

L'intelligence artificielle et sa place dans les sciences de la cognition

Reproduction de l'article : V. Rialle (1996), L'intelligence artificielle et sa place dans les sciences de la cognition, Bulletin de l'Association Française de l'Intelligence Artificielle. n° 26, pp. 8-12.

Table des matières

I - Introduction.....	1
II - Une histoire enchâssée dans celle des sciences de la cognition.....	2
1. Collusion entre l'IA et les sciences de la cognition.....	2
2. L'expression intelligence artificielle.....	3
III - L'Homme et son double, la Machine, ou le carrefour de l'IA et des SC.....	3
1. L'ordinateur est insurpassable.....	3
2. Mais cette base pragmatique.....	4
IV - IA «classique», connexionnisme, émergentisme.....	4
1. Il a été de bon ton.....	4
a) Le front des conflits.....	4
b) Il résulte des meilleures.....	4
2. La tombée en désuétude.....	5
V - L'IA et la psychologie cognitive.....	5
VI - L'IA forte : de l'homo au robot sapiens.....	7
VII - L'IA et la philosophie de l'esprit : l'esprit est-il computationnel ?.....	8
1. L'IA forte représente.....	8
2. Le jeu de la réflexion.....	9
VIII - Conclusion.....	9
IX - Bibliographie.....	10

I - Introduction

Quelle est la place de l'Intelligence Artificielle dans la mouvance des sciences de la cognition (SC) ? Une telle question inclut tout en la dépassant celle de la définition même de l'IA, mille fois rencontrée. Ce texte ne prétend pas donner une vue exhaustive de l'IA, vaste sujet qui se laisse difficilement "encapsuler" dans un article de synthèse, mais tente simplement de présenter de manière concise, donc nécessairement non exhaustive, et pédagogique le panorama des relations que l'IA entretient avec les SC.

La collusion entre les sciences de la cognition et l'intelligence artificielle est très étroite. Mais l'IA cognitive n'est pas toute l'IA : un nombre important de chercheurs ne se reconnaît pas dans un contexte de sciences cognitives. D'un autre côté, l'IA permet à un très grand nombre de thèmes de s'exprimer : de la quasi-totalité des disciplines des sciences de la cognition à la poésie, la théologie, la psychanalyse, etc. C'est dire la prégnance de l'IA non seulement dans la sphère scientifique mais aussi dans de multiples dimensions de l'expérience humaine actuelle.

Commençons comme il se doit par un peu d'histoire.

II - Une histoire enchâssée dans celle des sciences de la cognition

1. Collusion entre l'IA et les sciences de la cognition

Pour montrer l'étroite collusion qui existe entre l'IA et les sciences de la cognition, rien de tel qu'un bond d'un demi-siècle en arrière, lorsque tout était encore en germe. Dans le creuset des années 40, et bien qu'une longue tradition scientifique remontant au moins à Aristote ait expliqué cette émergence, tout semblait naître en même temps : le concept de "machine universelle" avait été proposé par Turing en 1936 et 1937, qui n'aura de cesse d'étudier la possibilité théorique d'une machine intelligente. Le célèbre "Test de Turing" scellera à sa manière l'une des plus étroites relations qui existe aujourd'hui entre IA et SC, puisqu'il n'ambitionne rien moins que de fournir le moyen de décider si une machine peut être considérée comme "intelligente" au regard de l'homme, c'est-à-dire dotée de toutes les capacités de raisonnement, de décision et, le test insiste là-dessus, de ruse habituellement reconnues à l'intelligence humaine. Von Neumann de son côté, à qui l'on doit la paternité (disputée à Eckert et Mauchly) de l'architecture générale de l'ordinateur, publiera son dernier ouvrage sous le titre "The Computer and the Brain". D'où la constatation que ces deux illustres pères fondateurs, à qui l'on attribue couramment une part importante de la paternité de l'informatique dans son ensemble, semblaient avoir pour préoccupations fondamentales la conception d'une machine possédant tous les attributs d'une intelligence humaine. Ils n'étaient déjà pas seuls et allaient être suivis d'un grand nombre : en 1943, W. McCulloch (psychiatre d'origine) et W. Pitts proposent le premier modèle mathématique de fonctionnement du neurone et montrent la capacité d'un réseau constitué de tels objets formels à reproduire certaines fonctions logiques. Quelques années plus tard (1949), le psychologue canadien D.O. Hebb introduit l'apprentissage par ajustement des poids synaptiques, notion qui connaîtra le succès que l'on sait sous le terme d'apprentissage hebbien ou règle de Hebb. Norbert Wiener de son côté apporte les fondements d'une théorie générale de la commande et du contrôle — la cybernétique —, dont on reconnaît aujourd'hui la portée et l'influence qu'elle eut sur l'IA. C'est la « première cybernétique », celle des années 40-50, qui semble-t-il, constitue le premier berceau de ce qui deviendra plus tard les sciences de la cognition (cf. **Cybernétique**).

