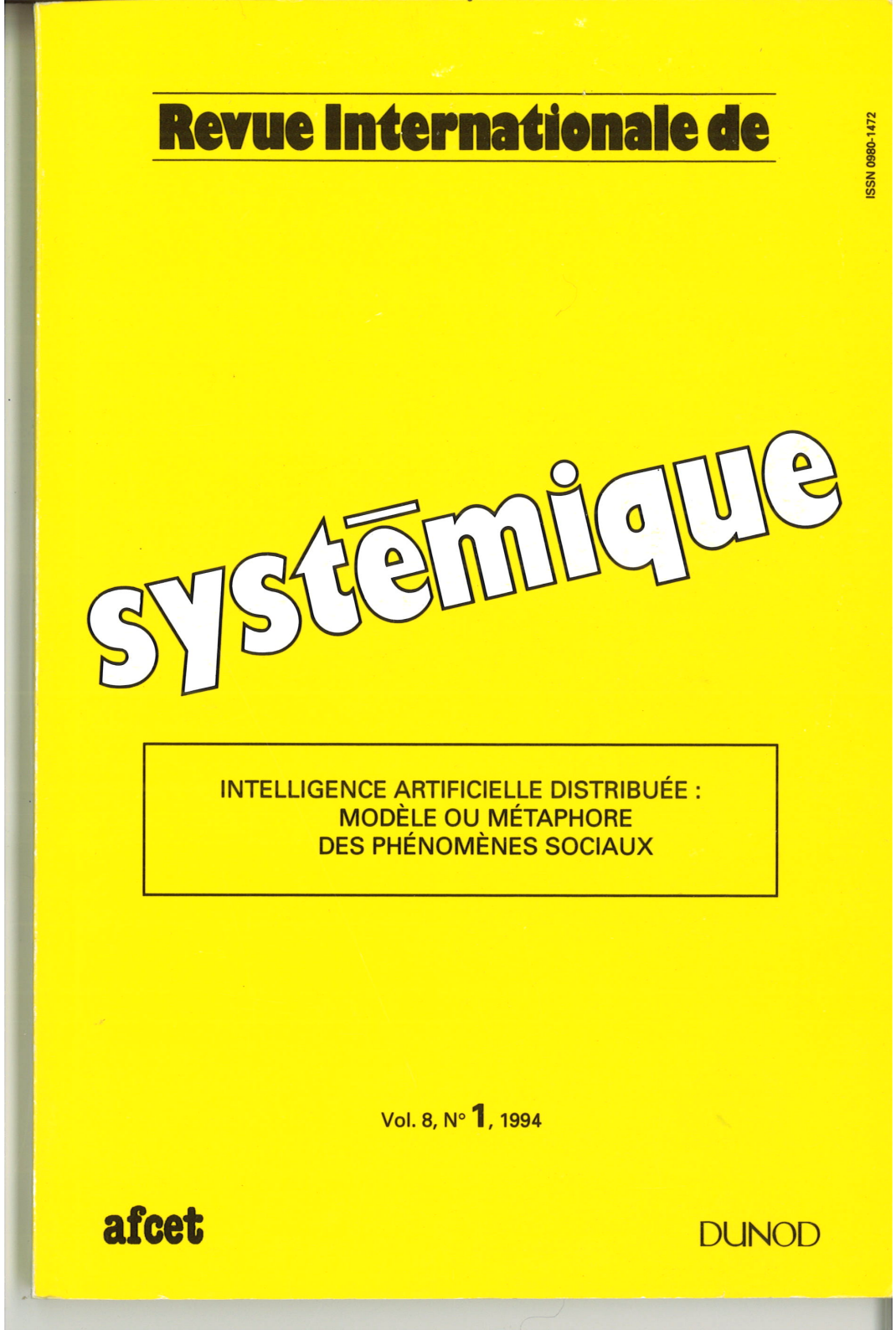

AFSCET

Revue Internationale de systēmique

Revue
Internationale
de Sytémique

volume 08, numéro 1, pages 51 - 63, 1994

Intelligence artificielle distribuée : entre simulation et métaphore

Bruno Blachimont

Numérisation Afscet, janvier 2016.

Creative Commons

C. BRASSAC, *Théorie des actes de langage et IAD, Journée « Systèmes Multi-Agents » du PRC-GDR Intelligence artificielle*, Montpellier, 1993.

P. CANNIZZO-PUYET, B. CARPUAT, M. P. GLEIZES, P. GLIZE, *TELEMAC : système multi-agent en télémédecine*, Convention Intelligence Artificielle, Paris, 1991.

P. R. COHEN, H. J. LEVESQUE, *Rational interaction as the basis for communication*, Intentions in communication P. R. COHEN, J. MORGAN, M. E. POLLACK Eds, MITPress, 1990.

A. DROGOUL, B. CORBARA, D. FRESNEAU, *Applying ethomodelling to social organization in ants, Biology and evolution of social insects,* J. BILLEN Ed., Leuven University Press, 1992.

