

Modèles de la pensée : Quels enjeux pour les chercheurs en sciences cognitives ?

Brigitte CHAMAK*

RÉSUMÉ. Cet article a pour objet l'analyse des modèles de la pensée proposés par les chercheurs en sciences cognitives et des enjeux sous-jacents à la construction de ce domaine. La première partie de ce travail porte sur l'analyse des textes, des discours, des références, des représentations, des images que se sont forgés les acteurs les plus actifs dans la construction des sciences cognitives. Si pour certains, la logique apparaît comme le meilleur moyen d'aborder le fonctionnement de la pensée, pour d'autres, c'est l'étude du cerveau qui leur semble l'approche la plus adéquate. L'objectif de cet article est de comprendre dans quel univers culturel les chercheurs en sciences cognitives évoluent, quelles positions philosophiques ils adoptent, quelles références ils utilisent et comment leurs pratiques et leur formation influent sur leurs conceptions. La deuxième partie aborde la question des enjeux disciplinaires, économiques et politiques attachés au processus d'institutionnalisation des sciences cognitives en France.

Mots clefs : sciences cognitives, modèles de la pensée, institutionnalisation.

ABSTRACT. Theories of the Mind and the Stakes for the Researchers in Cognitive Science. This paper is concerned with the analysis of the theories of the mind proposed by the researchers in cognitive science and the stakes linked to the construction of this domain. The first part of this work deals with the analysis of the texts, discourses, references and representations of the most famous actors of the construction of the cognitive science. Some maintain that the logic is the most suitable mean to study the cognitive processes, whereas others believe that it is the study of the brain functioning. The aim of the paper is to understand the cultural universe of the researchers in cognitive science, what philosophical choices they adopt and how their practices and formation affect their conceptions. The second part of the paper deals with the institutional aspects of the construction of the cognitive science in France.

Key words: cognitive science, theories of the mind, institutionalization.

INTRODUCTION

Le réalisme métaphysique et scientifique postule que l'Univers a des propriétés « en soi », indépendantes de l'observateur. Cette position présente l'avantage d'une simplification qui a favorisé la construction des systèmes de connaissance en Occident mais pour répondre aux problèmes qui résistent à une approche traditionnelle, le constructivisme et la systémique proposent une alternative qui consiste à étudier conjointement système observant et système observé (Tabary, 1991). Au sein du domaine des sciences cognitives, réalisme

* CESAMES, Centre de Recherche Psychotropes, Santé Mentale, Société, CNRS UMR 8136, INSERM U611, Université René Descartes - Paris 5, 45, rue des Saints-Pères 75270 Paris Cedex 06. E-mail : brigitte.chamak@paris5.sorbonne.fr.

et constructivisme coexistent, mais l'utilisation de ces termes recouvre bien des positions et des motivations. Afin de mieux comprendre comment des choix philosophiques et des idées se développent dans un contexte précis, l'étude de l'émergence des sciences cognitives et de l'enracinement historique et sociologique du cognitivisme, du connexionnisme et du constructivisme va servir ici de fil conducteur. Cet article, qui vise à mettre en évidence l'influence du contexte socio-historique sur les prises de position philosophiques, sert de prétexte à un exercice périlleux qui consiste à traiter de l'émergence d'un domaine scientifique sous un angle à la fois historique, sociologique et épistémologique.

Depuis plus de trente ans, les sciences cognitives connaissent un développement impressionnant. Le nombre de chercheurs, de livres, d'enseignants, d'étudiants, de publications dans ce domaine augmente tous les jours, mais il est très difficile de véritablement comprendre ce que recouvre cette étiquette « sciences cognitives » qui est utilisée dans des acceptations différentes. Les chercheurs en sciences cognitives expliquent qu'ils aspirent à regrouper diverses disciplines pour analyser les processus impliqués dans la formation et l'exploitation de la connaissance. Ils sont intéressés par l'étude du fonctionnement de la pensée et cherchent à décrire, expliquer, simuler les fonctions cognitives telles que le langage, le raisonnement, la perception, la compréhension, la mémoire, ou l'apprentissage. Ce type d'intérêt les conduit à proposer différentes théories de l'esprit qui ont des conséquences importantes puisque, en un sens, elles tentent de définir ce qui caractérise un être pensant.

L'interdisciplinarité prônée par les chercheurs en sciences cognitives est souvent mise à rude épreuve car, en fonction de leur discipline d'origine, ils produisent des discours différents qui correspondent à des conceptions concurrentes. Pour les uns, l'esprit étant un phénomène fonctionnel, ses fonctions relèvent de processus cognitifs dont il s'agit de caractériser les opérations en termes de modèles. Les modèles fonctionnels de l'esprit sont généralement associés à la métaphore de l'ordinateur et les conceptualisations des phénomènes cognitifs s'organisent autour de l'intelligence artificielle et de ses simulations (Andreewsky, 1991). L'hypothèse de base est que la pensée opère par un travail de traitement de l'information et les fonctions cognitives sont assimilées à des opérations logiques. L'objectif consiste à obtenir de la machine un comportement intelligent, ou plutôt qui serait jugé tel si c'était un être humain qui le produisait. Pour d'autres, l'esprit procédant du fonctionnement cérébral, son étude relève principalement des neurosciences. Il s'agit alors de mettre en évidence les mécanismes neurophysiologiques sous-jacents aux fonctions cognitives: détecter quelles molécules, quelles structures neurales sont impliquées dans telle ou telle fonction. Aux recherches sur le système nerveux central de caractère essentiellement biologique ou physiologique sont associées des simulations utilisant des « réseaux de neurones artificiels ». Alors que dans le premier cas, l'étude du fonctionnement du cerveau n'est pas vraiment pertinente, dans le second, elle est centrale. Les sciences cognitives constituent donc une nébuleuse complexe avec des niveaux d'analyse différents et des modèles concurrents. Les dangers qui menacent ce domaine encore jeune ont trait au réductionnisme rapportant tous les mécanismes de la pensée à un niveau d'explication logique ou neurologique et à l'hégémonie de disciplines comme l'intelligence artificielle ou les neurosciences.

Dans ce travail, je me suis tout particulièrement intéressée à l'analyse des textes, des discours, des références, des représentations, des images que se

forment les acteurs les plus actifs dans la construction des sciences cognitives. L'objectif est de comprendre dans quel univers culturel ils évoluent, quels types de position philosophique ils adoptent, quels types d'idéologie ils développent, comment leurs pratiques influent sur leurs conceptions et quels arguments ils mettent en avant pour tenter d'acquérir une certaine autonomie. Les relations entre recherches dites « fondamentales » et recherches « appliquées », entre pratiques et théories, les phénomènes de stratégies disciplinaires, de concurrence et de hiérarchie ; les rapports entre science et technologie, science et politique sont autant de questions soulevées tout au long de cet article qui met en lumière l'influence des phénomènes contingents dans l'orientation des sciences cognitives.

