un contexte herméneutique, c'est choisir un sens. Cela signifie donc que la sémiotique peut rendre compte de la différence phénoménologique si l'on sait assigner des contextes herméneutiques différents aux termes de la différence. Les langues naturelles et les langages artificiels se distinguent par leur parcours interprétatifs, et correspondent à des herméneutiques de nature différentes. Cela implique que les tenants de question diffèrent.

C'est en effet la conclusion à laquelle nous étions arrivé plus haut en distinguant les différents types de support matériel: la raison orale reposant sur un support corporel renvoie à une communication intersubjective où la question est l'Autre, la communication à autrui. La raison graphique repose quant à elle sur un support matériel externe inanimé qui renvoie à une communication fondée sur la médiation et le partage du support : la question est donc l'extériorité du support et de sa spatialité. C'est donc la question de l'objectivité et de sa constitution dans la spatialité matérielle externe opposée à la question de l'intersubjectivité et de sa constitution matérielle interne, c'est-à-dire corporelle.

L'intersubjectivité comme question de l'Autre et de l'intériorité opposée à l'objectivité comme question de l'objet et de l'extériorité renvoient à l'opposition du vivant au mort, ou plutôt du vivant à l'inanimé. L'intersubjectivité est la communication d'un corps animé qui communique un sens à un autre corps animé. L'inscription corporelle du sens définit une herméneutique de l'intersubjectivité dans la mesure où seul un support vivant peut participer à l'interprétation d'un message inscrit dans un autre corps vivant. Alors que l'inanimé externe peut inscrire dans sa propre matérialité les principes de son interprétation, la corporéité vivante constitue un espace interne d'inscription dont l'extériorisation dans la spatialité matérielle externe est contextuelle : seul un autre corps vivant peut déterminer quelle inscription corporelle peut capter le sens transmis lors d'une communication intersubjective. Alors que toute représentation externe, c'est-à-dire sur un support inanimé, perd son rapport au contexte, la raison orale et son inscription corporelle conservent un rapport à la situation de communication. Il faut donc opposer la raison graphique avec sa permanence et sa décontextualisation à la raison orale, son rapport au contexte et sa volatilité.

C'est cette opposition que l'on retrouve à l'œuvre entre les connaissances scientifiques et les connaissances phénoménologiques : la permanence temporelle du support matériel externe confère à l'écrit une réalité que l'on peut interroger mais non remettre en cause. L'interrogation ne peut aller que dans le sens d'une reconstruction, ou d'un enrichissement objectif. La matérialité en privilégiant la permanence autorise une cumulativité de sens et son enrichissement. Dans la mesure où l'inscription matérielle conditionne son interprétation, sa permanence fait que toute nouvelle inscription peut être rendue compatible avec les prescriptions interprétatives des inscriptions précédentes. En décontextualisant le sens pour l'inscrire sur un support, en d'autres termes en décontextualisant le sens pour en faire des prescriptions interprétatives intemporelles inscrites dans la structure matérielle du support, la raison graphique permet l'élaboration d'une cumulativité du sens qui est souvent revendiquée comme la propriété essentielle de la pensée scientifique.

La raison graphique peut décontextualiser en conservant néanmoins le sens dans la mesure où les prescriptions interprétatives de la structure matérielle constituent un principe de recontextualisation. C'est le support qui prescrit comment être contextualisé pour acquérir un sens. Si bien que la recontextualisation ne dépend pas du contexte effectif dans le lequel le lecteur effectue la recontextualisation : le sens graphique va donc participer à la constitution d'objets qui vont transcender les contextes particuliers ; ces objets sont alors universels et nécessaires : ce sont en d'autres termes des objets scientifiques. En délinéarisant la raison orale pour l'inscrire sur un support matériel externe, la raison graphique inaugure de nouveaux registres du sens puisque les principes constitutifs de ce sens graphique ne procèdent pas d'une raison plongée dans un contexte toujours changeant mais de prescriptions interprétatives inscrites dans la permanence du support. En effet, le support n'est pas seulement un stockage : le support mémorise un contenu dans la mesure où il prescrit comment il doit être lu. Le support n'est pas un espace de rangement, comme le laisse entendre la métaphore du stockage, mais un espace d'accès : la mémoire n'est pas le rangement du contenu mais l'accès au contenu.

La raison orale reste en revanche plongée dans l'évanescence du temps changeant et donc subordonnée à la temporalité du contexte. En constatant que le corps lui permet d'inscrire le sens qu'elle véhicule, on affirme de ce fait que l'inscription matérielle de la raison orale est soumise à la temporalité du vivant, complexe d'espace et de temps. Offrant la spatialité du corps, le vivant est soumis à la temporalité de son devenir, devenir qui le distingue par définition de la matérialité inerte. C'est pourquoi la raison orale n'a de sens qu'en contexte puisqu'elle est toujours soumise au contexte : le contexte d'une communication orale du fait de l'évanescence du support phonique, le contexte de l'inscription corporelle du fait de la temporalité du vivant.

C'est ce qu'illustrent les connaissances phénoménologiques: exprimées dans une écriture phonétique répliquant la raison orale, les connaissances phénoménologiques se décontextualisent dans l'écriture en perdant la possibilité de se recontextualiser par elles-mêmes, comme le permettent les connaissances scientifiques qui mobilisent la raison graphique. Les connaissances phénoménologiques n'ont de sens que pour un lecteur vivant dont la temporalité propre permet de recontextualiser, dans son propre contexte, l'écriture des connaissances phénoménologiques. En répliquant la raison orale sur un support qui n'est pas celui qui lui est propre, en inscrivant la raison orale sur un support externe au lieu du support vivant corporel, l'écriture phonétique s'abstrait du contexte oral sans pour autant gagner la possibilité de recontextualisation inscrite dans le support propre à la raison graphique. D'une certaine manière, l'écriture phonétique est une écriture sans raison graphique, c'est ce que montre la différence phénoménologique.

On peut donc constater un renversement dans le rôle de l'écriture dans les raisons orale et graphique : dans les connaissances phénoménologiques, l'écriture phonétique mémorise l'écriture orale en en faisant l'abstraction. La matérialité de l'écriture phonétique est un support d'inscription qui ne s'est pas substitué à la corporalité dans le rôle de constitution du sens : l'écrit phonétique n'est pas constitutif de son sens dans la mesure où il vient en retard sur le sens oral qui l'a constitué. C'est donc le corps, comme matérialité orale, qui est constitutif du sens oral. En revanche, dans les connaissances scientifiques, c'est l'écriture graphique qui devient constitutif du sens si bien que son passage à l'oral est désormais impossible.

Par conséquent, si dans les deux cas, la matière est bien constitutive du sens, il faut distinguer la temporalité du corps vivant soumis au devenir de l'absence de temporalité, ou une temporalité sur une échelle beaucoup plus longue, des supports matériels externes : de telle sorte que la raison orale se fonde sur la corporéité comme rapport à l'autre, et la raison graphique sur l'objet externe comme rapport à l'extériorité. La raison orale est une communication avec une *autre intériorité*, alors que la raison graphique est la communication par l'intermédiaire d'une *même extériorité.*

La nature des supports matériels permettant l'inscription du sens conditionne la nature du sens inscrit et ses conditions de réappropriation. Le corps propre, vivant, renvoie à la raison orale, l'écrit et l'imprimé à la raison graphique. Mais, nous avons souligné tout au long de ce travail que l'IA pouvait être caractérisée par le fait de retenir un support informatique et computationnel pour l'inscription des connaissances et du sens. Or, si la nature du support conditionne la nature du sens inscrit, il en résulte que l'IA doit permettre de constater l'émergence de nouvelles connaissances. Autrement dit, alors que l'écrit et l'imprimé, statiques, renvoient à la raison graphique (et la raison formelle avec la science), il faudrait, pour exploiter pleinement l'analogie, postuler que le computationnel renvoie à une « raison computationnelle », en entendant par là non pas que la raison soit computationnelle mais que le computationnel permet à la raison de constituer de nouvelles connaissances. Or, à ce moment, la spéculation doit s'arrêter pour laisser la place à l'expérimentation : il faut étudier expérimentalement la raison computationnelle et les connaissances qu'elles constituent. C'est le projet que nous proposons pour l'IA.

5 IA, herméneutique matérielle et artéfacture

A présent, nous pouvons rassembler les éléments proposés par la discussion de ce chapitre pour formuler un programme de recherches pour l'IA dans le cadre de l'herméneutique matérielle que nous venons d'esquisser.

Rappelons le problème que nous avons rencontré en considérant la pratique technologique de l'IA en tant qu'elle élabore des SBC dans le cadre d'une ingénierie des connaissances. L'IA est une ingénierie des connaissances dans la mesure où elle doit modéliser les connaissances d'un domaine, et donc à travers elles, le domaine lui-même, pour les opérationnaliser en SBC. Mais, en cela, l'IA ne se distingue pas d'autres disciplines modélisatrices comme les mathématiques appliquées, l'automatique, etc. Ce qui justifie de dégager une discipline scientifique et technologique à part entière, autonome et distincte des disciplines modélisatrice, c'est qu'elle s'intéresse particulièrement aux domaines non formels et aux connaissances exprimées en langue naturelle qu'ils recèlent. L'approche de l'IA est de formaliser les connaissances de manière à obtenir un modèle effectif du domaine : effectif car exprimé formellement et donc exploitable informatiquement, modèle car renvoyant aux connaissances du domaine, et donc au domaine lui-même.