P. FALZON, *Ergonomie cognitive du dialogue*, Presses universitaires de Grenoble, 1989.

J. FERBER, E. JACOPIN, *The framework of ECO-Problem Solving*, Decentralized AI N° 2, North Holland, 1991.

J. J. GABARRO, *The development of working relationships*, Intellectual teamwork, J. GALEGHER, R. E. KRAUT Eds, L. Erlbaum Associates, 1990.

M. R. GENESERETH, M. L. GINSBERG, J. S. ROSENSCHEIN, *Cooperation without communication*, *AAAI 86*, 5th international conference on AI, Philadelphia, Volume 1, p. 51-57.

M. P. GLEIZES, P. GLIZE, S. TROUILHET, *Social knowledge abilities of an agent*, Expersys 93, Editions IITT, 1993.

P. GMYTRASIEWICZ, E. DURFEE, D. WEHE, *The utility of communication in Coordinating, Intelligent Agents,* AAAI 91 July 14-19, p. 166-172.

H. P. GRICE, Logic and conversation, *Syntax and semantics*, Vol. 3, P. COLE et J. L. MORGAN Eds, N. Y. Academic Press, 1975.

G. SABAH, *CARAMEL: A flexible model for interaction between the cognitive processes underlying natural language understanding*, Sixth Israeli Conference on AI, Vision and Pattern Recognition, Tel Aviv, 1989.

J. R. SEARLE, *A taxonomy of illocutionary acts, Expression and Meaning, Studies in the Theory of Speech Acts,* Cambridge University Press, 1979, chapter 1, p. 1-29.

H. A. SIMON, Sur la complexité des systèmes complexes, *Revue internationale de systémique*, Vol. 4, n° 2, 1990.

S. TROUILHET, *Représentation et traitement des connaissances sociales chez l'agent : Application à l'environnement multi-agent Synergic*, Thèse de Doctorat de l'Université Paul Sabatier, Toulouse, 8 juillet 1993.

K. ZEGHAL, *Un modèle de coordination d'actions réactive appliqué au trafic aérien*, Premières journées francophones IAD & SMA, Toulouse 7-8 avril 1993.

REVUE INTERNATIONALE DE SYSTÉMIQUE
Vol. 8, N° 1, 1994, pp. 51 à 63

INTELLIGENCE ARTIFICIELLE DISTRIBUÉE : ENTRE SIMULATION ET MÉTAPHORE

Bruno BACHIMONT [1]

Résumé
Tant l'intelligence artificielle (IA) que l'intelligence artificielle distribuée (IAD) reposent sur l'hypothèse que la connaissance se laisse représenter symboliquement, et que la réflexion d'un agent rationel (IA) ou la communication dans une société d'agents (IAD) sont adéquatement reproduites par la manipulation de ces représentations. Au lieu d'être de simples simulations technologiques, l'IA et l'IAD se revendiquent comme de véritables explications scientifiques de la cognition, qu'elle soit sociale (IAD) ou individuelle (IA). Mais si, comme nous l'argumentons, la représentation symbolique n'est pas une explication de la connaissance, mais peut tout au plus, se laisser interpréter à titre métaphorique comme une connaissance, tant l'IA que l'IAD ne peuvent prétendre être une activité scientifique. Mais, le statut de simples techniques de simulation qui alors leur échoit est insuffisant pour rendre compte à la fois de la nature de leurs modèles et de leur fécondité. Notre théorie de l'artefacture veut montrer que c'est la métaphore des représentations comme connaissances qui permet à l'IA et à l'IAD de posséder une telle fécondité, à condition qu'on la prenne comme telle au lieu d'en faire une explication scientifique de la cognition.

Abstract
Artificial intelligence (AI) as well as distributed artificial intelligence (IAD) rely upon the hypothese that knowledge can be symbolically represented, and that manipulating symbolic representations can account for the cognition of a rational agent (AI) or a society of agents (IAD). Such an account aims at being not a simple technological simulation but rather a real scientific explanation of the social (IAD) or individual (AI) cognition. We claim in this paper that symbolic representations are not a scientific account of knowledge, but can at most be metaphorically interpreted as knowledge. Then, AI and DAI are only technological simulation rather than real science. But, in this case, the real nature of their models as well as their usefulness become incomprehensible. Our theory of artefacture is a tentative to show that the metaphor of symbolic

1. DIAM-Département de Biomathématiques, 91, boulevard de l'Hôpital, 75643 Paris Cedex 13.

Rev. intern. systémique. 0980-1472 Vol. 8/94/01/ $ 4.00/© Afcet Gauthier-Villars

representations as knowledge is what enables to understand their deep nature, providing that metaphor is understood as such and not as a scientific explanation of cognition.

De manière schématique, on peut dire que l'Intelligence Artificielle Distribuée (IAD) modélise des sociétés d'agents alors que l'Intelligence Artificielle (IA) plus classique modélise un seul agent rationnel. Par ce recours au distribué, l'IAD est souvent présentée comme une solution aux problèmes tant théoriques que pratiques rencontrés en IA. En posant que la connaissance est de nature sociale plutôt que psychologique, l'IAD pense mieux appréhender ce que sont les connaissances (meilleure adéquation épistémologique) et construire des systèmes plus efficaces (meilleure compétence technologique).