« La cybernétique aura : introduit la conceptualisation et le formalisme logico-mathématique dans les sciences du cerveau et du système nerveux ; conçu l'organisation des machines à traiter l'information et jeté les fondements de l'intelligence artificielle ; produit la "méta-science" des systèmes, laquelle a laissé son empreinte sur l'ensemble des sciences humaines et sociales, de la thérapie familiale à l'anthropologie culturelle en passant par l'économie, la théorie des jeux, la sociologie, les sciences du politique et bien d'autres ; fourni à point nommé à plusieurs "révolutions scientifiques" du XXe siècle, très diverses puisqu'elles vont de la biologie moléculaire à la relecture de Freud par Lacan, les métaphores dont elles avaient besoin pour marquer leur rupture par rapport à un ordre ancien ». ((Dupuis 1985 : 11), cité par (Pelissier et Tête 1995 : XVIII-XIX)).

Citation 1. Cybernétique

2. L'expression intelligence artificielle

Comme chacun sait, la date de naissance de l'expression intelligence artificielle est celle du séminaire d'été tenu en 1956 au Dartmouth College, États-Unis. Mais ce qu'il faut remarquer ici est le profil de ses organisateurs, premier noyau dur de la discipline, qui signait déjà une forte relation entre informatique et psychologie expérimentale : M. Minsky (psychologue), A. Newell (psychologue), H. A. Simon (psychologue, économiste), J. McCarthy (informaticien), C. Shannon (mathématicien).

Soulignons cependant que chez ces pionniers de l'IA, l'intelligence était essentiellement celle des raisonnements logiques et des processus d'inférence heuristique. Nous concevons aujourd'hui plus clairement que l'intelligence n'est pas que cela : elle inclut les processus perceptifs, l'apprentissage, l'auto-organisation, l'adaptation... et souvent y culmine. Ces notions étaient-elles concevables pour les pionniers de l'IA des années 40 ? Piaget, dès les années trente, en indiquait en tout cas la direction. La cybernétique était certainement sensible à ce point de vue, mais l'IA naissante n'en était pas là : l'intelligence à imiter était celle des "grosses têtes", savants, experts et logiciens ; pas encore celle qui se développe jour après jour chez l'enfant, ou encore l'intelligence collective et distribuée de nombreuses colonies animales.

Quelle qu'ait été l'histoire des conceptions de l'intelligence, tout s'est passé comme si dès les premières années, le projet de réaliser une machine telle que l'ordinateur avait, à côté de ses objectifs pratiques immédiats, celui de reproduire voire dépasser, l'intelligence humaine dans toutes ses dimensions. La suite sera une profusion de travaux théoriques et empiriques... et nombre de débats et de questionnements épistémologiques.

III - L'Homme et son double, la Machine, ou le carrefour de l'IA et des SC

Pour continuer de montrer la collusion entre l'IA et les SC, il n'est pas inutile de refaire ces parcours, quelque peu classiques mais bons à rappeler, qui vont de l'homme à la machine et réciproquement.