Or, ce que montre la pratique (cf. introduction), c'est que les modèles conçus en IA sont exprimés formellement mais ne représentent pas formellement le domaine : au contraire, ils s'interprètent linguistiquement en reprenant pour leur vocabulaire formel les termes linguistiques du domaine. C'est ainsi que les modèles IA représentent les connaissances du domaine.

Notre première partie a approfondi cette interprétatibilité des modèles IA en proposant une méthodologie explicitant et exploitant cette caractéristique.

Mais nous avions noté une seconde caractéristique des systèmes et des modèles en IA : c'est qu'ils participent à la constitution de nouvelles connaissances dans le domaine. Autrement dit, alors que le domaine n'est pas formel, le système formel automatique permet de constituer de nouvelles connaissances non formelles dans le domaine. Nous l'avions constaté de manière générale en nous référant aux travaux d'acquisition des connaissances, dans la mesure où ces derniers utilisent l'outil informatique pour constituer les modèles, et de manière plus précise dans la première partie où l'ontologie matérielle du domaine, système non formel des connaissances synthétiques *a priori* du domaine, était constituée par la médiation informatique.

Il nous semble qu'à présent nous avons des éléments de réponse pour expliciter ce fait et pour fonder à partir de ce fait un programme de recherches.

Nos investigations phénoménologiques et herméneutiques nous ont permis de poser l'hypothèse que toute connaissance est constitué par la médiation d'un support matériel externe à la conscience. La conscience est un pur renvoi à l'extériorité sans intériorité : il n'existe pas d'intériorité à la conscience. Autrement dit, il n'existe pas de structures *a priori* de la conscience. La médiation d'un support externe se comprend de manière analogique avec l'objet technique : l'objet est technique dans la mesure où il prescrit son usage sans le déterminer. Le support externe possède une structure qui prescrit sa lecture sans la déterminer.

Or, l'analyse transcendantale que nous avons effectuée à propos de l'informatique et de la logique mathématique qui la théorise, nous a permis de conclure sur le fait que l'informatique constitue un support matériel externe. Elle est un support à double titre : elle l'est en tant qu'elle mobilise des symboles pouvant s'interpréter linguistiquement, mais elle l'est surtout en tant qu'elle construit des calculs qui peuvent s'interpréter. En d'autres termes, l'objet empirique que l'informatique permet de poser comme un support matériel n'est pas un signe statique comme peut l'être une lettre tracée sur le papier, mais un signe dynamique comme l'est un calcul ou un algorithme.

Ce qu'apporte l'informatique, c'est un support dynamique où l'interprétabilité ne doit pas seulement concerner les expressions symboliques considérées à part les unes des autres, de manière statique, mais aussi les expressions symboliques dans leur comportement dynamique, en tant qu'elles sont enrôlées dans un algorithme.

Si l'IA permet, par ses modèles informatiques de constituer de nouvelles connaissances, c'est donc que le calcul constitue un support dynamique fournissant une matière à interprétation qui n'est proposée par aucun autre type de support. Il est donc normal que l'informatique constitue de nouvelles connaissances dans le domaine dans la mesure où elle fournit un nouveau type de support, puisque nous avons argumenté plus haut que chaque type support constitue un type particulier de connaissances.

L'informatique fournit ainsi une technique de la représentation symbolique dynamique. Le terme de « représentation » est à considérer pas seulement dans son sens de « construction syntaxique de symboles », mais aussi dans son sens de représentation de connaissances du domaine. En enrôlant des expressions symboliques possédant une interprétabilité linguistiques car construites à partir des signifiants symboliques de la langue, l'IA leur confère un contenu opératoire et effectif qui *prescrit l'interprétation des représentations sans la déterminer* : en effet, étant donné que nous nous situons dans des domaine non formels, le computationnel ne peut assumer le régional. Néanmoins, et c'est ce que nous avons rencontré dans notre méthodologie proposée dans la première partie, le computationnel prescrit un sens dans le régional ou, pour reprendre le vocabulaire de la première partie, le matériel effectif prescrit et constitue le matériel régional.

Il est clair que l'indétermination de la prescription de l'interprétation par le support dynamique fait de l' IA une discipline expérimentale et technologique. Elle prend deux dimensions, une dimension phénoménologique, une dimension scientifique :

- dans sa dimension phénoménologique, il s'agit de décrire comment le sens se constitue à partir de la prescription informatique ; c'est dans ce cadre que se meuvent nos projets expérimentaux, à savoir la méthodologie exposée dans la première partie et le projet HOSPITEXTE dont nous allons dire quelque mots ;
- dans sa dimension scientifique, il s'agit de théoriser avec les outils de modélisation habituel de la science la description phénoménologique précédente ;

L'IA possède donc une dimension où elle correspond à une phénoménologie du support dynamique, et une autre dimension où elle correspond à une théorie scientifique de la constitution de connaissance à partir des supports dynamiques.

La dimension phénoménologique est autonome et peut constituer un cadre de recherches pour l'IA considérée à travers les catégories de l'herméneutique matérielle. En effet, nous avons dit que le support dynamique constitue des connaissances dans le domaine. Il s'agit alors d'observer expérimentalement cette constitution. C'est l'objectif de notre projet HOSPITEXTE.

Le projet HOSPITEXTE consiste dans l'élaboration d'une station de lecture professionnelle pour la consultation, la constitution et la modification du dossier de spécialité dans un service hospitalier. Le projet HOSPITEXTE est un projet faisant suite au projet européen DOME (Séroussi *et al.*, 1996). Le résultat de ce projet, suite à une étude des besoins des praticiens hospitaliers en unité de soins, a été de mettre en évidence la nécessité d'informatiser le dossier de spécialité sous la forme d'un hypertexte. La réalisation d'une maquette (Bouaud & Séroussi, 1996) (voir figure 7.1) proposant une navigation hypertextuelle dans un dossier a permis de concrétiser et de confirmer ce besoin auprès des utilisateurs potentiels. La maquette a également permis de mettre en évidence la nécessité d'adjoindre aux possibilités de lecture navigationnelle des fonctionnalités d'écriture permettant de travailler les contenus textuels en fonction des objectifs pratiques de la lecture (reconnaître les facteurs de risque, l'histoire de la pathologie courante, éditer un nouveau compte rendu, etc.) : la lecture navigationnelle devient alors une lecture professionnelle.

L'objectif du projet HOSPITEXTE est de poursuivre DOME en réalisant un prototype fonctionnel de station hypertextuelle de lecture professionnelle du dossier de spécialité en médecine. Une telle station doit permettre l'accès et l'appropriation de tous les dossiers d'un service de spécialité clinique. La pratique médicale hospitalière repose en effet en partie sur l'exploitation d'un ensemble de documents réunissant les informations se rapportant à un patient : le dossier de spécialité. Bien que contenant des informations précieuses, le dossier s'avère difficile à manipuler et acheminer (encombrement physique – plus de 10 kg d'environ 150 documents, hétérogénéité des documents le composant – radio, notes, cahiers, etc.) et à consulter (difficulté à localiser l'information recherchée – examiner 150 documents).

Le but du projet est, d'une part, de dématérialiser le dossier (et de résoudre ainsi les problèmes d'encombrement et d'acheminement) et, d'autre part, de se doter, grâce à ce changement de support, d'outils informatiques de consultation et de recherche d'informations – et de résoudre ainsi les problèmes de consultation. Il est en effet crucial qu'une informatisation du dossier patient ne modifie pas la forme ni le contenu d'un dossier médical, mais seulement son support. En effet, le dossier est composé de documents dont la forme et le contenu présentent l'information médicale telle que les médecins veulent la rencontrer dans leur pratique.

Le principe du projet est de concevoir une station hypertextuelle de travail pourvue de fonctions de réécriture permettant à l'utilisateur professionnel de constituer lui-même de nouvelles connaissances à partir de ce que prescrit les liens et inférences dynamiquement effectués par le système[20]. En permettant par exemple à l'utilisateur de calculer des liens

20. Bien avant HOSPITEXTE, un prototype de station de lecture a été développé à l'initiative de B. Stiegler par la société AIS-Berger-Levrault, qui montre comment les fonctions de réécriture sont l'instrument technique pour la constitution de nouvelles connaissances (Chahuneau *et al.*, 1992).

en utilisant des techniques linguistiques du traitement automatique de la langue naturelle (TALN), il peut reconnaître un type de lien possédant un sens médical qu'il décide alors de fixer en une structure stable du dossier hypertextualisé. Par exemple, si le système est capable de retrouver par inférence les facteurs de risque d'un patient pour telle ou telle maladie ou ses antécédents (voir figure 7.2), l'utilisateur peut constater qu'une telle connaissance est systématiquement utile pour chaque consultation du dossier et que, par conséquent, elle doit figurer dorénavant comme un document à part entière du dossier, au même titre que les comptes rendus d'examen ou de sortie.