LOGIQUE, INFORMATION ET MODÉLISATION

En analysant les ouvrages qui, depuis les années quatre-vingt, s'attachent à définir et promouvoir les sciences cognitives, on pénètre rapidement dans un univers où l'information et la logique importent davantage que l'énergie ou la matière et où l'être vivant est transformé en une abstraction. La cybernétique et l'intelligence artificielle jouent un rôle crucial dans la mise en place de cette nouvelle vision.

1- Rôle de la cybernétique

La cybernétique qui apparaît à la fin des années quarante dans le contexte de la deuxième guerre mondiale est le fruit de rencontres entre chercheurs de disciplines différentes, en particulier de mathématiciens tels Norbert Wiener et John von Neumann, de neurophysiologistes et d'ingénieurs (Heims, 1991). Des problèmes d'ordre militaire sont à l'origine de la cybernétique: Norbert Wiener et un ingénieur, Julian Bigelow, étudiaient les aspects mathématiques de guidage et de contrôle des tirs anti-aériens (cf. article de Jérôme Segal). Ils devaient concevoir un système de prédiction capable non seulement de localiser un avion pour l'intercepter, mais également d'anticiper la stratégie du pilote. Ils firent appel à des neurophysiologistes et des psychologues pour mieux comprendre le mode de fonctionnement du pilote.

C'est au cours de réunions à Boston, juste avant la Seconde Guerre mondiale, que Wiener rencontre Arturo Rosenblueth, neurophysiologiste qui cherche à promouvoir des discussions interdisciplinaires. Wiener évoque les dérèglements des systèmes de rétroaction des machines et demande à Rosenblueth si de tels dysfonctionnements peuvent être observés dans le système nerveux. C'est le début d'une collaboration qui donne lieu en 1943 à la publication d'un article (Rosenblueth et al., 1943), souvent considéré comme l'un des textes fondateurs de la cybernétique, dans la mesure où il est interprété comme ayant supprimé la barrière ontologique qui sépare l'homme de l'animal et de la machine. Rosenblueth, Wiener et Bigelow proposent de comparer certains troubles du système nerveux aux dérégulations mécaniques (cf. article de Jérôme Segal). En 1942, lors d'un séminaire à New York consacré à l'étude de l'inhibition centrale du système nerveux, Rosenblueth expose les idées développées dans cet article. Sa présentation insiste sur quelques aspects des organismes vivants susceptibles d'être analysés à l'aune des réalisations de certaines machines analogiques. Le modèle proposé remplace la traditionnelle relation cause-effet d'un stimulus par une causalité circulaire nécessitant un

feedback négatif¹. Le neurophysiologiste Warren McCulloch qui était présent propose l'idée d'organiser après guerre un cycle de conférences qui deviendront les célèbres conférences Macy. Après cette rencontre, les échanges s'intensifient entre le groupe de McCulloch à Chicago et celui de Wiener au MIT. Wiener de son côté discutait régulièrement avec John Von Neumann². S'inspirant du modèle électrique du cerveau, ils échangeaient des idées sur la question des analogies entre organismes et machines. Bien que deux articles publiés en 1943 soient décrits comme fondateurs de la cybernétique (Rosenblueth et al., 1943; McCulloch & Pitts, 1943), ce n'est véritablement qu'au cours des conférences Macy que la cybernétique va prendre corps. L'importance accordée à la modélisation est centrale. À défaut de comprendre comment fonctionne véritablement le cerveau, des scientifiques vont concevoir des modèles formels, pures constructions intellectuelles, mais aussi des modèles matériels qui puissent être comparés à l'objet que l'on cherche à modéliser, c'est-à-dire le cerveau.

Les conférences Macy (*Conferences for Circular Causal and Feedback Mechanisms in Biological and Social Systems*) au nombre de dix, débutent en mars 1946, et c'est au cours de la troisième rencontre que Norbert Wiener propose le terme « cybernétique » (Heims, 1991). Des chercheurs de disciplines différentes (mathématiciens, logiciens, ingénieurs, neurophysiologistes, anthropologues...) se retrouvent ainsi régulièrement jusqu'en 1953. Des discussions très stimulantes caractérisent ces rencontres mais la notion même de modèle s'estompe progressivement. Une relation d'équivalence s'instaure entre les faits et leur description. La vision des ingénieurs semble rapidement s'imposer: les troubles du système nerveux sont comparés aux dérégulations mécaniques ; les analogies et métaphores homme/machine se multiplient ; la logique et la modélisation apparaissent comme primordiales ; tout est traduit en termes de communication et de contrôle. Les cybernéticiens cherchent à produire une théorie qui soit valable à la fois pour le vivant et l'inanimé, l'animal et la machine et pensent avoir trouvé dans le principe de rétroaction (ou *feedback*) le mécanisme commun aux différents niveaux d'organisation.

Avec la cybernétique, c'est une vision mécaniste de la pensée qui est proposée : le cerveau est comparé à une machine à calculer et l'être vivant à une machine auto-régulée. De l'importance accordée à l'information et à ses échanges devait émerger une nouvelle vision du monde, vision dominée par la culture des physiciens, toujours à la recherche de lois universelles, celle des mathématiciens-logiciens, qui raisonnent en termes de formalisation et modélisation, et celle des ingénieurs, qui recherchent l'efficacité. La cybernétique joue un rôle crucial dans l'introduction du formalisme logico-mathématique dans les sciences du cerveau et la conception de l'organisation des machines à traiter l'information. Favorisant la production de grands calculateurs, elle a concrétisé le vieux rêve de la machine pensante. Le modèle, subordonné au réel qu'il ne faisait que mimer, s'émancipe et devient l'égal du référent (Dupuy, 1994). Le modèle ne renvoie plus qu'à lui-même. Ces nouvelles machines vont focaliser

1 Le principe de feedback veut que l'information transmise aux mécanismes de contrôle d'un système fasse la différence entre une situation désirée et la situation existante, le mécanisme agit pour réduire la différence entre les deux situations.