Contrairement à ce qui est habituellement avancé, nous argumentons dans cet article que l'IAD rencontre les mêmes problèmes épistémologiques que l'IA car elle ne s'attaque pas à sa dimension proprement problématique, à savoir l'articulation des représentations aux connaissances. Comme elle, elle voit dans ses programmes des connaissances et des agents rationnels, usant ainsi de métaphores sans réel fondement, quoique fécondes en pratique. Prenant la métaphore pour des faits scientifiques, l'IA, distribuée ou non, possède un statut épistémologique confus qui ne peut être précisé qu'en clarifiant le rôle joué par les métaphores dans ses modèles. Notre théorie de l'artefacture tente d'y contribuer en montrant d'une part comment manipuler la métaphore sans se prendre au piège de la constituer comme explication, et d'autre part comment l'IA et surtout l'IAD se prêtent à ce schéma épistémologique et y trouvent une fécondité originale, rendant compte de leur activité tout en évitant les écueils épistémologiques.

I. DE L'IA A L'IAD

I.1. Niveau des connaissances

Depuis ses origines dans les années cinquante, l'intelligence artificielle affiche l'objectif de concevoir et réaliser des systèmes informatiques exhibant un comportement reconnu intelligent dans la résolution de certains problèmes ou l'effectuation de certaines tâches. Un tel objectif laisse entrevoir la possibilité de résoudre des problèmes pour lesquels les méthodes automatiques manquent, *i.e.* pour lesquels il faut faire preuve d'intelligence pour les résoudre. Par exemple, la planification, le raisonnement, la conception, sont des activités reconnues complexes, et la personne réalisant ces tâches se voit attribuer le qualificatif « d'intelligent », ce qui ne serait certes pas le cas si elle se contentait d'appliquer sans réfléchir une procédure bien établie.

Le but de l'IA est donc de concevoir des systèmes informatiques capables d'avoir ces activités dans la mesure où ces dernières se présentent comme des heuristiques, des réflexions, mais non comme le déroulement de procédures automatisées : il semble qu'une pensée consciente doive délibérer pour mener à bien la tâche à traiter. Cette pensée délibératrice porte sur des objets et les relations les unissant d'un univers de référence dans lequel le problème à résoudre ou la tâche à effectuer sont définis : *e.g.* des buts, plans, ateliers de conception, etc.

L'absence de méthode standard pour le traitement de ces tâches, la nécessité d'une délibération consciente interdisent de définir ces systèmes par les programmes qui les constituent même si, nécessairement, ils reposent sur eux. Bref, s'il faut *réaliser* le système par des programmes, on ne peut le *définir*, le caractériser par des programmes. En effet, puisqu'on ne connaît pas la procédure à suivre, il n'est pas possible de déterminer si le programme est correct ou non d'un point de vue informatique. Pour cela, il faut le voir à l'œuvre et le considérer en contexte : si le programme se manifeste en reproduisant la même démarche délibérative, ou, du moins, se manifeste par une considération pertinente de l'univers de référence, alors, le programme pourra être reconnu comme une réalisation informatique correcte de la résolution.

Pour reconnaître que le système fonctionne correctement, il faut retrouver dans le déroulement du programme les données du problème, les connaissances considérées, le raisonnement mené. Il est donc nécessaire de pouvoir associer au programme les objets, relations, *i.e.* les connaissances, de l'univers de référence. Si le programme est une représentation symbolique de l'univers de référence, alors il est possible d'assurer cette correspondance, l'univers de référence devenant ainsi la réalité dénotée par le programme qui en est la représentation symbolique. C'est pourquoi l'on aura non seulement des systèmes informatiques, mais des systèmes à base de connaissances (sbc), *i.e.* des programmes considérés depuis l'interprétation qu'ils possèdent dans l'univers de référence.

Par conséquent, les sbc se caractérisent par le fait que les tâches qu'ils doivent réaliser ne se définissent non pas par les programmes qui les accomplissent mais par les connaissances que ces programmes expriment et la signification qu'ils possèdent dans le contexte réel de la tâche. Les programmes sont interprétés en termes de connaissances : ce sont des états sémantiquement interprétés ou états intentionnels car ils se rapportent à des états de choses du monde (Pylyshyn, 1984). Cela signifie que les représentations syntaxiques du programme possèdent des interprétations dénotationnelles, *i.e.* des références, dans le monde et que le système est décrit à partir de ce que les programmes désignent et pas seulement à partir de leur structure syntaxique formelle. A la suite de Newell (Newell, 1982), on dit

que les tâches que les sbc doivent accomplir se définissent donc *au niveau des connaissances* indépendamment des formalismes informatiques de représentation choisis pour les traiter *au niveau des programmes.*