1. L'ordinateur est insurpassable

L'ordinateur est insurpassable pour toutes les opérations purement calculatoires, mais l'opérateur humain reste difficile à remplacer par une machine pour l'accomplissement d'un certain nombre d'actions demandant beaucoup de souplesse (e.g., réponses à un guichet de

renseignements, "fouille" dans des textes...), des tâches de spécialistes (application d'une expertise, traduction, interprétation, accomplissement d'actes complexes perceptifs et moteurs, ...) ou encore impliquant la vie, la justice ou l'éthique. Néanmoins certains objectifs, imposés notamment par l'essor technologique, ne peuvent être réalisés que par une machine suffisamment autonome : activités en milieux contaminés (nucléaires, biologiques, ...), en milieux inaccessibles (exploration/exploitation de planètes), activités de précision (e.g., robotique médicale), etc. Ces besoins exercent une pression permanente pour reculer les frontières de l'assumable par une machine. Mais la référence par excellence pour accomplir ces tâches reste la personne humaine, à travers notamment ses capacités de jugement en situation et de réaction face à l'imprévu (cf. le sauvetage d'Apollo 13). D'où l'importance prise ces dernières années par la notion d'ingénierie inverse, qui juxtapose deux questions :

Comment faisons-nous ? à laquelle tentent de répondre les sciences de la cognition.

Comment faire faire par une machine ? Dont la réponse est attendue d'une coopération entre diverses branches scientifiques et filières technologiques auxquelles concourent en priorité la robotique et l'IA.

2. Mais cette base pragmatique

Mais cette base pragmatique n'est pas neutre. Elle se déploie sur une toile de fond de déterminations biologiques, écologiques et politiques au sens fondamental des termes. L'explicitation de ces déterminations fait en partie l'objet des SC, notamment dans ses débats sur l'intentionnalité, la conscience et la subjectivité, et sur les rapports entre corps, cerveau, esprit, culture et champ social. Nous évoquerons, sans nous y attarder, ces débats par la voie des relations entre l'IA et la psychologie cognitive (chap. V), mais après un rappel au chapitre suivant, de l'opposition qui a prévalu il y a quelques années entre IA «classique» et connexionnisme.

IV - IA «classique», connexionnisme, émergentisme...

1. Il a été de bon ton

Il a été de bon ton, durant la période qui a suivi le regain du connexionnisme au début des années 80, d'opposer l'IA dite classique et le connexionnisme, comme si l'une de ces deux branches d'investigation pouvait prétendre modéliser à elle seule l'ensemble des processus cognitifs humains. Le débat a été inévitable et fructueux, puisque l'on parle aujourd'hui beaucoup plus de systèmes complexes, hybridant les approches ou distribuant les tâches à des modules spécialisés, sans distinction d'obédience théorique. Il est repris ici en tant que l'un des débats animés qui a pris corps au sein des SC (e.g., (Posner 1989)).

a) Le front des conflits

Le front des conflits tournait d'un côté autour de la rigidité du Système de Symbol Physique, et de l'autre côté autour de la difficulté posée par les réseaux de neurones à la compositionnalité des symboles et par suite aux opérations symboliques un tant soit peu complexes.

b) Il résulte des meilleures

Il résulte des meilleures analyses qu'aucune des deux approches ne saurait prétendre posséder les clés d'un modèle du fonctionnement cognitif humain, et que les oppositions se résorbent en une différence de niveaux de résolution (Lacharité 1994). Par ailleurs, la question de

l'inclusion ou de l'exclusion des réseaux neuronaux du champ de l'IA peut être considérée comme une affaire d'école plutôt classée (cf. **Ganascia**).

« (...) il est important de noter que d'un côté l'intelligence artificielle, dans sa version classique, ne comprend pas les réseaux neuronaux puisqu'elle s'est constituée en vue d'offrir une alternative à la cybernétique, mais que, d'un autre côté, dans sa version générique, l'intelligence artificielle comprend toutes les techniques susceptibles de doter les machines d'intelligence, et donc, en particulier les techniques promues par le connexionnisme. » (Ganascia 1993 : 211).