2 Depuis 1937, von Neumann assumait le rôle de conseiller en matière de recherche balistique auprès de l'armée américaine. À partir de 1941, il se consacre aux mathématiques appliquées aux explosions et à l'aérodynamique, puis il participe à l'élaboration de l'ENIAC (Electronic Numerical Integrator and Calculator), un calculateur électronique.

l'attention des chercheurs et permettre l'éclosion d'un nouveau domaine : l'intelligence artificielle.

2- Rôle de l'intelligence artificielle

Si la cybernétique abordait la compréhension des processus mentaux à travers la simulation d'une logique présumée des réseaux de neurones, l'intelligence artificielle (IA), qui fait son apparition en 1956 aux États-Unis, s'en démarque en se donnant pour objectif d'étudier l'esprit indépendamment de son support matériel, le cerveau.

Ce domaine est caractérisé par l'élaboration de modèles computationnels, de programmes d'ordinateur en vue de produire des machines au comportement "intelligent". La facilité à obtenir des financements pour cette discipline naissante s'explique par le fait que les recherches sur les interactions homme/machine, la vision artificielle et la traduction automatique intéressaient tout particulièrement les militaires en cette période de guerre froide.

Cependant des problèmes apparaissent assez rapidement car le présupposé que toute connaissance peut être appréhendée en termes purement logiques est rudement mis à l'épreuve. En 1966, un rapport rédigé à la demande du Département d'État américain critique les recherches sur la traduction automatique et conduit à une diminution importante des financements (Mc Corduck, 1979). En 1973, un rapport sur l'état des recherches en IA, produit par Sir James Lighthill à la demande du *Science Research Council*, centre de recherche britannique, est largement négatif (Lighthill et al., 1973). A la même époque, Hubert Dreyfus, professeur de philosophie à Berkeley, et Joseph Weizenbaum attaquent les approches et conceptions adoptées par les chercheurs en IA.

Depuis les années soixante, les débats s'intensifiaient autour des définitions de l'intelligence et des questions liées à la formalisation et au machinisme. L'intelligence étant définie par les chercheurs en IA comme la capacité à résoudre des problèmes, les machines pouvaient être qualifiées d'intelligentes. Hubert Dreyfus, qui a observé pendant quelques mois les pratiques des chercheurs en IA dans le but d'écrire une histoire de cette nouvelle discipline, a rédigé en 1965 un article intitulé *'Alchemy and Artificial Intelligence'* qui attaquait ce domaine, le comparant à une pseudo-science du Moyen-Age (Mc Corduck, 1979). En 1972, il publie *'What Computers can't do. The limits of AI'*, où il critique cette approche qui consiste, selon lui « à réduire la nature humaine à une mécanique ». Expliquant que préjugés philosophiques et naïveté caractérisent les hypothèses premières en IA, il insiste sur le fait qu'en dépit de nombreuses difficultés, les chercheurs en IA affichent un optimisme démesuré parce qu'ils sont convaincus que l'esprit fonctionne comme un système de manipulation de symboles et que toute connaissance peut être formulée explicitement. Il voit dans ces conceptions un danger et évoque le fait qu'entre nature et culture, la frontière n'est pas si marquée (Dreyfus, 1984: 361) :

« même les impératifs de l'instinct peuvent être détournés et dominés en vertu de paradigmes ; l'homme n'a donc pas une «nature» fixée une fois pour toutes. La nature humaine est de fait si malléable qu'il se pourrait bien qu'elle soit sur le point de changer une fois de plus. Si le paradigme informatique devient si puissant que les êtres humains commencent à se percevoir comme des calculateurs numériques, alors, puisque les machines, nous l'avons vu, ne pourront pas se comporter comme des êtres humains, peut-être verra-t-on

les humains se mettre petit à petit à se comporter comme des machines. »

Les analyses et commentaires de Dreyfus annonçaient le début d'une période où les attaques contre l'IA allaient déferler. En 1976, Joseph Weizenbaum, qui avait travaillé en IA au MIT, se retournait contre ce domaine et s'élevait contre « l'impérialisme de la raison instrumentale qui vient imposer à l'homme l'univers de la machine » (Weizenbaum, 1976). Il déplorait « l'atrophie de l'esprit humain qui fait confiance à la seule science pour interpréter la réalité » et estimait que les sciences de l'ordinateur, « au lieu de révolutionner le monde, ont été utilisées pour soutenir les institutions sociales et politiques américaines ». Weizenbaum s'insurgeait contre ce qui lui apparaissait comme un déplacement d'intérêt de l'homme vers la machine.

Au MIT, il avait conçu un système, ELIZA, qui cherchait à imiter le dialogue entre un individu et un psychologue. Utilisant la méthode rogérienne, qui consiste, pour faire parler un individu, à reprendre, en phrase interrogative, ce qu'il vient de dire, il a réussi à impressionner son entourage par les résultats obtenus. Weizenbaum travaillait en collaboration avec un psychiatre qui a voulu faire croire que ce programme pouvait prendre la place d'un thérapeute. Cette dérive n'a pas été appréciée par Weizenbaum qui craignait que la fabrication de machines visant à imiter, et donc à remplacer l'homme dans diverses tâches, finisse par reléguer l'être humain au second plan. Les intérêts économiques en jeu, les avantages à utiliser des machines plutôt que des hommes, lui faisaient redouter une orientation trop radicale dans un sens qu'il estimait néfaste.

Aux craintes de Dreyfus et de Weizenbaum s'ajoutent les questions du philosophe John Searle (Searle, 1980). Posant les problèmes de la conscience, de l'intentionnalité et du sens, il a engagé des débats qui se poursuivent encore aujourd'hui. Décrivant l'ordinateur comme une machine qui effectue des opérations formelles mais qui est dépourvue de connaissance sémantique et d'intention, il explique que la dimension proprement sémantique manque aux langages de programmation et que la machine ne connaît pas les états des choses auxquels sont censées renvoyer les formules. Faisant la distinction entre IA « faible » et « forte », il critique la seconde qui affirme que l'ordinateur est capable de penser, de comprendre, d'apprendre. Pour démontrer que l'ordinateur ne peut penser, tout simplement parce qu'il ne peut comprendre, Searle imagine une situation où des individus seraient contraints d'utiliser correctement, en suivant à la lettre des consignes précises, les divers signes de la langue chinoise sans pour autant la comprendre (*Chinese room*). Ils seraient alors dans la situation d'un ordinateur qui fonctionne selon la syntaxe mais non selon la sémantique : logiquement, mais sans comprendre, et donc sans penser véritablement comme nous le faisons, avec des états mentaux qui relèvent, chez l'humain, d'une infrastructure biologique, en même temps que biographique, psychologique, historique et sociologique.