En d'autres termes, le comportement dynamique de la station permet d'établir des rapports et des corrélations qui, si elles font sens au niveau médical, se stabilisent et constituent de nouvelles connaissances médicales sur le patient. L'objectif du projet, en permettant à l'utilisateur de choisir lui-même ce qui fait sens, est décrire le lien entre ce qui est prescrit par le comportement dynamique du système et ce qui est déterminé comme ayant un sens médical. L'objectif est d'observer quelles configurations de sens possibles proposées par le système, soit automatiquement, soit à la requête de l'utilisateur, vont se stabiliser pour constituer des connaissances médicales reconnues et nécessaires sur le patient. L'analogie qu'il faut ici dresser est celle du passage que nous avons décrit plus haut entre le *volumen* et le *codex*, puis au livre imprimé. Nous avions souligné comment la modification du support entraînait une modification des rapports possibles, c'est-à-dire du sens prescrit. Certains de ces rapports se sont stabilisés en nouvelles connaissances et nouvelles pratiques intellectuelles ; par exemple, l'innovation des index, tables des matières, addenda et errata sur le plan des pratiques, les progrès dans l'éloboration des lois de la mécanique céleste avec la constitution de tables astronomiques imprimées[21] sur le plan des connaissances.

En proposant une hypothèse permettant de comprendre pourquoi et comment l'informatique permet de constituer des connaissances dans des domaines non formels, nous répondons à la question posée à la fin de la première partie : quelle signification donner au fait que l'ontologie matérielle du domaine, c'est-à-dire les connaissances synthétiques *a priori* du domaine, est constituée par la médiation technique de l'informatique ? Notre hypothèse est donc que les structures transcendantales du domaine ne sont ni *a priori*, ni *a posteriori*, mais de nature technique. En d'autres termes, notre hypothèse consiste à dégager une troisième terme entre l'empirique et le transcendantal kantien, le technique. C'est donc ainsi qu'il faut appréhender le projet d'une herméneutique matérielle, c'est-à-dire comme le projet de décrire la constitution du sens non pas à partir des structures transcendantales *a priori* de l'esprit ou de la raison, mais à partir des structures de la « raison technique » qui décrivent comment l'esprit s'investit dans les objets techniques pour déterminer le sens qu'ils lui prescrivent, qu'ils lui adressent.

Enfin, on constate que nos considérations possèdent un impact opératoire direct sur la manière de faire de l'IA. La méthodologie que nous avons proposée dans la première partie permet de contrôler l'investissement linguistique de sens des systèmes formels automatiques pour en faire des modèles du domaine et de constituer de nouvelles connaissances du domaine, son ontologie matérielle. La plupart des travaux en IA insiste sur le fait qu'il y a élaboration de connaissances, mais peu sur le fait que cette élaboration repose pour sa possibilité sur l'outil informatique lui-même. En mettant l'accent sur la constitution informatique des connaissances en inscrivant l'IA dans le cadre de l'herméneutique matérielle, nous proposons de concevoir une artéfacture où l'artefact informatique est un support

21. Pour tous ces exemples, voir (Eisenstein, 1991).

dynamique de prescription du sens.

Connection DOME : Service de rhumatologie (Pr DAHAN), le 10/01/96 à 15:55:26

DOME Fiche de synthèse : Durand Marie

Durand Marie, **Fém**, (**6 mars 1962**), N°service : **83 12391, 90 54579 3**, NIP : **389639562**, [Maj : 31/8/95]

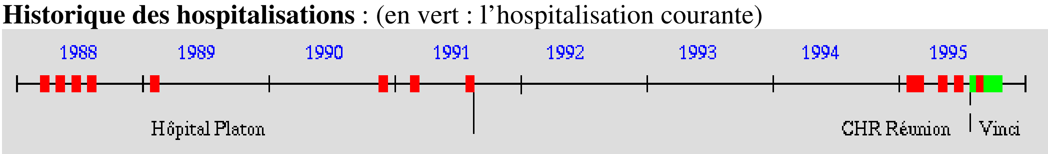
Synthèse | Examens | Imagerie | CRHs | Dossier | Recherche

Attention : *Pneumothorax sur fibroscopie.*

Motif de l'hospitalisation courante : *Dypsnée majeure d'effort*
Diagnostic principal : *Sclérodermie et polymyosite connues depuis 1987.*
Conclusion du CRH de l'hospitalisation précédente (19/5/95) :

> *Aggravation d'une fibrose pulmonaire sclérodermique dans le cadre d'une polymyosite cortico résistante.*

Historique des hospitalisations : (en vert : l'hospitalisation courante)



Suivi de l'état courant

Antécédents : classement par type (Alphabetique | Chronologique)

Histoire de la maladie

Historique des traitements

Soins infirmiers en cours...

Service de Rhumatologie du Pr DAHAN - Hôpital Léonard de Vinci, Paris.

Maquette réalisée par DIAM/SIM AP-HP (diam@biomath.jussieu.fr) pour le projet DOME.
Dernière modification : 23/06/1995 [16:13]

FIG. 7.1 – *Navigation hypertextuelle dans le dossier patient*

Connection DOME : Service de rhumatologie (Pr DAHAN), le 10/01/96 à 15:57:36

Fiche antécédents (par type) : Durand Marie

Durand Marie, **Fém**, (**6 mars 1962**), N°service : **83 12391, 90 54579 3**, NIP : **389639562**, [Maj : 31/8/95]

Synthèse | Examens | Imagerie | CRHs | Dossier | Recherche

Etat courant | Antécédents | HDM | Traitements

Attention : *Pneumothorax sur fibroscopie.*

Classement par type (Alphabétique | Chronologique)

Antécédents Médicaux :

- IVG (1985)
- Tuberculose pulmonaire (1987) *(mention CRH 17/4/88)*
- Abcès dentaire - trismus (Février 1988) *(mention CRH 15/3/88)*
- Septicémie à Staph dore Methi R, endocardite mitrale sur KT (Mars 1988) *(mention CRH 29/5/88)*
- Septicémie à candida albicans (Avril 1988)
- IVG (Janvier 1989)
- Aspergillose pulmonaire (Juin 1995)

Antécédents Chirurgicaux :

- Césarienne (1982) *(mention CRH 17/3/95)*

Rubriques "antécédents" des CRH précédents

Rubriques "conclusion" des CRH précédents

Service de Rhumatologie du Pr DAHAN - Hôpital Léonard de Vinci, Paris.

Maquette réalisée par DIAM/SIM AP-HP (diam@biomath.jussieu.fr) pour le projet DOME.
Dernière modification : 28/06/1995 [15:51]

FIG. 7.2 – *Représentation dans la maquette d'*HOSPITEXTE *des antécédents médicaux d'un patient. Cette fiche pourra être (dans le cadre du projet) constituée à l'aide d'outils automatiques de TALN ou par des réécritures explicites de l'utilisateur. L'hypertexte permet par exemple de stabiliser cette fiche comme un nouveau document à part entière du dossier, alors qu'elle n'existait pas en tant que telle dans le dossier patient.*

Chapitre VIII

Conclusion
Technologies et sciences cognitives

Notre travail dans ce mémoire a consisté à vouloir inscrire notre pratique d'ingénieur dans une réflexion philosophique pour d'une part mieux la comprendre et d'autre part pour lui donner une orientation. Travaillant à la conception de systèmes à base de connaissances médicales, nous sommes confronté au problème d'opérationnaliser informatiquement des connaissances dont le seul médium d'expression est la langue naturelle. Or, en travaillant à ce problème, nous sommes arrivés à un double constat qui a motivé notre réflexion.

Nous avons en effet constaté que les systèmes formels conçus en IA reposent sur l'interprétatibilité linguistique de leurs signifiants symboliques pour représenter les connaissances du domaine et sur leur formalité pour être effectif *sans qu'il y ait de rapport direct entre les deux*. En d'autres termes, un système formel modélise le domaine non pas parce qu'il est formel, mais parce qu'il est interprétable linguistiquement dans les termes des connaissances du domaine, elles-mêmes exprimées en langue naturelle. Cependant, nous avons constaté que, bien que la formalité ne soit pas au principe du caractère modélisateur des systèmes formels, l'effectivité qu'elle implique permet la constitution de nouvelles connaissances dans le domaine, bien que ce dernier ne soit pas formel.

Notre réflexion s'est alors concentré sur deux points particuliers. D'une part comment contrôler l'interprétabilité linguistique pour que les systèmes formels puissent fonctionner comme des modèles du domaine ? D'autre part, comment comprendre la possibilité de constituer par des outils formels des connaissances non formelles ?

La première question nous a conduit à élaborer une méthodologie permettant d'assurer une interprétation des expressions symboliques des systèmes formels qui en fasse des connaissances du domaine. Etant donné que l'interprétation d'expressions formelles est conditionnée par l'interprétation des primitives syntaxiques permettant de les construire, nous avons donc élaboré une méthodologie permettant de construire l'ontologie du domaine, c'est-à-dire les primitives munies de leur signification du formalisme de représentation.