Citation 2. Ganascia

2. La tombée en désuétude

La tombée en désuétude des oppositions entre IA symbolique et connexionnisme a été accélérée par l'avancée de la réflexion sur le concept d'intelligence. Au centre de cette réflexion se trouvent moins la notion d'intelligence, quelque peu indéfinissable et par trop générale, que celles, plus affinées, d'intentionnalité, d'interprétation, d'émergence, d'autonomie, de vie artificielle, de grammaire cognitive, etc. Autant de facettes qui semblent évoquer plus fidèlement les caractères de l'intelligence que l'on constate aussi bien dans la nature en général que dans les activités proprement humaines. Ces dernières tiennent dans les SC le haut du pavé, et le défi consiste à comprendre l'activité rationnelle de l'homme en tant que partie intégrante d'une dynamique biologique et pulsionnelle d'un côté, d'une dynamique sociale, culturelle et historique de l'autre côté. Ces dynamiques sont combinées, contraintes ou amplifiées par une capacité presque hypertrophiée chez l'homme : celle de produire en permanence du sens, du contre-sens ou de la croyance par le langage et l'activité représentationnelle. Vaste programme... mais qui présente l'avantage de relativiser et donc de ramener une certaine modestie au sein des débats ponctuels entre telle ou telle école.

V - L'IA et la psychologie cognitive

Dans l'introduction de son ouvrage collectif consacré au débat entre l'IA et la psychologie cognitive, J.-F. Le Ny, nomme intelligence artificielle cognitive, la part de l'IA qui s'intéresse à l'étude du fonctionnement cognitif humain (cf. **Conception**) :

« Appelons « cognitive » cette conception ; elle peut a priori sembler plus agréable au psychologue. Mais elle n'est pas forcément plus confortable, ni pour le spécialiste de l'intelligence artificielle, ni, en définitive, pour nous. (...) Ce fonctionnement cognitif humain, encore faut-il d'abord le connaître et, cette fois-ci, au sens fort du terme. Nos collègues de l'intelligence artificielle cognitive, devenus par là nos cousins, comme je le disais en commençant, peuvent bien alors s'adresser aux spécialistes de la psychologie scientifique que nous sommes censés être, plutôt qu'à leur intuition subjective. Mais en quels termes répondrons-nous à leurs questions spécifiques ? » (Le Ny 1993, 13).

Citation 3. Conception

Il y souligne aussi à juste titre que cette IA cognitive n'est pas forcément la plus efficace en termes d'ingénierie (e.g., le principe du réacteur est plus efficace que celui du battement d'aile

pour construire un avion). De son côté, G. Tiberghien en trace succinctement le programme (cf. IA cognitiviste):

« (...) une partie seulement de l'IA s'intéresse aux phénomènes cognitifs et à la façon dont ceux-ci sont produits par le cerveau. L'IA « cognitiviste » est fondamentalement bionique :

a) elle s'inspire de certaines propriétés symboliques ou sub-symboliques de la cognition humaine afin de produire des logiciels simulant des propriétés similaires dans des environnements de résolution de problèmes particuliers ;

b) elle fournit également des formalismes qui sont des langages de description aptes à saisir certaines dimensions des représentations symboliques ou sub-symboliques (langage de prédicats, réseaux sémantiques, systèmes de production, langage objet, réseaux connexionnistes, systèmes hybrides) ;

c) enfin, l'IA offre des systèmes computationnels qui permettent de simuler, de façon précise, des modèles psychologiques ou neuropsychologiques ; cela oblige les chercheurs à définir des cahiers des charges rigoureux ce qui peut augmenter la portée heuristique de leurs tentatives de modélisation. » (Tiberghien 1993 : 822).

Citation 4. IA cognitiviste

Une phrase de S. Harnad est également significative de l'importance de l'IA pour le spécialiste en psychologie cognitive : «(...) en tant que psychologue cognitiviste, je m'intéresse à l'intelligence des machines uniquement dans la mesure où elle peut être considérée comme un modèle "réaliste" de l'intelligence humaine (ou de celle d'autres organismes)» (Harnad 1994 : 66).