Cette période de critique correspond aussi à l'apparition et au développement d'un nouveau domaine : celui des sciences de la cognition, qui paraît donner l'opportunité d'estomper les problèmes liés aux postulats exposés par les chercheurs en IA. Faisant partie d'une entreprise plus vaste, ce domaine allait être beaucoup plus difficile à attaquer puisque des approches et des conceptions très variées, voire contradictoires, pouvaient y être développées. L'émergence des sciences cognitives semble avoir conduit à une acceptation de

la validité de l'IA, en lui conférant une assise épistémologique via la philosophie analytique, et en atténuant, du moins en apparence, les positions extrêmes qui lui ont valu tant de critiques.

SCIENCES COGNITIVES ET MODÈLES DE PENSÉE

1- Cognitivism et fonctionnalisme

Le premier centre en sciences cognitives a été créé à l'université de Harvard en 1960 par deux psychologues, Jérôme Bruner et George Miller, mais c'est au milieu des années quatre-vingt que se sont multipliés les livres consacrés aux sciences cognitives. L'histoire de ce qui est présenté comme un nouveau domaine révolutionnaire commence à être abordée, non pas par des historiens mais par ceux-là mêmes qui sont impliqués dans sa construction. Ainsi le psychologue Howard Gardner publie, en 1985, une *Histoire de la Révolution cognitive* (Gardner, 1985). Pour lui, le noyau central de la science cognitive³ est constitué par la psychologie et l'informatique, et le lien entre ces deux domaines, via les techniques de l'information, passe par le cognitivism. L'enjeu sous-jacent est le développement de la psychologie au sein des sciences cognitives pour assurer à cette discipline, présentée comme en voie de déclin, une ouverture grâce à une collaboration avec des chercheurs en informatique.

Le postulat central du cognitivism réside dans l'hypothèse selon laquelle la cognition est essentiellement la manipulation de représentations symboliques. A l'origine de ce modèle, caractéristique de l'IA classique, se trouve l'idée que toutes nos réflexions pourraient être décrites sous forme d'une suite d'opérations logiques et que « penser, c'est calculer ». Cette expression avait été utilisée de manière métaphorique au XVII^e siècle par Wilhelm Gottfried Leibniz, précurseur de la logique symbolique. Avec le travail de Von Neumann et la naissance de l'ordinateur, cette expression prend un sens neuf : pour un cognitiviste, la pensée humaine procède de la même façon qu'un programme informatique. Elle combine des opérations logiques effectuées sur des symboles abstraits.

Le mathématicien Marvin Minsky, l'un des fondateurs de l'IA, résumait ainsi le problème de la représentation des connaissances (Minsky 1975):

« de quelles méthodes de représentation et de traitement des connaissances faut-il doter un programme pour qu'il puisse affronter, avec une souplesse comparable à celle de l'esprit humain, des situations incertaines ? »

Il a décrit un système permettant d'incorporer des connaissances dans des programmes sous forme de stéréotypes qui peuvent s'adapter à des circonstances variables, plutôt que sous forme d'ensemble de données fixes et définies de manière univoque. Le programme est conçu pour sélectionner le schéma qu'il suppose le mieux adapté et tenter de l'ajuster le mieux possible aux données qu'il découvre peu à peu.

Dans un ouvrage publié en 1985, Marvin Minsky décrit l'esprit comme une société constituée d'agents (Minsky 1985). L'esprit est assimilé à la société des hommes, composée de réseaux où les agents sont interconnectés et le monde de la concurrence et de l'entreprise semble lui servir de modèle pour décrire le

³ Il utilise le terme « science cognitive » au singulier pour insister sur l'idée d'une science unifiée.

fonctionnement de l'esprit. Ainsi, lorsqu'il se pose la question d'un "conflit à l'intérieur des esprits", il explique que :

« plus un conflit interne entre les subordonnés d'un agent se prolonge, plus la position de cet agent s'affaiblit par rapport à ses concurrents. Si ces problèmes ne sont pas résolus rapidement, d'autres agents prendront le contrôle et ceux qui étaient précédemment activés seront «congediés» ».

Adoptant la vision de l'ingénieur, Marvin Minsky définit l'intelligence comme l'ensemble des processus qui nous permet de résoudre des problèmes, et, puisqu'« il est possible de programmer un ordinateur pour qu'il résolve n'importe quel problème par tâtonnements et sans que nous sachions à l'avance comment il le résoudra », il conclut qu'une machine peut être considérée comme intelligente.

Cognitiviste comme Marvin Minsky, le psychologue Jerry Fodor décrit un état mental comme étant caractérisé par une relation qu'entretient l'esprit avec une certaine proposition exprimée dans un langage formel, le «mentalais» (Fodor, 1975). Sa théorie de la modularité postule que l'esprit humain est constitué par des dispositifs cognitifs hautement spécialisés, répondant chacun à des principes propres et destinés à traiter de façon automatique un type très limité d'informations en provenance de l'extérieur (Fodor, 1983). Il distingue les processus périphériques, modulaires et cloisonnés, et les processus centraux, non modulaires, qui exploitent les informations fournies par les systèmes périphériques. Exhumant la psychologie des facultés de Franz Josef Gall et s'inspirant des travaux du linguiste Noam Chomsky⁴, il adhère à l'innéisme et assimile les facultés cognitives à des modules de traitement de l'information. Fodor considère la pensée comme un automate traitant des «entrées» (*input*) et produisant des «sorties» (*output*); l'état interne du système correspondrait à nos états mentaux et nos représentations mentales. Cette théorie computationnelle de l'esprit et du comportement considère les individus comme des systèmes traitant de l'information.