Etant donné la nature non formelle des domaines que nous abordons, nous avons dû concevoir des ontologies non formelles de connaissances, les ontologies matérielles ou régionales. Ces ontologies conditionnent la structure des connaissances du domaine et contiennent les connaissances synthétiques *a priori* du domaine ou région, renouant ainsi avec un thème husserlien.

Mais, la constitution de ces ontologies matérielles repose en sa possibilité sur le recours à l'effectivité des mécanismes symboliques informatiques. En d'autres termes, en traitant la première question, nous nous sommes retrouvé confronté à la seconde question.

Il fallait donc l'aborder. Pour cela nous avons dans un premier temps élaboré une épistémologie transcendantale de l'informatique, qui nous a permis de montrer qu'elle possédait sa logique transcendantale et qu'à ce titre elle était une science expérimentale de la nature. Mais, construisant un savoir expérimental sur les calculs, l'informatique n'explique pas en quoi elle intervient dans la constitution des connaissances d'un domaine non formel. Par conséquent nous avons dû entreprendre dans un second temps une analyse phénoménologique de la connaissance pour comprendre comment la médiation technique pouvait intervenir dans la constitution des connaissances. C'est cette analyse qui nous a conduit à la notion d'une herméneutique matérielle selon laquelle toute connaissance procède d'une médiation technique pour sa constitution, dans la mesure où elle est l'interprétation qui constitue la détermination du sens prescrit par l'objet technique lui-même.

Dans ces conditions, le thème de l'IA se précisait pour devenir l'étude de la médiation technique de l'informatique comme science du calcul dans la prescription qu'elle opérait au niveau du sens et la détermination que ce sens pouvait recevoir au niveau du domaine. A ce moment, notre réflexion s'arrête dans la mesure où elle doit à présent se faire expérimentation.

En résumé, nos propositions sont les suivantes :

- en effectuant une analyse transcendantale de l'informatique, nous avons proposé de voir dans l'informatique un discours objectivant sur le réel. L'informatique est une physique des signes, un physique du calcul. Elle repose sur des principes synthétiques *a priori* et possède, comme discours d'objectivation, sa logique transcendantale.

- l'analyse transcendantale de l'informatique a également permis de proposer de voir dans la logique mathématique qui la théorise une discipline scientifique où le dispositif transcendantal pouvait être appliqué. Pour cela, il fallait se confronter à deux difficultés : la première selon laquelle il fallait comprendre en quoi la logique n'était pas seulement analytique, mais synthétique, retenant dans la logique mathématique la synthéticité des mathématiques et non l'analyticité de la logique formelle ; la seconde selon laquelle le dispositif transcendantal est originalement conçu pour analyser le discours scientifique de la physique, explicitant les principes philosophiques de la construction de l'objet scientifique, et par conséquent il fallait montrer en quoi la logique mathématique construisait une objectivité qui lui était propre. En outre, cette objectivité logique doit s'articuler à l'objectivité informatique. Pour traiter ces problèmes, nous nous sommes inspirés des propositions que J. Petitot a formulées concernant la nature transcendantale des mathématiques et leur articulation à la physique. Pour rester dans les termes de cette analogie, on peut dire que nous avons proposé de voir dans l'informatique une *physique du calcul* et dans la logique mathématique une *mathématique du calcul*.

- enfin, nous avons déplacé la question du transcendantal sur la question de la technique, en proposant de voir dans la technique un troisième terme entre l'empirique et le transcendantal. En effet, nous avions constaté comment la technique informatique nous avait permis de constituer les connaissances synthétiques *a priori* du domaine. L'herméneutique matérielle permet d'expliciter l'hypothèse selon laquelle cette médiation technique, ici informatique, est essentielle pour la constitution des connaissances. En effet, nous avons posé que la conscience ou l'esprit ne possède aucune intériorité dans laquelle il serait possible d'établir des structures *a priori*, analytiques ou synthétiques. L'esprit rencontre dans son extériorité sémiotique ou technique les structures lui permettant de constituer ses connaissances. Ces structures sont matérielles, et soumises au devenir. Contrairement au transcendantal, le technique donne une dimension historique à la constitution transcentantale des connaissances en autorisant d'une part leur évolution et d'autre part leur accumulation.

Comme on a pu le constater, il n'a pas été question dans ce travail de sciences cognitives. En effet, notre projet correspond à une technologie de la représentation symbolique, et non à une science de l'esprit. Nous aimerions par conséquent conclure ce travail sur le fait que l'IA telle que nous la reformulons comme artéfacture dans le cadre d'une herméneutique matérielle correspond à une technologie cognitive qui ne correspond pas une mise en application technologique d'une science cognitive à laquelle elle serait subordonnée. En effet, l'herméneutique matérielle consiste à affirmer que le technique est originaire dans la mesure où toute connaissance procède dans sa constitution d'une médiation technique. Le technique est originaire dans la mesure où c'est le nom que nous avons donné au fait qu'il se trouve toujours, dans notre extériorité, des objets ou choses qui s'adressent à nous en prescrivant leur usage ou interprétation sans qu'il soit possible d'assigner une origine à cette prescription, sans qu'il soit possible de remonter à un auteur de cette prescription. C'est pourquoi, en proposant de voir en l'IA une technologie cognitive décrivant comment une physique du calcul peut servir de support à la constitution de connaissances, nous pensons lui donner l'indépendance à laquelle elle a droit et l'originalité (*originarité*?) qui la légitime.

Références bibliographiques

Akkermans H., van Harmelen F., Schreiber G. & Wielinga B. (1993). A formalization of knowledge-level models for knowledge acquisition. *International Journal of Intelligent Systems*, **8**(2), 169–208.

Alquié F. (1950). *La découverte métaphysique de l'homme chez Descartes.* Epiméthée. Paris: Presses Universitaires de France.

Auroux S. (1995). *La révolution technologique de la grammatisation.* Mardaga.

Auroux S. (1996). *La philosophie du langage.* Collection 1er Cycle. Paris: Presses Universitaires de France.

Bachelard S. (1957). *La logique de Husserl.* Epiméthée. Paris: Presses Universitaires de France.

Bachimont B. (1994). *Le contrôle dans les systèmes à base de connaissances. Contribution à l'épistémologie de l'intelligence artificielle.* Langage, Raisonnement, Calcul. Paris: Hermès. Seconde édition revue et augmentée.

Bachimont B. & Collinot A. (1993). Des connaissances aux représentations : Compiler ou interpréter ? In *Proceedings of the JAC'93 Conference.*

Barbaras R. (1994). *La perception, Essai sur le sensible.* Optiques, Philosophie. Paris: Hatier.

H. Barreau, Ed. (1989). *La mathématique non standard.* Fondements des sciences. Paris: Editions du CNRS.

Bechtel W. (1988). *Philosophy of Mind. An Overview for Cognitive Science.* Lawrence Erlbaum Associates.

Belaval Y. (1960). *Leibniz critique de Descartes.* Collection TEL. Paris: Gallimard.

Blanché R. (1970). *La logique et son histoire & d'Aristote à Russell.* Collection U. Paris: Armand Colin.

Bottéro J. (1987). *Mésopotamie L'écriture, la raison et les dieux.* Bibliothèque des histoires. Paris: Gallimard.

Bouaud J., Bachimont B., Charlet J. & Zweigenbaum P. (1995). Methodological principles for structuring an "ontology". In *Proceedings of the IJCAI'95 Workshop on "Basic Ontological Issues in Knowledge Sharing."*, Montral, Canada.

Bouaud J., Bachimont B., Charlet J., Zweigenbaum P. & Boisvieux J.-F. (1994). Acquisition and structuring of an ontology within conceptual graphs. In *ICCS'94 Workshop on Knowledge Acquisition using Conceptual Graph Theory*, p. 1–25, University of Maryland, College Park, MD.

BOUAUD J., BACHIMONT B. & ZWEIGENBAUM P. (1996a). Processing metonymy: a domain-model heuristic graph traversal approach. In *Proceedings of COLING'96*, Copenhagen, DK.

BOUAUD J., BACHIMONT B. & ZWEIGENBAUM P. (1996b). Traitement de la métonymie basé sur un modèle du domaine et sur une recherche heuristique de chemin dans des graphes. In *Actes de TALN'96*, Marseille.

BOUAUD J. & SÉROUSSI B. (1996). An experiment towards a document-centered hypertextual computerised patient record. In *Submitted to MIE'96*, Copenhagen, DK.

BOURIGAULT D. (1994). *LEXTER : un Logiciel d'EXtraction de Terminologie. Application à l'acquisition des connaissances à partir des textes.* PhD thesis, Ecole des Hautes Etudes en Sciences Sociales.

BRACHMAN R. & SCHMOLZE G. (1985). An overview of the KL-ONE knowledge representation system. *Cognitive Science*, **9**(11), 216–260.

BRACHMAN R. J. (1979). On the epistemological status of semantic networks. In N. FINDLER, Ed., *Associative Networks: Representation and Use of Knowledge by Computers*, p. 3–50. Academic Press.