Il ne faut cependant pas renverser le mouvement : que l'intelligence des machines intéresse la psychologie cognitive ne fait pas de la modélisation psycho-cognitive une branche de l'IA, même si cette modélisation vise une simulation sur ordinateur. Le but premier de la psychologie cognitive n'est pas de fabriquer des machines cognitives, mais d'étudier la cognition humaine. Il reste néanmoins que le spécialiste de l'IA et celui de la psychologie cognitive ont en partage, jusqu'à un certain point, la cognition humaine : « Nous explorons le territoire des activités cognitives humaines (mémoire et connaissance, calcul et raisonnement, langage et communication, perception et action) pour déterminer dans quelle mesure les manifestations de certaines d'entre elles peuvent être réalisées par une machine. » (Jorrand 1995).

Le projet de l'IA considérée comme une branche des sciences de la cognition consiste donc à comprendre l'intelligence pour la reproduire sous forme de modèles implémentés dans des ordinateurs. Le but est clairement de produire de l'artificiel, comme le rappelle Ganascia : «Si l'on veut saisir pleinement ce que désigne l'IA, il faut bien voir que ce qui importe au premier chef dans l'IA, ce n'est pas l'intelligence, mais c'est l'artificiel : l'intelligence artificielle est d'abord un artifice, une ruse au moyen de laquelle on prête une intelligence aux machines dans le but de les asservir à nos propres fins » (Ganascia 1993 : 210). De plus, il n'est absolument pas nécessaire de reproduire les mécanismes de l'intelligence naturelle pour qu'un système interactif puisse être considéré comme "intelligent" (cf. Jorion).

« (...) si l'utilisateur d'un système intelligent a l'impression que celui-ci comprend une intention, qu'il manipule des significations, qu'il exprime des idées ou qu'il est conscient, il est indifférent que ces objets jouent ou non un rôle effectif dans les mécanismes implémentés dans la machine. Autrement dit, dans la mesure où un utilisateur pourra - lui - toujours attribuer des intentions, reconnaître des idées ou des significations dans le fonctionnement d'un système naturel ou artificiel, le seul problème qui se posera au concepteur d'un système intelligent sera que la machine manifeste ces propriétés au niveau phénoménal. Lequel est le seul qui importe à l'utilisateur humain, puisque c'est à ce niveau qu'il juge de l'intelligence et qu'il mobilise le schéma de prévisibilité qui s'applique à ce qui est intelligent en général ». (Jorion 1990 : 12).

Citation 5. Jorion

Il demeure que, dès que l'on arrive au point c) du programme énoncé par G. Tiberghien (cf. **IA cognitiviste**), la relation entre l'IA et la psychologie cognitive débouche sur un problème central : celui de la causalité des représentations mentales, sur lequel Fodor notamment appuie sa théorie de la modularité de l'esprit et dont P. Engel en donne un résumé clair (cf. **Engel**) :

« (...) la psychologie cognitive postule que le comportement des organismes s'explique par des représentations mentales douées de contenus et causalement efficaces (des « inférences », des « plans », des « bases de données », des « représentations de connaissances »). Mais comment des représentations douées de contenus peuvent-elles être causalement efficaces, s'il est vrai que, selon le postulat scientifique, toute causalité est physique ? (...) La solution selon Fodor vient en deux temps. Premièrement, les attitudes propositionnelles sont des relations à des propositions, entendues au sens de symboles concrets « inscrits » dans le cerveau. (...) En d'autres termes, les états mentaux comme les croyances et les désirs, sont des états fonctionnels de l'organisme, qui se définissent par leurs relations fonctionnelles et causales à des stimuli, à des sorties comportementales, et à d'autres états mentaux (...). Deuxièmement on peut comprendre la relation causale entre des contenus sémantiques de représentation et des comportements si l'on recourt à la métaphore de l'ordinateur, car les ordinateurs ont précisément la propriété de relier les propriétés causales de symboles à leurs propriétés sémantiques, par le biais de leur structure syntaxique. » (Engel 1995 : 53-54).

Citation 6. Engel

Ainsi posée, cette problématique de la causalité des représentations mentales débouche sur le terrain de la philosophie de l'esprit. Mais avant d'y venir, il est nécessaire, pour comprendre certains débats animés aussi bien chez les spécialistes de l'IA qu'en philosophie et en SC, d'aborder le point de vue extrémiste de l'intelligence artificielle.