Fodor, comme beaucoup de cognitivistes, adhère au fonctionnalisme, courant qui s'intéresse davantage aux fonctions qu'au support matériel de ces fonctions (autonomie du formel relativement au matériel). Zenon Pylyshyn, qui dirige un centre de science cognitive aux États-Unis, considère que le fonctionnalisme est la théorie de base unifiant le domaine de la science cognitive (Pylyshyn, 1984). Il présente la cognition comme une forme de calcul et soutient que le contenu sémantique des états mentaux est codé de la même façon que celui des représentations d'ordinateur. En 1984, il assignait à la science cognitive le but de devenir un domaine scientifique original dont le sujet porterait sur les «choses connaissantes», choses qui partagent une base commune : la manipulation de représentations symboliques. Il remettait en question l'idée suggérée par l'utilisation de l'expression «simulation sur ordinateur», c'est-à-dire une imitation plutôt qu'une proposition sérieuse décrivant comment sont réellement les choses. Penser à un modèle comme à une métaphore a, selon lui, des conséquences fâcheuses. Il n'est pas utile de se réfugier derrière une méta-

⁴ Noam Chomsky est la figure de proue de la grammaire générative. Il a tenté de reconstituer une grammaire universelle du langage humain, de formaliser la description des structures linguistiques afin d'en donner un modèle mathématique (projet de grammaire formelle). Pour lui, la capacité de parler serait innée, inscrite dans notre cerveau et prédéterminerait biologiquement la forme générique des seules grammaires possibles pour les langues humaines.

phore : pour qu'un programme d'ordinateur soit considéré comme un modèle de la cognition, le programme doit correspondre au processus effectivement réalisé par un être humain. La simulation est ainsi assimilée à l'objet ou à la fonction à imiter.

Comme l'avaient fait les cybernéticiens avant eux, les cognitivistes conçoivent la machine universelle de Turing comme un modèle pour le cerveau ou la pensée. En 1936, Alan Turing décrivait une machine imaginaire, un dispositif théorique capable de résoudre tous les problèmes traités sous forme d'algorithmes (Turing, 1936). Turing, qui, avec Von Neumann, participa activement à la conception des premiers calculateurs électroniques, était très intéressé par la production de modèle du fonctionnement du cerveau. Pour lui, la pensée était une sorte de calcul logique, non spécifique du cerveau humain : si une machine était correctement programmée, son comportement pourrait être semblable à celui d'une personne pensante et la machine pourrait, par conséquent, être considérée comme une machine pensante. Cette conception a été adoptée par les chercheurs en intelligence artificielle et par beaucoup de chercheurs en sciences cognitives.

2- Connexionnisme

Depuis le début des années quatre-vingt, un autre courant, le connexionnisme, se développe. Il prend pour modèle le réseau de neurones formels (ou réseau neuromimétique), inspiré du modèle proposé en 1943 par deux cybernéticiens : Warren McCulloch (neurophysiologiste) et Walter Pitts (logicien) (McCulloch & Pitts 1943). Dans cette modélisation en réseau, les caractéristiques physiologiques et chimiques du neurone sont laissées de côté pour faire place aux neurones idéalisés considérés comme des appareils à seuil pouvant être soit actifs, soit inactifs, laissant ou non passer le courant en fonction des impulsions excitatrices et inhibitrices qu'ils reçoivent. Cet article annonce un tournant en introduisant l'idée que la logique peut être considérée comme la discipline appropriée à la compréhension du cerveau et de l'activité mentale. Les approches connexionnistes contemporaines utilisent une stratégie comparable qui consiste à construire un système cognitif en partant de composants simples susceptibles de se relier entre eux par des connexions. L'idée de base est que pour résoudre des problèmes, la pensée humaine ne procède pas par une suite de déductions logiques ; c'est de l'interaction entre micro-unités d'information qu'émerge une solution. Alors que dans un modèle cognitiviste (système d'intelligence artificielle classique) le « calcul » est conduit séquentiellement sous la direction d'un centre de contrôle, dans un réseau connexionniste, les « calculs » sont menés en parallèle, sans contrôle central, par le simple effet d'interactions locales. Alors que le cognitiviste considère la cognition comme un « calcul » sur des représentations internes ou mentales, le connexionniste envisage la cognition comme l'émergence d'états globaux dans un réseau de composants simples avec des règles locales qui gèrent les opérations individuelles et des règles de changement qui gèrent les liens entre les éléments. D'un point de vue pratique, cette méthode mathématique de résolution de problèmes vise à trouver une solution optimale à un problème donné.

La recherche sur les réseaux de neurones prit un tournant lorsqu'en 1958 Frank Rosenblatt réalisa, pour le compte de la Marine américaine, le perceptron, modèle de perception visuelle (Rosenblatt, 1962). Ce réseau, capable d'« apprentissage », était arrangé en couches de neurones superposées. Les neurones d'entrée simulaient l'activité de la rétine, les neurones de sortie

classaient les traits que le système reconnaissait et les neurones de la couche cachée réalisaient les calculs intermédiaires. Trouver un réseau qui donne la bonne réponse à tous les exemples était qualifié d'apprentissage et le procédé pour y arriver était nommé règle ou algorithme d'apprentissage. Il existe des algorithmes qui modifient l'architecture du réseau en le rendant de plus en plus complexe, jusqu'à ce que tout l'ensemble d'apprentissage soit appris. D'autres approches, au contraire, partent d'un réseau complexe, et le simplifient en supprimant les connexions inutiles.

L'architecture trop simple du perceptron ne permettait de réaliser qu'un ensemble limité de toutes les fonctions possibles de ses entrées : les fonctions linéaires séparables. Cette limitation a été mise en exergue par Marvin Minsky et Seymour Papert, deux pionniers de l'IA, pour critiquer cette approche qui entraînait en compétition avec le cognitivisme (Minsky & Papert, 1969). Les recherches sur les réseaux de neurones ont ainsi été abandonnées et il a fallu attendre dix ans pour que la méthode des réseaux neuronaux, améliorée par l'introduction de plusieurs couches cachées, réussisse à s'imposer.

De même que le cognitivisme, le connexionnisme représente un ensemble de méthodes de modélisation et de simulation. L'ordinateur est toujours considéré comme l'outil de base mais il n'est plus qu'un simple outil de simulation, et non une idéalisation fondatrice pour l'étude de l'esprit.

Pourtant, l'esprit est encore comparé à un programme fonctionnant comme un système de manipulation de symboles, et la conscience à un système d'exploitation de l'esprit. Dans les années 1990, Daniel Dennett, philosophe américain, assimilait la conscience à l'activité d'une machine virtuelle implantée dans le cerveau (Dennett 1994). Il était d'accord avec Philip Johnson-Laird, auteur en 1988 de 'L'ordinateur et l'esprit', pour dire que *"toute théorie de l'esprit doit le traiter comme un automate"* et expliquait les phénomènes de la conscience humaine en termes d'opérations réalisées par une machine virtuelle, une sorte de programme d'ordinateur qui façonne les activités du cerveau :

« en principe, un robot bien «programmé» avec un ordinateur composé d'éléments de silicone serait conscient, aurait un moi. Ou plutôt il aurait un moi conscient dont le corps serait un robot et dont le cerveau serait l'ordinateur. »

Il poursuivait :

« A présent que le flux de la conscience a été repensé sous forme des opérations d'une machine virtuelle réalisée dans le cerveau, il cesse d'être 'évident' que nous sommes victimes d'une illusion quand nous imaginons qu'un tel flux pourrait se loger dans le cerveau électronique d'un robot. »

Il précisait que, pour lui, il ne faisait pas de doute que :

« la vie puisse s'expliquer en termes de choses qui ne sont pas elles-mêmes vivantes. »

Constatant qu'en fait il avait remplacé une famille de métaphores et d'images par une autre, il affirmait que les métaphores sont les outils de la pensée et que personne ne peut réfléchir sur la conscience sans faire appel à elles.