CASSIRER E. (1957). *Zur modernen Physik.* Oxford University Press.

CHAHUNEAU F., LÉCLUSE C., STIEGLER B. & VIRBEL J. (1992). Prototyping the ultimate tool for scholarly qualitative research on texts. In *Proceedings of the 8th Annual Conference of the UW Center for the New Oxford English Dictionary and Text Research*, Waterloo.

CHANDRASEKARAN B. (1986). Generic tasks in knowledge-based reasoning: High-level building blocks for expert systems design. *IEEE Expert*, **1**(3), 23–30.

CHANDRASEKARAN B. (1990). Design problem solving: A task analysis. *Artificial Intelligence Magazine*, **11**(4), 59–71.

CHANDRASEKARAN B. & JOHNSON T. R. (1993). Generic tasks and task structures: History, critique and new directions. In J.-M. DAVID, J.-P. KRIVINE & R. SIMMONS, Eds., *Second Generation Expert Systems*, p. 232–272. Springer-Verlag.

CHARLET J. (1992). ACTE: acquisition des connaissances par interprétation d'un modèle causal. *Revue d'Intelligence Artificielle*, **6**(1–2), 99–129.

CHARLET J., KRIVINE J.-P. & REYNAUD C. (1996). Causal model-based knowledge acquisition tools: Discussion of experiments. *International journal of Human-Computer Interaction.* A paraître.

CHAUVIRÉ C. (1995). *Peirce et la signification, introduction à la logique du vague.* philosophie d'aujourd'hui. Paris: Presses universitaires de France.

CHELLAS B. F. (1980). *Modal Logic, an Introduction.* Cambridge: Cambridge University Press.

CLANCEY W. J. (1985). Heuristic classification. *Artificial Intelligence Journal*, **27**, 289–350.

CLANCEY W. J. (1989). Viewing knowledge bases as qualitative models. *IEEE-Expert.*

CLANCEY W. J. (1993). The knowledge level reinterpreted: Modeling socio-technical systems. *International Journal of Intelligent Systems*, **8**(1), 33–49.

COCCHIARELLA N. (1991). Formal ontology. In H. BURKHARDT & B. SMITH, Eds., *Handbook of Metaphysics and Ontology*. Münich: Philosophia Verlag.

CORI P. & LASCAR R. (1993). *Logique mathématique.* Masson.

COUTURAT L. (1901). *La logique de Leibniz.* Paris.

CRAMPE-CASNABET M. (1989). *Kant, une révolution philosophique.* Philosophie présente. Paris: Bordas.

DAILLE B. (1995). Repérage et extraction de terminologie par une approche mixte statistique et linguistique. *Traitement Automatique de la Langue*, **36**(1/2), 101–118.

DASTUR F. (1995). *Husserl, des mathématiques à l'histoire.* Presses Universitaires de France.

DAVIS R. (1979). Interactive transfer of expertise: Acquisition of new inference rules. *Artificial Intelligence*, **12**(2), 409–427.

P. DE ROUILHAN & F. RIVENC, Eds. (1992). *Logique et fondements des mathématiques ; Anthologie (1850-1914).* Payot. publication de l'Institut d'Histoire et Philosophie des Sciences et Techniques.

DE SAUSSURE F. (1985). *Cours de Linguistique Générale.* Paris: Payot.

DELAHAYE J.-P. (1995). Arguments et indices dans le débat sur le réalisme mathématique. In J.-M. SALANSKIS & M. PANZA, Eds., *Le platonisme mathématique*, p. 23 – 47. Masson.

DELEUZE G. (1963). *La philosophie critique de Kant.* Paris: Presses Universitaires de France.

DENNETT D. (1991). *Consciousness Explained.* Little, Brown and Company. Traduction française chez Odile Jacob par P. Engel, *la conscience expliquée*, 1993.

DERRIDA J. (1968). *De la grammatologie.* Editions de Minuit.

DESANTI J.-T. (1992). *Réflexions sur le temps ; Variations philosophiques 1.* Figures. Grasset. Conversations avec Dominique-Antoine Grisoni.

DESCARTES R. (1963). *Œuvres philosophiques de Descartes*, volume I : 1618-1637. Paris: Garnier. Edition critique établie par F. Alquié.

DESCARTES R. (1967). *Œuvres philosophiques de Descartes*, volume II : 1638-1642. Paris: Garnier. Edition critique établie par F. Alquié.

DESCARTES R. (1973). *Œuvres philosophiques de Descartes*, volume III : 1643-1650. Paris: Garnier. Edition critique établie par F. Alquié.

M. DETIENNE, Ed. (1981). *Les savoirs de l'écriture dans la Grèce ancienne.* Presses Universitaires de Lille.

DILTHEY W. (1947). *Le monde de l'esprit.* Aubier. 2 volumes.

DOJAT M. & PACHET F. (1995). Effective domain-dependent reuse in medical knowledge bases. *Computers and Biomedical Research*, **28**(6), 403–432.

DORLAND (1994). *Dorland's Illustrated Medical Dictionary.* W.B. Saunders Company.

DREYFUS H. (1984). *Intelligence artificielle : Mythes et limites.* Flammarion.

H. DREYFUS & H. HALL, Eds. (1982). *Husserl, Intentionality and Cognitive Science.* A Bradford Book. Cambridge Mass: The MIT Press.

DREYFUS H. L. (1972). *What Computers Can't do.* New York: Harper Row.

DREYFUS H. L. (1991a). Husserl et les sciences cognitives. *Les Études Philosophiques*, **1**, 1–30.

DREYFUS H. L. (1991b). L'épiphénoménologie de Husserl. *Les Études Philosophiques*, **1**, 57–78.

DUMMETT M. (1991). *Les origines de la philosophie analytique.* nrf essais. Paris: Gallimard.

DUPUY J.-P. (1994). *Aux origines des sciences cognitives.* Paris: Editions La Découverte.

DUURSMA C. (1992). Interpretation models and problem solving methods. In B. NEUMANN, Ed., *Proceedings of the ECAI'92 Conference*, p. 572–576.

ECO U. (1985). *Lector in Fabula & Le rôle du lecteur ou la Coopération interprétative dans les textes narratifs.* Paris: Grasset & Fasquelle.

ECO U. (1988a). *Le signe; Histoire et analyse d'un concept.* Bruxelles: Editions Labor.

ECO U. (1988b). *Sémiotique et philosophie du langage.* Formes Sémiotiques. Paris: Presses Universitaires de France.

ECO U. (1992). *Les limites de l'interprétation.* Grasset & Fasquelle.

EISENSTEIN E. L. (1991). *La révolution de l'imprimé dans l'europe des premiers temps modernes.* anthropologie des sciences et des techniques. Paris: Editions La Découverte.

EISLER R. (1994). *Kant-Lexicon.* Gallimard.

ENGEL P. (1989). *La norme du vrai; essai sur la philosophie de la logique.* Gallimard.

EVERAERT-DESMEDT N. (1990). *Le processus interprétatif; introduction à la sémiotique de Ch. S. Peirce.* Philosophie et Langage. Mardaga.

FEFERMAN S. (1988). Hilbert's program relativized: Proof-theoretical and foundational reduction. *The Journal of Symbolic Logic*, **53**, 364–384.

FELBER H. (1987). *Manuel de terminologie.* Paris: Unesco.

FISETTE D. (1994). *Lecture frégéenne de la phénoménologie.* Tiré à part. l'éclat.

FODOR J. (1975). *The language of Thought.* The MIT Press.

FODOR J. A. (1981). Methodological solipsism considered as a research strategy in cognitive psychology. In J. HAUGELAND, Ed., *Mind Design*, p. 307 – 338. The MIT Press.

FODOR J. A. (1983). *La modularité de l'esprit.* Propositions. Paris: Minuit.

FODOR J. A. (1987). *Psychosemantics.* A Bradford Book. The MIT Press.

FOLLESDAL D. (1982). Husserl's Notion of Noema. In H. DREYFUS & H. HALL, Eds., *Husserl, Intentionality and Cognitive Science*, p. 73 – 80. The MIT Press.

FREGE G. (1971a). Le projet d'une idéographie. In C. IMBERT, Ed., *Frege: Ecrits philosophiques et logiques*. Paris: Editions du Seuil.

FREGE G. (1971b). Über Sinn und Bedeutung. In C. IMBERT, Ed., *Frege: Ecrits philosophiques et logiques*. Paris: Editions du Seuil.

GADAMER H.-G. (1960). *Vérité et méthode*. Le Seuil.

GAINES B. R., SHAW M. L. G. & WOODWARD J. B. (1993). Modeling as framework for knowledge acquisition methodologies and tools. *International Journal of Intelligent Systems*, **8**(2), 155–168.

GANASCIA J.-G. (1990). *L'âme-machine, Les enjeux de l'intelligence artificielle*. Paris: Editions du Seuil.

GANASCIA J.-G. (1993). *L'intelligence artificielle*. Dominos. Paris: Flammarion.

GILSON E. (1981). *L'Être et l'Éssence*. Paris: Librairie Philosophique Jean Vrin.