VI - L'IA forte : de l'homo au robot sapiens

Outre les diverses tendances et positions théoriques ou épistémologiques qui constituent l'IA, on peut observer, tout au long de son histoire, deux pôles d'attraction antagonistes nommées, en particulier par J. Searle, "IA forte" et "IA modérée" (plutôt que "faible").

L'IA modérée est une position pragmatique qui consiste à ne se référer qu'à ce qui est de l'ordre du pratique et du réalisable dans un proche ou moyen avenir, et qui vise à la réalisation

de tâches difficiles mais concrètes et de première importance sur les plans industriel, économique ou social (traduction automatique, “fouille” de textes, résolution de problèmes complexes dans divers domaines, systèmes de perception-action en temps réel, etc.).

L'IA forte de son côté offre une position extrême quant à la conception même de l'esprit, de la machine et finalement de l'homme. Actuellement, cette position se nourrit des meilleures avancées des neurosciences — celles notamment qui ouvrent des voies de recherche sur les liens entre raisonnements, décisions, émotions, et états corporels, ou celles qui explicitent le continuum entre cognition biologique et cognition sociale. Elle reprend ces thèmes pour tracer un programme d'élaboration complète d'un esprit dans une machine. L'IA forte n'hésite pas à esquisser une sorte de fantasme de réalité virtuelle absolue dont une version amusante, de type sentimentalo-humaniste, est ici citée à titre d'exemple : «The emotional states observed in higher animals and humans — curiosity, love, fear, hate, jealousy, etc. — are an integral part of, and contribute to, social intelligence. Likewise, higher intelligence must also include as part of its repertoire social intelligence. From the point of view of human requirements, we need neither computers that hate, nor computers that are unduly afraid. We want computers that love — computers that cherish and care for their human masters, and perhaps also for their machines colleagues». (Stonier 1992 : 169-170).

Mais l'IA forte n'en reste pas là. Remplacer peu à peu des parties de notre cerveau et de notre corps par des éléments micro-électroniques, l'un des grands maîtres de l'IA l'envisage avec le plus grand sérieux (cf. **Minsky**).

« Mieux nous connaissons notre cerveau, plus nous pourrons l'améliorer (...). À mesure que ces inventions s'imposeront, nous les relierons à nos cerveaux, en insérant des milliers d'électrodes microscopiques dans le corps calleux, grand fuseau nerveux où circulent les données du cerveau. Toutes les parties du cerveau seront équipées de nouveaux accessoires. Enfin, nous remplacerons toutes les parties du corps et du cerveau et nous remédierons à tous les défauts et à toutes les blessures qui écourtent notre vie ». (Minsky 1994, : 123).

Par conséquent,

« Laisserons-nous la terre à des robots ? Oui, mais ils seront nos enfants. Nous devons nos esprits à la vie et à la mort de toutes les créatures qui nous ont précédés dans la lutte nommée évolution. Tout cet effort ne doit pas se transformer en un immense gâchis : c'est notre mission. » (Minsky 1994, : 126).

Citation 7. Minsky

Ainsi, exempts du traumatisme de la naissance, grâce à l'accouchement sans douleur permis par le clonage robotique («un milliard de copies en moins d'un an » (Minsky 1994, : 124 !), nos arrières-arrières-....- arrières petits-enfants, les robots-humains parfaits et immortels, qui conserveront dans leur prodigieuse mémoire la totalité des connaissances, réflexions philosophiques, poèmes, tragédies et partitions musicales ..., seront-ils à même de se demander ce qui a bien pu conduire ce malheureux petit homme du vingtième siècle à opérer cette curieuse mutation consistant à se transformer en robot immortel et parfait ?

On retrouve, dans cette implosion radicale du questionnement ontologique, le thème de la transfiguration rencontré dans la métaphysique : l'homme, par son perfectionnement, accède à l'immortalité. Mais ici, l'ange qu'il devient est un robot. L'autre envolée de l'IA forte dans la

métaphysique se fait par le thème du Golem : l'homme reste ici un homme biologique, mais il produit un double, le robot, censé au moins l'égaliser en tous points.