I.2. Les agents rationnels

Cette description en termes de connaissances des programmes mis en œuvre dans un sbc revient à considérer ce dernier comme un *agent rationnel* résolvant un problème en raisonnant à l'aide de ses connaissances. On dit en effet que les programmes s'interprètent comme des croyances, buts et actions manipulés conformément à un principe de rationalité, selon lequel un agent accomplit des actions pour réaliser ses buts en fonction de ses croyances. Un sbc est donc réussi s'il est possible de reconnaître dans ses opérations informatiques les agissements d'un agent rationnel traitant un problème à l'aide de ses connaissances. C'est d'ailleurs dans cette mesure que l'on est capable de dire si le problème est effectivement résolu : en effet, ce dernier est formulé en termes de connaissances, et doit être traité et résolu dans ces mêmes termes. Pour comprendre ou reconnaître la solution proposée par le système, il faudra interpréter le résultat informatique comme un état intentionnel.

I.3. Agent rationnel et cognitivisme computationnel

A ce niveau se pose la question de savoir quel statut épistémologique le programme possède par rapport au comportement intentionnel visé. L'alternative est la simulation ou l'explication. Dans le premier cas, on dit que le programme simule la rationalité d'un agent, mais que le système informatique lui-même n'est pas un agent rationnel. Il lui manque pour cela les principes physiques ou biologiques responsables chez l'humain de sa rationalité. On dira alors avec Searle (Searle, 1981) qu'il s'agit d'IA faible. Dans le second cas, on dit que le programme est l'explication scientifique de la rationalité ; en d'autres termes, ce qui fait qu'un système informatique ou un humain sont rationnels, c'est le fait qu'ils possèdent chacun un programme responsable de leurs activités.

Le cognitivisme computationnel [1] adopte ce second point de vue : il considère que les programmes sont des explications à la base du comportement rationnel. Si on comprend un système informatique ou un humain en lui associant des connaissances, c'est parce qu'il possède un programme. Par conséquent, l'IA devient la science de la rationalité computationnelle, dans la mesure où elle

expérimente des modèles symboliques pour expliquer le comportement rationnel. L'IA se distingue alors peu de la psychologie cognitive puisqu'elles partagent l'objet d'étude et la méthode.

I.4. La société d'agents

Le principal grief reproché aux sbc construit sur le modèle de l'agent rationnel est qu'ils ne peuvent appréhender qu'une petite partie des connaissances disponibles, pour la raison qu'une bonne partie de ces dernières résident dans l'interaction sociale entre les individus et non dans la réflexion psychologique individuelle d'un seul agent. Cela signifie que si l'on veut les représenter dans un système, il faudra reproduire une société de plusieurs agents délibérant entre eux et échangeant connaissances et informations. On peut comprendre cela en remarquant que les connaissances considérées en IA sont véhiculées par la langue naturelle (cf. *infra*) et que cette dernière est essentiellement destinée à la communication sociale d'agents, et pas seulement à la réflexion interne d'un seul esprit.

Outre le fait de mieux pouvoir tenir compte des connaissances, l'approche multi-agents permet d'éviter les problèmes classiques des sbc mono-agent, à savoir la fragilité et la rigidité. La fragilité correspond au fait que le contrôle de la résolution est rigide, non opportuniste, c'est-à-dire que la meilleure action n'est pas choisie à chaque instant de la résolution : si la solution est atteinte, elle ne l'est pas de manière efficace. La rigidité n'est pas indépendante de la fragilité, puisqu'un contrôle inefficace peut conduire à l'échec dû à l'explosion combinatoire.

I.5. Société d'agents et cognitivisme computationnel

Tant l'IAD que l'IA voient dans les systèmes construits les véritables explications scientifiques au fait qu'un agent rationnel possède des connaissances. Ainsi, un sbc sera la théorie computationnelle expliquant comment un agent (*e.g.* un humain) possède des connaissances. Une société d'agents pour sa part pourra être l'explication du comportement social ou l'explication de la cognition d'un agent rationnel unique, sa rationalité étant répartie sur plusieurs sous-agents conçus à leur tour comme des agents uniques. De la même manière que l'étude d'un agent rationnel donnait lieu à une psychologie cognitive, on obtient avec l'étude d'une société d'agents une sociologie cognitive (Gasser, 1991). Cette dernière veut expliquer et étudier les interactions sociales par les échanges entre des programmes construits comme des agents rationnels au sens de Newell. Si bien que les échanges avec la sociologie (au sens habituel, celle s'occupant des humains) s'effectue à

double sens : d'une part cette dernière devient une inspiratrice pour concevoir des modèles cognitifs et d'autre part, des expériences computationnelles sont menées pour expliquer les interactions entre agents humains. L'IAD s'inscrit donc dans une perspective cognitiviste puisqu'elle prétend naturaliser le social comme l'IA prétendait naturaliser le psychologique. La méthode employée restant computationnelle, il s'agit toujours bien du cognitivisme computationnel.