GINSBERG M. L. (1991). Knowledge interchange format: the KIF of death. *Artificial Intelligence Magazine*, **12**(3), 57–63.

GOCHET P. & GRIBOMONT P. (1990). *Logique, Tome 1 et 2*. Hermès.

GÖDEL K. (1989). Sur les propositions formellement indécidables des *Principia Mathematica* et des systèmes apparentés I. In *Le théorème de Gödel*. Seuil.

GOODY J. (1979). *La raison graphique, la domestication de la pensée sauvage*. Le sens commun. Paris: Les Editions de Minuit.

GOODY J. (1985). *La logique de l'écriture*. Armand Colin.

GRANEL G. (1968). *Le sens du temps et de la perception chez E. Husserl*. Bibliothèque de Philosophie. Paris: Gallimard.

GREFENSTETTE G. (1994). *Exploration in Automatic Thesaurus Discorvery*. Kluwer Academic Publishers.

GREIMAS A. J. (1986). *Sémantique structurale*. Presses Universitaires de France.

GRONDIN J. (1993). *L'universalité de l'herméneutique*. Epiméthée. Presses Universitaires de France.

GRUBER T. (1993). A translation approach to portable ontology specifications. *Knowledge Acquisition*, **5**, 199–220.

GRUBER T. (1995). Toward principles for the design of ontologies used for knowledge sharing. *International Journal of Human-Computer Studies*, **43**(5/6), 907–928.

GUARINO N. (1992). Concepts, attributes and arbitrary relations. *Data & Knowledge Engineering*, **8**, 249–261.

GUARINO N. (1994). The ontological level. In R. CASATI, B. SMITH & G. WHITE, Eds., *Philosophy and the Cognitive Sciences*. Vienna: Hölder-Pichler-Tempsky.

GUARINO N. (1995). Formal ontology, conceptual analysis and knowledge representation. *International Journal of Human-Computer Studies*, **43**(5/6), 625–640.

GUARINO N., CARRARA M. & GIARETTA P. (1994). *An Ontology of Meta-Level Categories.* Rapport interne 6/93, LADSEB-CNR.

GUENANCIA P. (1986). *Descartes.* Philosophie présente. Paris: Bordas.

HABERT B., BARBAUD P., DUPUIS F. & JACQUEMIN C. (1995). Simplifier des arbres d'analyse pour dégager les comportements syntaxico-smantiques des formes d'un corpus. *Cahiers de Grammaire*, **19**, 1–19.

HABERT B. & JACQUEMIN C. (1993). Noms composés, termes, dénominations complexes : problématiques linguistiques et traitements automatiques. *Traitement Automatique de la langue*, **34**(2), 5–42.

HABERT B. & NAZARENKO A. (1996). La syntaxe comme marche-pied de l'acquisition des connaissances : bilan critique d'une expérience. In *Actes des Journées d'Acquisition des Connaissances.*

HAHN, NEURATH & CARNAP (1985). La conception scientifique du monde. In A. SOULEZ, Ed., *Manifeste du Cercle de Vienne et autres écrits*, p. 105–130. Presses Universitaires de France.

HAREL D. (1978). *First-Order Dynamic Logic.* Lecture Notes in Computer Science No. 68. Springer-Verlag.

HARRISON T. (1992). *Principes de Médecine Interne.* Médecine-Sciences Flammarion, 5 édition française, traduite de la 12e édition américaine edition.

HAUGELAND J. (1981). Semantic engines: an introduction to mind design. In J. HAUGELAND, Ed., *Mind Design*, p. 1 – 34. Cambridge Mass: The MIT Press.

HAUGELAND J. (1985). *Artificial Intelligence & The Very Idea.* Cambridge Mass: The MIT Press.

HAUGELAND J. (1989). *L'esprit dans la machine ; Fondements de l'intelligence artificielle.* Paris: Odile Jacob.

HAYES P. J. (1985). The Second Naive Physics Manifesto. In J. R. HOBBS & R. C. MOORE, Eds., *Formal Theories of the Commonsense World*, p. 1–36. Ablex.

HAYES-ROTH B. (1985). A Blackboard Architecture for Control. *Artificial Intelligence*, **26**, 252 – 321.

HEIDEGGER M. (1953). *Kant et le problème de la métaphysique.* Collection TEL. Paris: Gallimard. Kant und das Problem der Metaphysik, 1965, trad. de Waelhens et Biemel.

HEIDEGGER M. (1971). *Qu'est-ce qu'une chose ?* Collection TEL. Paris: Gallimard. Die Frage nach dem Ding, 1962, trad. Reboul et Taminiaux.

HEIDEGGER M. (1982). *Interprétation phénoménologique de la "Critique de raison pure" de Kant.* Bibliothèque de de Philosophie Série : œuvre de Martin Heidegger. Paris: Gallimard.

HEIDEGGER M. (1986). *Être et Temps.* Bibliothèque de de Philosophie Série : œuvre de Martin Heidegger. Paris: Gallimard.

HILBERT D. (1972). Sur l'infini. In J. LARGEAULT, Ed., *Logique mathématique: textes.* Paris: Armand Colin.

HJEMSLEV L. (1971). *Prolégomènes à une théorie du langage.* Minuit.

HOBBES T. (1651). *The Leviathan.*

HOBBS J. R. (1995). Sketch of an ontology underlying the way we talk about the world. *International Journal of Human-Computer Studies*, **43**(5/6), 819–830.

J. R. HOBBS & R. C. MOORE, Eds. (1985). *Formal Theories of the Commonsense World.* Ablex.

HODGES (1985–1986). Truth in a structure. *Proceedings of the Aristotelian Society*, p. 135–151.

HOOKWAY C. (1985). *Peirce.* Routledge.

HOOKWAY C. (1994). Iconicity and Logical Form. *Histoire, Épistémologie, Langage*, **16**(1), 53 – 64.

HUSSERL E. (1928). *Méditations Cartésiennes.* Paris: Librairie Philosophique Jean Vrin. Cartesianische Meditationen, 1928, trad. Pfeiffer et Lévinas.

HUSSERL E. (1950). *Idées Directrices pour une Phénoménologie.* Collection TEL. Paris: Gallimard. Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie, erstes Buch, 1913, trad. Ricœur.

HUSSERL E. (1957). *Logique Formelle et Logique Transcendantale.* Epiméthée. Paris: Presses Universitaires de France. Formale und Transzendentale Logik. Versuch einer Kritik der logischen Vernunft, 1929, trad. Bachelard.

HUSSERL E. (1959). *Recherches Logiques.* Epiméthée. Paris: Presses Universitaires de France. Logischen Untersuchungen, 1895, trad. Scherer, Elie et Kelkel.

HUSSERL E. (1964). *Leçons pour une phénoménologie de la conscience intime du temps.* Epiméthée. Paris: Presses Universitaires de France. Vorlesungen zur Phänomenologie des inneren Zeitbewusstseins, 1928, trad. Dussort.

HUSSERL E. (1968). *L'origine de la géométrie.* Presses Universitaires de France. Traduction et introduction de J. Derrida.

HUSSERL E. (1982). *Idées directrices pour une phénoménologie et une philosophie phénoménologique pures ; livre second : Recherches phénoménologiques pour la constitution.* Epiméthée. Paris: Presses Universitaires de France. Ideen zu einer reinen Phänomenologie und phänomenologischen Philosophie, zweites Buch : Phänomenologische Untersuchungen sur Konstitution, 1952, trad. Escoubas.

HUSSERL E. (1991). *Expérience et Jugement.* Epiméthée. Paris: Presses Universitaires de France. Erfahrung und Urteil, 1950, trad. Souche-Dague.

HUSSERL E. (1993). *La phénoménologie et le fondement des sciences.* Epiméthée. Paris: Presses Universitaires de France. trad. Tiffanau.

JACKENDOFF R. S. (1987). *Consciousness and the Computational Mind.* The MIT Press.

JACOB P. (1980a). *De Vienne à Cambridge.* Gallimard.

JACOB P. (1980b). *L'empirisme logique.* Propositions. Paris: Minuit.

D. JANICAUD, Ed. (1995). *L'intentionnalité en question : entre phénoménologie et recherches cognitives.* Paris: Librairie Philosophique Jean Vrin.

KANT E. (1965). *Prolégomènes à toute métaphysique future qui pourra se présenter comme science.* Paris: Librairie philosophique Jean Vrin. 1783, trad. Gibelin.

KANT E. (1980). *Critique de la Raison Pure.* Paris: Gallimard. Kritik der reinen Vernunft, 1781, Edition de la Pleëiade.

KANT E. (1982). *Logique.* Paris: Librairie Philosophique Jean Vrin. Logik, 1800, trad. Guillermit.

KANT E. (1984a). *Critique de la Faculté de Juger.* Paris: Librairie Philosophique Jean Vrin. Kritik der Urteilkraft, 1790, trad. Philonenko.

KANT E. (1984b). *Critique de la Raison Pure.* Quadriges. Paris: Presses Universitaires de France. Kritik der reinen Vernunft, 1781, trad. Tremesaygues et Pacaut.