VII - L'IA et la philosophie de l'esprit : l'esprit est-il computationnel ?

1. L'IA forte représente

L'IA forte représente une curiosité qui peut tantôt amuser tantôt inquiéter, mais qui, en tout cas, se développe, comme tous les intégrismes, sur un arrêt de la réflexion. Il existe en revanche une problématique des plus fécondes autour de questions fondamentales telles que la conscience, l'interprétation, l'intentionnalité, etc., englobées sous le terme très générique de sciences de l'esprit. On parle aussi de philosophie de l'esprit, de philosophie cognitive (Rialle et Fisette 1996) et de fonctionnalismes (l's est important, cf. INTELLECTICA n° 21 (2), 1995, richement consacré à ces thèmes). Dans cette veine, on assiste depuis quelques années à une floraison de livres et d'articles dont les titres juxtaposent le mot esprit au mot ordinateur ou à l'un de ses équivalents paradigmatiques, tel qu'intelligence artificielle, machine, "computation". La question récurrente qui s'y trouve débattue, sous diverses formes (elle est posée ici sous une forme simplifiée au maximum) est la suivante : l'esprit est-il computationnel ? En clair : le cerveau, si l'on admet l'indissociabilité corps-cerveau-esprit, fonctionne-t-il dans ses diverses activités comme une machine de traitement de symboles ? Si la réponse est positive, deux conséquences s'imposent alors : d'une part, on aura réussi à percer l'énigme de l'apparente dualité corps-esprit, qui est aujourd'hui l'un des points les plus discutés en philosophie de l'esprit, avec toutes les tentatives de naturalisation de l'esprit; d'autre part, on saura construire une machine intelligente. Si a contrario la sphère cognitive n'était pas computationnelle, serait-il encore possible d'en rendre compte scientifiquement ? Mais la question recouvre plusieurs pièges, à commencer par celui du contenu du terme "scientifique".

2. Le jeu de la réflexion

Le jeu de la réflexion dans lequel nous entraîne la question de l'esprit computationnel nous conduit à une série d'interrogations "délicieuses", pour reprendre un terme d'Hofstadter, qui intéressent l'IA "philosophique", mais dont la place manque ici pour les aborder. Disons simplement que, dans l'hypothèse soutenue actuellement (Jackendoff, Dennett, etc) selon laquelle la conscience pourra être, tôt ou tard, expliquée — c'est-à-dire explicitée en des termes rationnels au moyen d'un langage de description approprié —, ne sera-t-elle pas du même coup programmable ? Si notre conscience peut être révélée de manière entièrement rationnelle, c'est-à-dire sans recours à aucun moment de son explication à une forme de pensée magique ni à une boîte noire impénétrable, cette explication ne sera-t-elle pas du même coup réductible à un algorithme ? C'est une question à considérer comme malicieuse, comme toutes les questions faussement naïves, mais y pénétrer demanderait plus de développement.

VIII - Conclusion

De ce parcours "en diagonale" des relations entre l'IA et les SC, retenons que ces deux vastes domaines scientifiques sont historiquement et pratiquement indissociables. De plus, l'ordinateur constituant aujourd'hui l'aune ou la métaphore la plus employée pour comparer, opposer, jauger l'intelligence, l'IA se trouve d'emblée convoquée dans tous les débats philosophiques sur les capacités cognitives de l'homme. De fulgurants raccourcis sont pris au