II. L'ARTEFACTURE

Le programme d'étude pour l'IA et l'IAD semble donc tracé : l'explication du cognitif par les programmes. Sciences à part entière, elles en possèdent les caractéristiques standard : elles édifient des théories, effectuent des expériences, étudient la nature. Mais toutes ces thèses sont suspendues à l'affirmation qu'une connaissance d'un agent puisse être prise en compte par une représentation symbolique. Or, cela nous semble contestable, comme nous allons l'argumenter, même si, du fait de la briéveté de cet article, nous devons renvoyer le lecteur à nos autres travaux (Bachimont, 1992 ; Bachimont, 1993 ; Bachimont, 1994) pour le détail de ces discussions.

II.1. Connaissance et représentation

De manière générale, l'IA peut être définie comme la résolution (assistée ou automatique) de problèmes effectuée par la modélisation des connaissances dont on dispose sur eux. Alors qu'une approche informatique s'attache à *coder numériquement* une théorie formelle de la résolution du problème (*e.g.* codage informatique d'une modélisation mathématique d'un processus physique), une approche de type IA s'attache à *représenter symboliquement* les connaissances dont on dispose sur le problème (*e.g.* représentation symbolique de ce que l'on sait sur le problème). Comme le rappelle Newell (Newell, 1980), ces deux façons de voir sont équivalentes du point de vue formel, car elles se ramènent toutes deux à du calcul. Mais, passer d'une conception où l'ordinateur est un *number cruncher* à celle où il est un *symbol processor* change profondément la manière dont on veut résoudre les problèmes posés.

Par exemple, alors que le codage informatique permet de simuler numériquement une théorie physico-chimique du mécanisme cardiaque pour résoudre un problème de diagnostic, l'approche de type IA représentera les connaissances du médecin sur le diagnostic. Dans le premier cas, on utilise un modèle quantitatif qu'une

simulation numérique permet d'utiliser, et dans le second, on utilise un modèle qualitatif qu'une représentation symbolique permet d'opérationaliser. Cela revient à constater que la distinction codage numérique *versus* représentation symbolique est une distinction portant sur le type des connaissances prises en compte dans les modèles et non sur la contrepartie informatique obtenue. C'est ainsi qu'on opposera les connaissances quantitatives *scientifiques* [2] exprimées dans les langages formalisés fournis par les mathématiques aux connaissances qualitatives *phénoménologiques* exprimées informellement dans la langue naturelle. Dans le premier cas, on s'intéresse au monde tel qu'il peut être défini par une description ou modèle formel, en tant qu'il respecte les lois formelles de la description. On résout alors les problèmes en codant numériquement les connaissances scientifiques du modèle formel. Dans le second cas, on s'intéresse au monde et à sa structure qualitative telle qu'une description linguistique, en langue naturelle, ou modèle qualitatif, le décrit. On résout alors les problèmes en représentant symboliquement les connaissances phénoménologiques du modèle qualitatif.

Mais s'il est bien connu que le codage numérique des lois physiques en conserve le sens mathématique et physique, on peut se demander ce qui permet d'affirmer que les représentations symboliques des connaissances en conservent le sens phénoménologique. Nous avons vu plus haut que, pour le cognitivisme computationnel, le niveau des programmes est considéré comme l'explication scientifique des raisonnements portant sur des connaissances : si on résout certaines tâches ou problèmes en termes de connaissances, c'est que le système ainsi compris est computationnel. C'est une explication psychologiste de la connaissance puisque le computationnel est la cause mentale (esprit = programme) du comportement rationnel compris en termes de connaissances. Ce psychologisme s'articule à un formalisme explicitant le rapport entre le programme et son interprétation : cette dernière est la dénotation, la référence du programme dans le monde. Par conséquent, pour le cognitivisme, l'IA construit des modèles qualitatifs (Clancey, 1989) formels du monde où la réalité mondaine est la dénotation, la référence, des représentations symboliques. La représentation reflète par sa structure syntaxique la structure qualitative du monde visée par les connaissances phénoménologiques et c'est pour cela que le modèle symbolique formel de l'IA conserve le sens des connaissances phénoménologiques du modèle qualicatif informel.

II.2. De l'explication à la métaphore

Il faut, pour que cette solution soit valable, que le contenu phénoménologique des connaissances, telles qu'elles sont formulées dans la langue naturelle, soit un

contenu dénotationnel. Les représentations formelles seraient manipulées en fonction de leur forme syntaxique et recevraient un contenu sémantique en fonction de cette forme, selon une sémantique formelle. Or ceci est impossible. La sémantique de la langue n'est pas une sémantique formelle (Rastier, 1991) : les représentations de la langue, les connaissances phénoménologiques, sont manipulées en fonction de leur signification, de leur sens, et il n'est pas possible de faire dépendre ce sens de la forme syntaxique. Il n'y a pas de sémantique formelle de la langue ou des connaissances. Le cognitivisme (Haugeland, 1985) estime son projet possible dans la mesure où le sens phénoménologique se ramène au sens formel : s'il remarque bien que les connaissances sont manipulées en fonction de leur sens, c'est pour conclure que cette manipulation se ramène à de la sémantique formelle, ce qui est erroné à notre sens. L'utilisation et la mise en œuvre des connaissances ne correspondent pas à l'exploitation formelle de représentations syntaxiques.