KANT E. (1990). *Premiers Principes métaphysique de la science de la nature.* Paris: Librairie Philosophique Jean Vrin. 1786, trad. Gibelin.

KATHLEEN D. (1995). A linguistic ontology. *International Journal of Human-Computer Studies*, **43**(5/6), 809–818.

KERAVNOU E., DAMS F., WASHBROOK J., DAWOOD R., HALL C. & SHAW D. (1992). Background knowledge in diagnosis. *Artificial Intelligence in Medicine*, **4**, 263–279.

KOYRÉ A. (1966). *Etudes galiléennes.* Paris: Hermann.

LAIRD J. E., NEWELL A. & ROSENBLOOM P. S. (1987). SOAR : An architecture for general intelligence. *The Journal of Artificial Intelligence*, **33**, 1–64.

LALEMENT R. (1990). *Logique, Réduction, Résolution.* Etudes et Recherches en Informatique. Paris, France: Masson.

LAMBEK J. & SCOTT P. (1986). *Introduction to higher order categorical logic.* Cambridge University Press.

LARGEAULT J. (1972). *Logique mathématique – Textes.* Armand Colin.

LARGEAULT J. (1992). *Intuitionisme et théorie de la démonstration.* Vrin.

LASSAIGNE R. & DE ROUGEMONT M. (1993). *Logique et fondements de l'informatique ; logique du 1er ordre, calculabilité et lambda calcul.* Hermès.

LAURIÈRE J.-L. (1987). *INTELLIGENCE ARTIFICIELLE Résolution de Problèmes par l'Homme et la Machine.* Paris: Eyrolles.

LENAT D. & FEIGENBAUM E. (1991). On the thresholds of knowledge. *Artificial Intelligence*, **47**(1–3), 185–250.

LENAT D. & GUHA R. (1990). *Building Large Knowledge-Based Systems ; Representations and Inference in the Cyc Project.* Reading, Massachusetts: Addison Wesley.

LENAT D., PRAKASH M. & SHEPHERD M. (1986). Cyc: Using Common Sense Knowledge to Overcome Brittleness and Knowledge Acquisition Bottlenecks. *Artificial Intelligence Magazine*, **6**(4), 65–85.

LERAT P. (1995). *Les langues spécialisées.* Linguistique nouvelle. Paris: Presses universitaires de France, 2ème édition edition.

LEROI-GOURHAN A. (1964). *Le geste et la parole*, volume Tome 1 : Technique et Langage. Paris: Albin Michel.

LÉVINAS E. (1988). *En découvrant l'existence avec Husserl et Heidegger.* Paris: Librairie Philosophique Jean Vrin.

LÉVINAS E. (1989). *Théorie de l'intuition dans la phénoménologie de Husserl.* Paris: Librairie Philosophique Jean Vrin.

LINSTER M. (1993). Closing the gap between modeling to make sense and modeling to implement systems. *International Journal of Intelligent Systems*, **8**(2), 209–230.

W. G. LYCAN, Ed. (1990). *Mind and Cognition, A Reader.* Basil Blackwell.

MALHERBE J.-F. (1981). *Epistémologies Anglo-Saxonnes.* Presses Universitaires de France.

MARTIN H.-J. (1996). *Histoire et pouvoirs de l'écrit.* Bibliothèque de l'Évolution de l'Humanité. Albin Michel. avec la colloboration de Bruno Delmas.

MARTIN R. (1992). *Pour une Logique du Sens.* Linguistique Nouvelle. Paris: PUF, 2ème édition edition.

MARTINET A. (1980). *Eléments de linguistique générale.* Armand Colin.

MASINI G., NAPOLI A., COLNET D., LÉONAR D. & TOMBRE K. (1991). *Les langages à objets. Langages de classes, langages de frames, langages d'acteurs.* InterEditions.

MCDERMOTT J. (1988). Preliminary steps toward a taxonomy of problem-solving methods. In S. MARCUS, Ed., *Automating Knowledge Acquisition for Expert Systems*, p. 225–255. Kluwer.

MENDELSON E. (1987). *Introduction to Mathematical Logic.* Advanced Books & Software. Monterey, California: Wadsworth & Brooks/Cole.

MIKHEEV A. & FINCH S. (1995). A workbench for acquisition of ontological knowledge from natural language. In *Proceedings of the Seventh Conference of the European Chapter of the Association for Computational Linguistics*, p. 194–201.

MILLER I. (1984). *Husserl & Perception and Temporal Awareness.* A Bradford Book. Cambridge, Ma: The MIT Press.

NARAYANAN A. (1990). *On being a Machine Volume 2: Philosophy of Artificial Intelligence.* Ellis Horwood Series in Artificial Intelligence Foundations and Concepts. Ellis Horwood Limited.

NECHES R., FIKES R., FIHIN T., GRUBER T., PATIL R., SENATOR T. & SWARTOUT W. R. (1991). Enabling technology for knowledge sharing. *AI magazine*, **12**(3), 36–56.

NEF F. (1988). *La logique et le langage.* Hermès.

NEWELL A. (1980). Physical Symbol Systems. *Cognitive Science*, **4**, 135 – 183.

NEWELL A. (1982). The Knowledge Level. *Artificial Intelligence*, **18**, 87–127.

NEWELL A. & SIMON H. A. (1981). Computer science as empirical inquiry: Symbols and search. In J. HAUGELAND, Ed., *Mind Design*, p. 35 – 66. The MIT Press.

NORMAN D. A. (1983). Some observations on mental models. In D. GENTNER & A. L. STEVENS, Eds., *Mental Models*, p. 7–14. Lawrence Erlbaum Associates.

NYGREN E. & HENRIKSSON P. (1992). Reading the medical record. I. *Computer Methods and Programs in Biomedicine*, **39**, 1–12.

NYGREN E., JOHNSON M. & HENRIKSSON P. (1992). Reading the medical record. II. *Computer Methods and Programs in Biomedicine*, **39**, 13–25.

PACHERIE E. (1993). *Naturaliser l'intentionnalité.* Paris: Presses Universitaires de France.

PASCAL G. (1957). *La pensée de Kant.* Pour connaître. Paris: Bordas.

PATOČKA J. (1988). *Qu'est ce que la phénoménologie?* Krisis. Paris: Millon.

PATOČKA J. (1992). *Introduction à la phénoménologie de Husserl.* Krisis. Paris: Millon.

PEIRCE C. S. (1960). *Collected Papers.* Harvard University Press.

PEIRCE C. S. (1967). *The Annotated Catalogue of the Papers of C.S. Peirce.* University of Massachusetts Press.

PEIRCE C. S. (1976). *The New Elements of Mathematics.* Mouton. 4 volumes, édités par C. Eisele.

PEIRCE C. S. (1978). *Ecrits sur le signe.* Paris: Editions du Seuil. rassemblés, traduits et commentés par Gérard Deledalle.

PEIRCE C. S. (1984). *Textes anticartésiens.* Paris: Aubier. traduits et introduits par Joseph Chenu.

PEIRCE C. S. (1992). *Reasoning and the Logic of Things; The Cambridge Lectures of 1898.* Harvard University Press. préfacé par K.L. Ketner et H. Putnam.

PETITOT J. (1986). *Morphogenèse du Sens.* Formes Sémiotiques. Paris: Presses Universitaires de France.

PETITOT J. (1991a). De la physique à la forme et au sens. Actualité de l'épistémologie transcendantale. In F. MARTY, Ed., *L'épistémologie transcendantale et le problème de l'objectivité*, Entretiens du Centre Sèvres, p. 59–110. Paris: Osiris.

PETITOT J. (1991b). Idéalités mathématiques et réalité objective. In *Hommage à Jean-Toussaint Desanti*, p. 213–282. Paris: T.E.R.

PETITOT J. (1992a). Continu et objectivité. In J.-M. SALANSKIS & H. SINACEUR, Eds., *Le labyrinthe du continu*, p. 200–220. Paris: Springer-Verlag.

PETITOT J. (1992b). *Physique du sens.* Paris: Editions du CNRS.

PETITOT J. (1993). Phénoménologie naturalisée et morphodynamique : la fonction cognitive du synthétique a priori. *Intellectica*, **17**(2), 79–126.

PETITOT J. (1995). Sheaf mereology and Husserl's morphological ontology. *International Journal of Human-Computer Studies*, **43**(5/6), 741–763.

PHILONENKO A. (1968). *L'œuvre de Kant.* A la recherche de la vérité. Paris: Librairie Philosophique Jean Vrin.

PHILONENKO A. (1989). *L'Ecole de Marbourg.* A la recherche de la vérité. Paris: Librairie Philosophique Jean Vrin.

POTTIER B. (1992). *Sémantique Générale.* Linguistique Nouvelle. 75006 Paris: PUF, 1ère édition edition.

PUTNAM H. (1989). *Représentation et Réalité.* nrf essais. Paris: Gallimard.

PUTNAM H. (1991). Peut-on encore parler de vérité et de réalité? sur *La Crise des sciences européennes*. In J. POULAIN, Ed., *Critique de la raison phénoménologique*, p. 83–98. Cerf.