passage par quelques prophètes de l'IA forte, raccourcis qui ont néanmoins pour effet positif de stimuler, chez ceux qui ne souhaitent pas les emprunter, la réflexion critique et épistémologique. Mais à force de "grignoter" du terrain sur la part spécifiquement humaine de l'intelligence, on ne voit pas pourquoi nous n'arriverions pas à produire une machine qui illusionne jusqu'à un degré élevé, une machine qui apprenne comme un enfant, hésite, s'émerveille, explique ses hésitations, demande de l'aide, se met en colère lorsque les choses ne vont pas, etc. La question cependant est moins celle de savoir si cela est réalisable ou non, que celle de comprendre ce qui, dans nos fondements ou notre culture, pousse une frange non négligeable de nos contemporains à ce que ce soit une machine qui nous parle d'amour! Chemin faisant, l'IA rencontre, redécouvre et renouvelle — plutôt qu'elle ne les "découvre" — quelques problèmes fondamentaux : ceux notamment qui concernent la définition de l'intelligence et ses déterminations, la définition de l'esprit, de la conscience et de l'intentionnalité, la relation cerveau-esprit, la formation des concepts et des croyances, l'apprentissage de la langue et sa relation à l'intelligence, son usage en tant que médiateur dans les dialogues et dans des textes, etc. "Tout cela ne dépasse-t-il pas le cadre de l'IA ?", objectera-t-on. La réponse est bien sûr : oui. Ces questions dépassent en effet largement le champ de l'IA, mais aujourd'hui, tout le monde se retrouve "sur le pont" — neuropsychologues (Damasio, Crick, Edelman...), mathématiciens (Penrose...), philosophes (Dennett, Putnam, Searle...), etc. — pour percer les mystères des fonctions supérieures de l'homme. Tout un chacun sans exception se réfère à un moment ou à un autre à l'IA ou à l'ordinateur. Ce qui domine néanmoins, c'est que personne ne semble être vraiment mieux placé que les autres pour aborder ces questions et qu'une certaine pulsion épistémophilique pousse des gens de l'IA, comme dans d'autres disciplines, à contribuer à cette recherche pour laquelle la multidisciplinarité constitue une condition de progrès.

IX - Bibliographie

- Dupuis, J. P. (1985). L'essor de la première cybernétique. Cahiers du CREA 7.
- Engel, P. (1995). Introduction à la philosophie de l'esprit. Paris, La Découverte.
- Ganascia, J. G. (1993). L'Intelligence Artificielle. Paris, Flammarion.
- Harnad, S. (1994). L'ancrage des symboles dans le monde analogique à l'aide de réseaux neuronaux : un modèle hybride. Modèles de la cognition : vers une science de l'esprit. V. Rialle et D. Payette, Univ. du Québec à Montréal, Dép. de Philo. LEKTON vol. IV (2): 65-83.
- Jorion, P. (1990). Principes des systèmes intelligents. Paris, Masson.
- Jorrand, P. (1995). On nous dit que l'IA est morte... Bulletin de l'AFIA (Association Française pour l'Intelligence Artificielle) 21(avril): 3-4.
- Lacharité, N. (1994). Le problème de l'architecture dans le débat entre l'approche connexionniste et les approches «classiques» de la représentation. Modèles de la cognition : vers une science de l'esprit. V. Rialle et D. Payette. Univ. du Québec à Montréal, Dépt. de Philosophie. LEKTON, 4(2): 37-64.
- Le Ny, J. F., Ed. (1993). Intelligence Naturelle et Intelligence Artificielle. Paris, Presses Univ. de France.
- Minsky, M. (1994). Laisserons-nous la terre à des robots ? Pour la Science 206: 120-126.
- Pelissier, A. et A. Tête, Eds. (1995). Sciences cognitives : textes fondateurs. Paris, Presses Univ. de France.
- Posner, M., Ed. (1989). Foundation of Cognitive Science. Cambridge, MA, MIT Press.

Rialle, V. (1996). L'intelligence artificielle et sa place dans les sciences de la cognition. Bulletin de l'Association Française de l'Intelligence Artificielle(26): 8-12.

Rialle, V. et D. Fissette, Eds. (1996). Penser l'esprit : des sciences de la cognition à une philosophie cognitive, Presses Univ. de Grenoble.

Stonier, T. (1992). Beyond Information : the natural history of intelligence. Berlin, Springer-Verlag.

Tiberghien, G. (1993). Les sciences cognitives : un nouveau programme scientifique ? Dictionnaire critique de la communication. L. Sfez. Paris, Presses Univ. de France: 817-824.