Or, si l'on refuse la solution cognitiviste, les modèles édifiés en IA et IAD ne peuvent plus être des explications scientifiques. Si le sens des représentations n'est pas leur contenu dénotationnel, il n'est pas légitime de rapporter la compréhension des systèmes en termes de connaissances à des programmes. L'explication devient métaphore : ce n'est qu'à titre métaphorique que le système computationnel a des buts ou des croyances, même s'il existe un rapport dénotationnel entre les états du programme des états du monde. Car, comme nous l'avons dit, un tel rapport dénotationnel ne suffit pas à faire des programmes une explication scientifique des connaissances. Si bien que l'on se paye de mots, quand on parle d'agent rationnel, de société d'agents, de connaissances en IA et IAD. Le cognitivisme prend donc les métaphores qu'il utilise dans ses programmes pour des observations scientifiques, en croyant que le mot suffit à créer le phénomène.

Que reste-t-il à ces disciplines, dans cette perspective dévastatrice ? La situation est complexe car elles ne peuvent être ni des sciences, comme nous venons de le voir, ni de simples technologies informatiques. Prises comme des sciences, on s'installe alors dans la métaphore, où l'on prend la simulation pour le phénomène. Comme technologies informatiques, elles pratiquent de simples simulations de modèles mis au point dans le cadre d'autres disciplines, où le rapport au réel s'installe autrement que dans la métaphore. Pourtant, ce statut technologique n'est pas satisfaisant dans la mesure où il devient dès lors impossible de distinguer les programmes de l'IA et de l'IAD des réalisations des autres disciplines utilisant l'informatique comme outil : considérer les programmes comme des explications scientifiques de connaissances permet de construire des systèmes originaux et de traiter des problèmes jusque là inaccessibles au traitement informatique. En

supprimant cette dimension, leur intérêt devient alors incompréhensible et l'on jette le bébé avec l'eau du bain. Par conséquent, l'IA et l'IAD doivent se construire sur la métaphore pour traiter des connaissances et doivent s'en extraire pour être une science.

Notre proposition de l'IA comme artefacture consiste non pas à éviter la métaphore, car c'est à travers elle que s'exprime l'intérêt des systèmes conçus, mais à la thématiser comme telle, au lieu de la maquiller en fait scientifique. Cette attitude permet de conserver toute l'originalité des approches IA et IAD en évitant l'écueil d'en faire des sciences qu'elles ne sont pas. Pour étayer et éclairer notre propos, nous présentons brièvement notre théorie de l'artefacture.

II.3. Bi-modélisation et validation

L'artefacture considère l'IA comme une entreprise de bi-modélisation des connaissances phénoménologiques qu'une étape ultime de validation sémiotique vient justifier.

Les connaissances phénoménologiques sont des connaissances portant sur la structure qualitative du monde telle qu'elle s'exprime dans les descriptions linguistiques. Les connaissances phénoménologiques sont donc des représentations interprétées véhiculant un sens analogue à celui que l'on trouve dans la langue naturelle. L'expression des connaissances phénoménologiques en représentations interprétées s'effectue à l'aide d'un *langage de représentation*. La langue naturelle est le langage de représentation par excellence, par définition même des connaissances phénoménologiques. Elle se caractérise par le fait que ses expressions se manipulent en fonction de ce qui est exprimé, *i.e.* en fonction de leur signification, et non en fonction de la forme syntaxique de ces expressions. Par extension, tout langage utilisé pour exprimer des connaissances où les expressions obtenues sont manipulées en fonction de leur sens et non en fonction de leur forme syntaxique constitue un langage de représentation. C'est ainsi que les langages de l'IA comme Lisp ou Prolog peuvent être considérés comme des langages de représentation s'ils sont utilisés pour exprimer des connaissances. Mais, comme on va le voir, cette utilisation ne coïncide pas avec l'exploitation informatique qui en est faite.

L'expression des connaissances dans un langage de représentation permet d'obtenir des *structures de représentation*. Ces structures caractérisent les connaissances phénoménologiques que l'on veut prendre en compte dans la réalisation d'un sbc. Exprimer les connaissances dans un langage de représentation correspond donc à une étape de *modélisation sémiotique* : une connaissance

phénoménologique étant une description linguistique d'objets structurant qualitativement le monde, son étude ressortit donc aux sciences sociales en général et sémiotiques en particulier (Rastier, 1987 ; Rastier, 1991). Le but d'une telle modélisation est de caractériser les connaissances et leur signification dans leur contexte d'utilisation. Il s'agit de savoir, lors de cette étape, ce que les mots, locutions, concepts, etc., utilisés par les experts et/ou futurs utilisateurs signifient, ce qu'ils valent pour eux. Par exemple, les concepts de signes, syndromes, diagnostic, manifestation, possèdent une acception particulière en médecine qu'il faut élucider avant et indépendamment de toute mise en œuvre informatique. On dira donc que la modélisation sémiotique détermine le *contenu sémiotique* des structures de représentation exprimées dans le langage de représentation.