PYLYSHYN Z. W. (1984). *Computation and Cognition, Toward a Foundation for Cognitive Science.* Cambridge Ma: The MIT Press.

QUINE W. O. (1977). *Relativité de l'ontologie et autres essais.* Aubier. Ontological relativity and other essays, 1969, trad. française de J. Largeault.

RASTIER F. (1987). *Sémantique Interprétative.* Formes Sémiotiques. Paris: Presses Universitaires de France.

RASTIER F. (1991). *Sémantique et Recherches Cognitives.* Formes Sémiotiques. Paris: Presses Universitaires de France.

RASTIER F., CAVAZZA M. & ABEILLÉ A. (1994). *Sémantique pour l'analyse.* Paris: Masson.

RICHARD J.-F. (1990). *Les activités mentales.* Paris: Armand Colin.

RICOEUR P. (1965). *De l'interprétation, essai sur Freud.* Seuil.

RICOEUR P. (1969). *Le conflit des interprétations, Essais d'herméneutique.* Seuil.

RICOEUR P. (1985). *Temps et récit: le temps raconté (tome 3).* Paris: Seuil.

RICOEUR P. (1986). *Du texte à l'action, Essais d'herméneutique, II.* Seuil.

RICOEUR P. (1987). *A l'école de la phénoménologie.* Paris: Librairie Philosophique Jean Vrin.

RIVELAYGUE J. (1992). *Leçons de métaphysique allemande, Tome II: Kant, Heidegger, Habermas.* Paris: Grasset.

RODIS-LEWIS G. (1971). *L'œuvre de Descartes.* A la recherche de la vérité. Paris: Librairie Philosophique Jean Vrin.

SAINT AUGUSTIN (1964). *Confessions.* Paris: GF-Flammarion.

SALANSKIS J.-M. (1991). *Herméneutique Formelle.* Fondements des sciences. Paris: Editions du CNRS.

SALANSKIS J.-M. (1992). Modes du continu dans les sciences. *Intellectica*, **13–14**(1–2), 45–78.

SALANSKIS J.-M. & PANZA M. (1995). *Le platonisme mathématique.* Masson.

J.-M. SALANSKIS & H. SINACEUR, Eds. (1992). *Le labyrinthe du continu*. Paris: Springer-Verlag.

SCHÉRER R. (1962). *La phénoménologie des "Recherches Logiques" de Husserl*. Epiméthée. Paris: Presses Universitaires de France.

SCHLICK M. (1985). Le vécu, la connaissance, la métaphysique. In A. SOULEZ, Ed., *Manifeste du Cercle de Vienne et autres écrits*, p. 183–197. Presses Universitaires de France.

G. SCHREIBER, B. WIELINGA & J. BREUKER, Eds. (1993). *KADS: A Principled Approach to Knowledge-based System Development*. Academic Press.

SEARLE J. (1981). Minds, Brains and Programs. In J. HAUGELAND, Ed., *Mind Design*. The MIT Press.

SÉROUSSI B., BOUAUD J., BACHIMONT B., CHARLET J. & ZWEIGENBAUM P. (1996). Document management system as the framework of a computer-based patient record. In *Proceedings of the SCAMC'96 Conference*. in submission.

SHANK R. (1990). What is AI anyway? In D. PARTRIDGE & Y. WILKS, Eds., *The foundations of artificial intelligence; A Sourcebook*, p. 1 – 13. Cambridge University Press.

SHANKER S. (1987). *Wittgenstein and the Turning-Point in the Philosophy of Mathematics*. State Universitate of New-York Press.

SHORTLIFFE E. & BUCHANAN B. (1975). A model of inexact reasoning in medicine. *Mathematical Biosciences*, **23**, 351–379.

SIMPSON S. (1985). Reverse mathematics. In A. NERODE & R. SHORE, Eds., *Recursion Theory*, Proceedings of Symposia in Pure Mathematics, XLII, Amer. Math. Soc., p. 461–471. Providence. R.I.

SMITH B. & CASATI R. (1993). La physique naïve, un essai d'ontologie. *Intellectica*, **17**(2), 173–197.

SMOLENSKY P. (1988). On the proper treatment of connectionism. *The Behavioral et Brain Sciences*, **11**(1), 1–23.

SODERLAND S., FISHER D., ASELTINE J. & LEHNERT W. (1995). CRYSTAL: Inducing a Conceptual Dictionary. In *Proceedings of the Fourteenth International Joint Conference on Artificial Intelligence*: Morgan Kaufmann.

SOWA J. (1984). *Conceptual Structures*. Addison Wesley.

SOWA J. (1991). Toward the expressive power of natural language. In J. SOWA, Ed., *Principles of Semantic Networks*, p. 157–189. Morgan Kaufmann.

SOWA J. (1995). Top-level ontological categories. *International Journal of Human-Computer Studies*, **43**(5/6), 669–685.

STEELS L. (1990). Components of expertise. *Artificial Intelligence Magazine*, **11**(2), 29–49.

STEELS L. (1993). The componential framework and its role in reusability. In J.-M. DAVID, J.-P. KRIVINE & R. SIMMONS, Eds., *Second Generation Expert Systems*, p. 273–298. Springer-Verlag.

STIEGLER B. (1994). *La technique et le temps ; Tome I : la faute d'Epiméthée.* Paris: Galilée.

STIEGLER B. (1995). Ce qui fait défaut. *Césure*, p. 233–278.

STIEGLER B. (1996). *La technique et le temps ; Tome II : la désorientation.* Galilée.

THIBAUD P. (1975). *La logique de Charles Sanders Peirce De l'algèbre aux Graphes.* Editions de l'université de Provence.

TIERCELIN C. (1993). *La pensée-signe; Études sur C.S. Peirce.* Editions Jacqueline Chambon.

TRAN-DUC-THAO (1992). *Phénoménologie et matérialisme dialectique.* Paris: Gordon & Breach.

TURING A. (1990). Computing machinery and intelligence. In M. BODEN, Ed., *The Philosophy of Artificial Intelligence*, Oxford Readings in Philosophy. Oxford University Press.

TURING A. M. & GIRARD J.-Y. (1995). *La machine de Turing.* Paris: Seuil.

VAN HEIJENOORT J. (1967). *From Frege to Gödel, A Source Book in Mathematical Logic, 1879-1931.* Harvard University Press.

VAN HEIJST G., FALASCONI S., ABU-HANNA A., SCHREIBER G. & STEFANELLI M. (1995). A case study in ontology library construction. *Artificial Intelligence in Medicine*, **0**, 0.

VAN HEIJST G., SCHREIBER G. & WIELINGA B. J. (1996). Usiing explicit ontologies in kbs development. *International Journal of Human-Computer Studies.* seminal paper as call for submission for a special issue on ontologies, to appear.

VISETTI Y.-M. (1991). Des systèmes experts aux systèmes à base de connaissances : à la recherche d'un nouveau schéma régulateur. *Intellectica*, **12**, 221–279.

VUILLEMIN J. (1955). *Physique et métaphysique Kantiennes.* Dito. Paris: Presses Universitaires de France.

WEIZENBAUM J. (1966). ELIZA: a computer program for the study of natural language communication between man and machine. *Communication of the* ACM, **9**(1), 36–45.

WIELINGA B., VAN DE WELDE W., SCHREIBER G. & AKKERMANS H. (1993). Towards a unification of knowledge modelling approaches. In J.-M. DAVID, J.-P. KRIVINE & R. SIMMONS, Eds., *Second Generation Expert Systems*, p. 299–335. Springer-Verlag.

WIELINGA B. J. & BREUKER J. A. (1984). Intepretation of verbal data for knowledge acquisition. In *Proceedings of the ECAI'84 Conference*, p. 41–50.

WINOGRAD T. (1990). Thinking machines: Can there be? are we? In D. PARTRIDGE & Y. WILKS, Eds., *The foundations of artificial intelligence ; A Sourcebook*, p. 167–189. Cambridge University Press.

WITTGENSTEIN L. (1967). *Wittgenstein und der WienerKreis.* Suhrkamp & Blackwell. Notes d'entretiens prise par Fr. Waismann, éditées par MacGuiness.

YASUHARA A. (1971). *Recursive Function Theory & Logic.* Academic Press.

ZWEIGENBAUM P., BACHIMONT B., BOUAUD J., CHARLET J. & BOISVIEUX J.-F. (1995a). Issues in the structuration and acquisition of an ontology for medical language understanding. *Methods of Informations in Medicine*, **34**(1), 15–24.

ZWEIGENBAUM P., BACHIMONT B., BOUAUD J., CHARLET J. & BOISVIEUX J.-F. (1995b). A multi-lingual architecture for building a normalised conceptual representation from medical language. In R. M. GARDNER, Ed., *Proc 17th Annu Symp Computer Applications in Medical Care*, New Orleans.

ZWEIGENBAUM P. & CONSORTIUM MENELAS (1994). MENELAS: an access system for medical records using natural language. *Computer Methods and Programs in Biomedicine*, **45**, 117–120.

Index des notions