A la lecture de ce type d'ouvrages, on constate qu'un nouvel univers conceptuel est élaboré, un univers qui fabrique de nouvelles métaphores, de nouvelles images, parfois inspirées de la science-fiction. C'est une autre façon de concevoir le vivant, la mise en place d'un cadre de pensée où le langage

mathématique paraît le plus pertinent pour appréhender l'activité mentale. L'idée de représentation est au cœur de ces approches qui fabriquent des modèles, c'est-à-dire des représentations schématiques d'un objet, d'un processus ou d'une fonction.

3- Constructivisme, systémique, auto-organisation et enaction

Au sein même des sciences cognitives sont proposées d'autres conceptions souvent inspirées de la phénoménologie qui écarte l'idée de représentation. Le concept central de la phénoménologie est celui d'intentionnalité. Le rapport de la conscience aux phénomènes est décrit comme un rapport intentionnel. Pour la phénoménologie, la forme la plus typique de la connaissance, qui sert du reste de fondement à toutes les autres, est la perception.

Ainsi, à la vision représentationnelle, majoritaire chez les cognitivistes et les connexionnistes, Francisco Varela substituait une alternative, qu'il nommait enaction (Varela, 1989a). Comparant les données obtenues en neurophysiologie et les modèles cognitivistes, Francisco Varela signalait deux failles principales du cognitivisme (Varela, 1989b) : la première a trait à l'utilisation de règles appliquées séquentiellement, alors que les données neurophysiologiques sont davantage en faveur d'un traitement parallèle. La deuxième faille est liée au fait que le traitement symbolique est localisé, c'est-à-dire que la perte ou la détérioration d'une partie des symboles ou des règles conduit à un mauvais fonctionnement, alors que dans le cerveau une réorganisation est possible. Francisco Varela constatait que le modèle connexionniste était plus proche de ce qui est observé en neurophysiologie, mais non complètement satisfaisant. Pendant de nombreuses années, il a fait appel aux conceptions constructivistes proposées dans les années cinquante par le physicien Heinz von Foerster, chef de file de la seconde cybernétique.

3-1- Constructivisme et cybernétique

Postulant que l'observateur modifie le phénomène observé, ce mouvement s'intéresse à l'ensemble observateur/observé et se demande comment éviter que l'observateur vienne perturber l'objet sous observation. Définissant un « principe d'ordre à partir du bruit », von Foerster exposait l'idée de la construction par l'homme de sa réalité. Cette vision constructiviste, présentée en 1949, semble trouver ses sources il y a près d'un siècle lorsque des mathématiciens se sont interrogés sur le problème des fondements des mathématiques et l'origine des nombres (donnés par la nature ou construits par l'homme).

Dans les années cinquante, deux participants aux conférences Macy, von Foerster et l'anthropologue Gregory Bateson⁵, vont contribuer à l'élaboration du constructivisme contemporain. S'inspirant de la théorie de l'information et de la communication, Bateson publie en 1951 un ouvrage qui présente la communication comme la matrice dans laquelle se trouvent enchâssées les activités humaines (Bateson, 1951). En 1952, la Fondation Rockefeller subventionne un programme de recherche sur les paradoxes de l'abstraction dans la communication, à l'Université de Stanford, sous la direction de Bateson qui s'entoure d'une équipe rapidement surnommée le Groupe de Palo Alto. L'objectif du groupe est d'appliquer aux comportements de communication les réflexions de Whitehead et Russell sur les paradoxes logiques et l'auto-

⁵ Gregory Bateson a été à l'origine de l'École de Palo Alto qui a mis en évidence la pluralité des niveaux de communication contenus dans un message. L'approche systémique de la communication utilisée par l'École de Palo Alto est, en particulier, appliquée en thérapie familiale.

référence. Travaillant dans le domaine de la psychiatrie, mais utilisant une approche anthropologique, Bateson étudie le comportement de jeunes schizophrènes et développe sa théorie de la « double contrainte » (*double bind*)⁶. Cette théorie va susciter une recherche thérapeutique connue sous le nom de thérapie systémique. La vision du thérapeute extérieur à la famille fait place à celle où le thérapeute est partie prenante du système dans lequel il intervient.

3-2- Constructivisme, auto-organisation et seconde cybernétique

Un autre courant fondé sur les théories de l'auto-organisation, qui prend sa source dans la seconde cybernétique, émerge dans les années soixante-dix. En 1969, en réaction contre les aspects mécanistes privilégiés par certains cybernéticiens, de sévères critiques sont formulées lors du symposium Alpach intitulé « Au-delà du réductionnisme », organisé par Arthur Koesler. Sont dénoncés pêle-mêle la science mécaniste et réductionniste, la conception du cerveau vu comme un ordinateur, la biologie moléculaire, le *behaviorisme* et la psychanalyse. Ce symposium met en valeur la science générale des systèmes et la pensée systémique y est représentée, entre autres, par Ludwig von Bertalanffy qui, en 1968, a publié la *Théorie Générale des Systèmes*. Lors du symposium, le thème de l'auto-organisation est utilisé pour critiquer la première cybernétique.

C'est à cette époque que le constructivisme s'épanouit via quelques textes fondateurs comme celui de Jean Piaget qui publie en conclusion de son Encyclopédie de la Pléiade : *Logique et connaissance scientifique*, un article intitulé 'Les courants de l'épistémologie scientifique contemporaine', qui restaure le statut des épistémologies constructivistes. Von Foerster, Bateson, Piaget, Watzlawick ont tous contribué à remettre en cause les formes objectivistes du réalisme. Si Kant défendait l'idée que toute connaissance étant dépendante de nos sens et des propriétés de notre esprit, la réalité profonde du monde ne nous est pas accessible et que nous devons par conséquent nous contenter d'analyser l'apparence des choses, il a introduit les jugements synthétiques a priori pour rendre compte des questions épistémologiques posées par l'élaboration de la connaissance par le sujet. Ce sont ces nécessités a priori qui se trouvent remises en question par Piaget qui insiste sur l'idée de construction.