Pour prendre en compte ces connaissances décrites dans un langage de représentation, il faut disposer d'une contrepartie informatique, obtenue lors d'une étape de *modélisation computationnelle*. La modélisation s'effectue par la représentation dans des structures symboliques informatiques des structures de représentations dégagées lors de la modélisation sémiotique. Les structures symboliques sont exprimées dans un *langage de programmation* et sont formelles, *i.e.* ininterprétées. Cela signifie que l'ordinateur manipule ces structures symboliques, ou *structures de programmation*, uniquement en fonction de leur forme syntaxique et indépendamment de leur signification et qu'il leur attache un *contenu computationnel* qui se réduit à l'utilisation qu'il en fait. La modélisation computationnelle est une *construction* d'objets informatiques : elle ne vise pas à dénoter des objets déjà présents, mais à les construire dans la dynamique du calcul symbolique. Les contenus sémiotiques sont modélisés par le contenu computationnel associé aux structures de programmation.

L'IA peut donc se définir comme la modélisation de structures de représentation, exprimées dans un langage de représentation (*e.g.* la langue naturelle), par des structures de programmation, exprimées dans un langage de programmation (*e.g.* Lisp). Modéliser au niveau des connaissances, c'est exprimer les connaissances dans un langage de représentation et les opérationaliser, c'est les modéliser dans un langage de programmation au niveau symbolique. Si on prend l'exemple des règles de production, on constate qu'on utilise les *mêmes* règles comme structures de représentation et de programmation. Manipulées par l'expert, ces règles véhiculent le sens qu'il leur confère et à travers lequel il les utilise. Ces règles *représentent* des connaissances, elles les expriment. Ces mêmes règles, manipulées par le moteur d'inférence, deviennent des structures de programmation et sont alors des instructions. Elles possèdent alors un sens computationnel (l'instruction) qui *n'est pas* le sens qui leur est conféré par l'utilisateur en tant que structure de

représentation. Le but de la modélisation est alors d'approximer le sens de la représentation par le sens computationnel. Il faut bien se rappeler qu'un langage de programmation est un modèle de calculabilité (il a été prouvé que les règles de production en sont un), et donc qu'il calcule, mais ne représente rien. L'appellation de *langage de représentation* pour des langages comme KRL ou KL-ONE est un abus de langage (!) et purement métaphorique car, si ces langages sont bien utilisés comme des langages de représentation, ils sont conçus comme des langages de programmation.

Le but d'une telle bi-modélisation est donc d'étudier dans quelle mesure la modélisation computationnelle permettant à l'ingénieur de construire un système permet à l'utilisateur de retrouver dans le système et ses actions les objets phénoménologiques qu'il vise à travers ses connaissances, caractérisées lors de la modélisation sémiotique. Le système et la modélisation computationnelle sont ainsi un support devant être interprété par l'utilisateur. La *validation sémiotique* d'un sbc consiste donc à évaluer la capacité de l'artefact construit à susciter les bonnes interprétations, à réveiller les bonnes connaissances (*i.e.* pertinentes) chez l'utilisateur lui permettant de constituer phénoménologiquement les objets phénoménologiques dont se composent les raisonnements qu'il pratique pour résoudre ses problèmes.

II.4. Artefacture et IAD

L'artefacture thématise le fait que la métaphore est le guide sémiotique pour concevoir et construire des systèmes informatiques. Mais, si le système construit fournit la matière à une interprétation sémiotique, à la métaphore, il ne sera jamais l'explication scientifique de la cible de l'interprétation : si on interprète un système par des buts et des croyances, ce n'est pour cela que les programmes sont l'explication de ces buts et croyances. De même pour l'IAD, où il n'y a de société d'agents que dans l'esprit du concepteur. L'IAD ne renouvelle en rien l'IA d'un point de vue scientifique.

En revanche, du point de vue de l'artefacture, il nous semble que l'IAD enrichisse considérablement l'IA classique en développant la métaphore. En effet, l'artefacture opère à deux niveaux. A un premier niveau, il doit être possible de retrouver dans le système les connaissances caractérisées lors de la modélisation sémiotique. A un deuxième niveau, le système comme métaphore permet de constituer de nouvelles connaissances. Son fonctionnement, échappant à notre compréhension scientifique du fait de sa complexité doit être expliqué au niveau

des connaissances. Mais, rebelle aux métaphores déjà éprouvées, il suscite de nouvelles métaphores pour l'appréhender et le comprendre. On obtient alors de nouvelles connaissances phénoménologiques, qui peuvent, comme guide sémiotique, conduire à de nouvelles réalisations scientifiques.

Dans cette dimension constitutive des connaissances, l'IAD adopte des modèles informatiques complexes excédant la compétence interprétative des métaphores de l'IA classique (basiquement celle de l'agent rationnel). Il faut donc de nouvelles métaphores pour les comprendre, tirées de la société d'agents. C'est cette création de connaissances et métaphores qui donne l'illusion d'une science expérimentale où l'on observe des phénomènes : il y a effectivement de nouvelles productions informatiques nécessitant de nouvelles compréhensions, de nouvelles interprétations, de nouvelles métaphores. Il n'y a donc pas d'observations de phénomènes donnant lieu à des explications scientifiques (toujours pour la raison que le comprendre métaphorique n'est pas l'expliquer scientifique, que la sémantique des connaissances n'est pas dénotationnelle), mais des comportements à interpréter.

La floraison des études et « expérimentations » en IAD prouve à la fois la force de l'illusion (de nature transcendantale : inhérente à notre manière de penser, mais surmontable par la critique épistémologique) et la puissance des réalisations à motiver des interprétations sans cesse renouvelées. A cet égard, l'IAD est certainemnet une artefacture réussie, même si elle n'est pas toujours pensée comme telle.

III. CONCLUSION

Disons le tout net, ni l'IA ni l'IAD ne sont des sciences. Elles sont à la fois bien moins et bien plus que des sciences. L'artefacture est la proposition théorique de considérer la métaphore comme le lieu essentiel des recherches en IA et en IAD. En effet, il s'agit, non de prendre des métaphores pour des réalités, mais de considérer ces métaphores comme des intermédiaires interprétatifs nécessaires d'une part pour reconnaître dans les systèmes les connaissances qu'ils doivent utiliser et d'autre part pour être capable de comprendre leur fonctionnement et donc de les utiliser. Il faut donc thématiser explicitement le recours aux métaphores et comprendre comment le système se prête à une telle interprétation sans posséder pour autant les propriétés dont on l'investit pour le comprendre. La métaphore n'est pas démiurgique, comme l'ont oublié l'IA et l'IAD pré-artefacturelles.

Notes et références

1. Nous appelons cognitivisme toute tentative de naturaliser/matérialiser l'esprit. Quand ce dernier est appréhendé comme un ordinateur, on a un cognitivisme computationnel. Quand l'esprit correspond à des invariants cognitifs émergeant des dynamiques physiques sous-jacentes d'un substrat, on a un cognitivisme physique (Petitot, 1992).

2. Nous utilisons la notion de *science* non pour établir une hiérarchie des savoirs, mais pour repérer des modes de fonctionnement des connaissances. Une science est un savoir où les connaissances fonctionnent comme un calcul et vise une réalité appréhendée par la mesure expérimentale homogène au calcul. D'autres types de savoir (*e.g.* sciences sociales) existent, et s'ils ne reçoivent pas le titre de science ici, ce n'est pas pour les dévaloriser, mais pour les distinguer. Ainsi, l'IA et l'IAD ne sont pas pour nous des sciences car elles ne fonctionnent pas comme des sciences (cf. *infra*). Elles constituent des savoirs tout aussi rationnels et véridiques que ce que nous appelons science, mais selon une rationalité autre qui ne peut se réduire ni s'assimilier à la rationalité scientifique.

B. BACHIMONT, *Le contrôle dans les systèmes à base de connaissances. Contribution à l'épistémologie de l'intelligence artificielle. Langage, Raisonnement, Calcul*, Paris, Hermès, 1992.

B. BACHIMONT, Nature, culture et artefacture : la place de l'intelligence artificielle dans les sciences cognitives, *Intellectica, 17,* 1993, p. 213-238.

B. BACHIMONT, Connaissances et modèles en intelligence artificielle : entre dénotation et construction, in *Proceedings of the RFIA'94 Conference*, Paris, 1994, vol. 2, p. 541-552.

W.-J. CLANCEY, Viewing knowledge bases as qualitative models. *IEEE-Expert*, 1989.

L. GASSER, Social Conceptions of Knowledge and Actions: DAI Foundations and Open Systems Semantics, *Artificial Intelligence, 47*, 1991, p. 107-138.

J. HAUGELAND, *Artificial Intelligence & The Very Idea*, Cambridge Mass., The MIT Press, 1985.

A. NEWELL, Physical Symbol Systems, *Cognitive Science, 4*, 1980, p. 135-183.

A. NEWELL, The Knowledge Level, *Artificial Intelligence, 18*, 1982, p. 87-127.

J. PETITOT, *Physique du sens*, Editions du CNRS, 1992.

Z. W. PYLYSHYN, *Computation and cognition, Towards a Foundation for Cognitive Science*, Cambridge Mass., The MIT Press, 1984.

F. RASTIER, *Sémantique Interprétative*, Formes Sémiotiques, Paris, Presses Universitaires de France, 1987.

F. RASTIER, *Sémantique et Recherches Cognitives*, Formes Sémiotiques, Paris, Presses Universitaires de France, 1991.

J. SEARLE, Minds, brains and programs, in J. Haugeland, Ed., *Mind Design*, The MIT Press, 1981.