La phénoménologie considère la rencontre sujet/objet comme le point de départ de toute connaissance. C'est à chaque observateur de construire une connaissance phénoménale à partir des effets du réel sur son organisme. La connaissance de la nature devient connaissance des effets de la nature sur l'observateur (Tabary 1991). Jean Piaget défend des thèses qui aboutissent à une conclusion semblable (Piaget 1959). Il reprend la dialectique sujet/objet et fait de la connaissance une construction élaborée par le sujet à partir des relations entre sujet et objet. Selon Piaget, c'est tout le psychisme qui est construit à partir de l'activité cérébrale du sujet en contact avec l'environnement, traduisant simultanément une prise de conscience de soi et une connaissance du réel. Le psychisme relève donc de l'activité biologique et d'une histoire individuelle.

⁶ Dans une situation de double contrainte, une injonction paradoxale est émise. Dans un contexte de relations intenses et dans des situations répétées, celui qui reçoit cette injonction est mis dans l'incapacité de sortir du cadre fixé par ce message paradoxal qui peut engendrer des troubles à caractère pathologique. Gregory Bateson et al. 1956. *Toward a Theory of Schizophrenia*, Behavioral Science, 1, 251-264.

Ces conceptions et la notion d'autonomie vont être développées par Umberto Maturana, Francisco Varela et Henri Atlan. En 1973, Pierre Vendryès définissait explicitement la théorie de l'autonomie (un système autonome est défini comme un système qui obéit à ses propres lois) (Vendryès, 1973). Ce qui permet l'autonomie est le fait pour un organisme de posséder plusieurs façons d'exister, dont l'une au moins assure le maintien de l'identité et de l'équilibre vis-à-vis des conditions présentes de l'environnement. Maturana et Varela vont définir la notion d'autopoïèse comme la propriété des systèmes autonomes complexes à renouveler leurs propres constituants. Un système autopoïétique est décrit comme un réseau de composants qui régénèrent, par leurs interactions et transformations, le réseau qui les a produits. Il engendre et spécifie continuellement sa propre organisation. Selon Varela (Varela, 1989a),

« la notion d'autopoïèse est nécessaire et suffisante pour définir l'organisation des êtres vivants. »

Pour éviter d'utiliser le terme «constructivisme», Francisco Varela conçu la notion d'enaction. Partant du principe que notre sens commun est constamment requis pour configurer notre monde d'objets et que ce dernier résulte de notre histoire, il remettait en question l'idée d'un monde aux propriétés prédéfinies. Avec l'enaction, il proposait une «voie moyenne» (entre positivisme et constructivisme) où observateur et phénomène observé se définissent l'un l'autre. L'idée de base est que les structures cognitives émergent de schèmes sensori-moteurs récurrents qui permettent à l'action d'être guidée par la perception et à la perception d'être co-constituée par le mouvement et l'action. Cette approche, déjà envisagée par Merleau-Ponty, étudie la manière dont le sujet parvient à guider ses actions dans une situation locale donnée. Décrivant le cerveau comme « un organe qui construit des mondes plutôt qu'il ne les réfléchit », il avançait l'idée que les facultés cognitives étaient inextricablement liées au vécu historique.

Mais constatant que ses réflexions ne conduisaient pas directement à des réalisations pratiques, Francisco Varela se plaignait du manque d'appuis institutionnels et critiquait la vision utilitariste. Il s'élevait contre le poids de la technologie qui limite le champ des recherches et concentre les énergies sur ce qui est créateur de marchés immédiats (Varela, 1989b). Il proposait « de s'éloigner des considérations pragmatiques de l'ingénieur », alors même que les sciences cognitives paraissaient se structurer de plus en plus autour des applications pratiques liées aux technologies de l'information.

3-3 Le constructivisme et les sciences de l'artificiel

Un courant constructiviste bien différent de ceux décrits précédemment est représenté par Herbert Simon (Simon, 1969) et Jean-Louis Le Moigne (Le Moigne, 1994) qui utilisent le constructivisme et la théorie des systèmes pour promouvoir les sciences de l'artificiel et les sciences de l'ingénieur. Le Moigne présente l'épistémologie constructiviste comme (Le Moigne, 1994 : 123) :

« une épistémologie de l'invention, ou plus correctement de la poïèse : la production originale par le faire [...] Elle ne vise plus à découvrir le vrai plan de câblage d'un univers dissimulé sous l'enchevêtrement des phénomènes ; elle vise à inventer, construire, concevoir et créer une connaissance projective. »

Il ne s'agit donc plus de découvrir les mécanismes et le fonctionnement du monde dans lequel nous vivons mais d'inventer un autre monde où l'artifice

créé par l'homme serait considéré comme aussi « réel » ou « irréel » que ce qui nous entoure. Remettant en question la notion de « réalité », il devient possible de considérer les entités artificielles construites par l'homme comme des êtres vivants, et c'est ce que font les chercheurs en vie artificielle. Créé en 1984 à l'Institut de Santa Fé, ce domaine de recherche se donnait pour objectif de créer des êtres artificiels sur ordinateur. Le terme de « vie artificielle » renvoie ici à des systèmes computationnels considérés comme vivants. Christopher Langton, un des pionniers de ce domaine, définissait ainsi son objet d'étude (Langton, 1989):

« la vie fabriquée par l'homme plutôt que par la nature. »

Alors que la biologie s'intéresse aux bases matérielles de la vie, la vie artificielle s'occupe de sa forme dynamique sans référence à sa matière. L'objectif de Langton était de dégager les lois générales de la « *vie telle qu'elle pourrait être* ». Il expliquait qu'il s'agissait de :

« voir ce qui se passe quand on met ensemble des choses très simples, comment elles évoluent vers des comportements très compliqués. »

Utilisant les concepts d'auto-organisation, les chercheurs en vie artificielle perçoivent l'organisme vivant comme un ensemble d'éléments simples en interaction et ont recours au concept d'émergence. La vie est conçue comme un processus émergent qui naît de l'organisation des interactions d'un grand nombre de molécules non vivantes. Leur credo est que la vie n'est pas une propriété de la matière mais plutôt de l'organisation de cette matière, et que saisir son essence revient, en fait, à saisir son organisation abstraite, sa logique. La méthode analytique des biologistes est critiquée. Il s'agit ici de générer, dans des systèmes artificiels, des comportements ressemblant à la vie. L'approche qui est choisie s'inscrit dans une mouvance fonctionnaliste.

Comme les cybernéticiens avaient simplifié le comportement des neurones en les assimilant à des appareils à seuil soit actifs, soit inactifs, Stuart Kauffman de l'Institut de Santa Fé propose d'imaginer le génome comme :

« un ordinateur parallèle où les unités élémentaires, les gènes, prennent simultanément les valeurs 1 ou 0 selon qu'ils sont actifs ou inactifs. »

Le jeu consiste à trouver la solution optimale d'un problème via l'utilisation des algorithmes génétiques (AG). Les AG constituent un mode de programmation d'ordinateur qui s'inspire des modèles néo-darwiniens. Le codage du problème s'effectue sous la forme de chaînes de bits qui sont assimilées à la fois à des « chromosomes » et à des « individus ». Afin de déterminer l'optimum d'une fonction mathématique, la démarche consiste à trouver une chaîne binaire adaptée. Dans ce but, les AG accomplissent des opérations selon quatre étapes : évaluation des 'individus' (indice de performance), sélection, reproduction (échange de fragments de 'chromosomes'), remplacement des 'individus-parents' par des 'individus-fils'. On recommence ensuite à l'étape 1 jusqu'à l'obtention d'une solution acceptable au problème à résoudre (Holland, 1992).

Dans ce type de représentation, l'individu est assimilé à un chapelet d'informations et le monde à un lieu hostile où règne la compétition. Tout est fondé sur le déterminisme génétique et l'individualisme compétitif (Helmreich, 1998). Les métaphores aident la simulation à prendre sens et la comparaison

devient rapidement identification. Les chercheurs redéfinissent, selon leurs critères, ce qu'est un individu, un chromosome, la vie. Les distinctions s'estompent entre les entités artificielles qu'ils conçoivent et les modèles naturels qu'ils cherchent à imiter. S'ils s'inspirent en général de la biologie, ils ne tiennent pas compte du fait que les biologistes donnent à un moment donné une définition qui est fonction des découvertes, des avancées techniques, et qui change donc au cours du temps. Par exemple, si on se réfère à la définition d'un virus donnée en 1953, aucune structure particulière n'est précisée (Luria, 1953)⁷, alors qu'en 1978, la définition proposée fait entrer en jeu la structure (Luria, 1978) :

« les virus sont des entités dont le génome est composé d'acide nucléique qui se réplique dans les cellules vivantes utilisant la machinerie synthétique cellulaire et provoquant la synthèse d'éléments spécialisés qui peut transférer le génome viral à d'autres cellules. »

Si on considère le concept abstrait de virus comme entité transmissible nécessitant la machinerie de l'hôte pour sa répllication (définition proposée en 1953), nul besoin de structure spécifique et les informaticiens peuvent s'appropriier le terme de virus. Choisir une définition *ad hoc* puis identifier les entités artificielles aux entités naturelles apparaît un schéma courant dans le monde des chercheurs en vie artificielle.

Le constructivisme utilisé pour mettre en avant les sciences dites de l'ingénieur n'a plus grand chose à voir avec celui prôné par Piaget, von Foerster ou Varela. Il sert de justification à la fabrication d'un monde où l'artifice est assimilé au non-artifice. La volonté d'Herbert Simon et de Jean-Louis Le Moigne de défendre les sciences de l'ingénieur correspond à un besoin de reconnaissance et révèle une certaine révolte contre le peu de prestige des recherches dites « appliquées » qui subissent mépris et condescendance de la part de ceux qui estiment réalisés des recherches dites « fondamentales ». Pourtant, nombre de ces recherches sont moins « pures » qu'il n'y paraît et des recherches dites « finalisées » peuvent apporter des contributions décisives. Pierre Bourdieu nous rappelle que c'est surtout dans les périodes de crise que s'affirment les images stéréotypées (Bourdieu, 1997 : 44). Cette rivalité est liée au fait que :

« les chercheurs « purs » voient bien que la reconnaissance sociale et le poids politique que les chercheurs « appliqués » obtiennent des autorités politiques, et dont témoignent leurs nombreuses participations aux responsabilités et aux pouvoirs temporels (cabinets ministériels...) ont pour contrepartie, bien souvent, des abdications ou des démissions scientifiques et surtout des renoncements à l'autonomie ».

Les oppositions entre deux univers qui ont du mal à se comprendre mettent en évidence l'anxiété des chercheurs qui se sont éloignés des applications pratiques et la révolte des ingénieurs qui se sentent mésestimés. Cet antagonisme se double de celui qui oppose les fonctionnalistes aux défenseurs d'une théorie neuronale de la pensée.

7 « Entités submicroscopiques capables d'être introduites dans des cellules vivantes spécifiques et de se reproduire dans ces cellules ».

4- Théorie neuronale de la pensée

Dans les années 1980-1990, ce sont les conceptions de Gérard Edelman qui ont trouvé, un large écho auprès de la communauté des neurobiologistes. Prix Nobel de médecine en 1972, il a édifié une théorie globale de la mémoire et de la conscience bâtie autour du principe de la sélection progressive des liaisons qui s'établissent entre neurones. Son ouvrage, *Neural darwinism* (Edelman, 1987), constitue une tentative de synthèse entre neurobiologie, évolutionnisme et génétique où il expose la théorie des groupes neuronaux. Il suppose que les mécanismes de la perception et de la mémoire reposent sur le principe que, parmi une infinité de connexions qui peuvent s'établir au cours du développement cérébral, seuls certains réseaux vont être stimulés par les actions du sujet et par les informations qu'il reçoit. Des connexions sont sélectionnées, d'autres disparaissent, d'où le nom de « darwinisme neuronal ». Edelman explique comment les fonctions du cerveau se développent en relation avec celles du corps à mesure que celui-ci se modifie sous l'effet de la croissance et de l'expérience. La construction du cerveau se traduirait par l'établissement de cartes neuronales qui mettraient en connexion des groupes neuronaux.

Avec *Topobiology* (Edelman, 1988), Edelman montre comment les cellules expriment dans le temps et dans l'espace des gènes qui régissent la production de molécules morphorégulatrices, insistant sur l'importance des événements épigénétiques. En 1989, il se penche sur la mémoire et la conscience pour expliquer que l'action de la conscience présente vient modifier la représentation du passé : ainsi le souvenir est perçu comme une invention à partir du présent qui mêle aux images enregistrées le travail de la perception (Edelman, 1989). Enfin, en 1992, il propose une *Biologie de la Conscience* (Edelman, 1992) : cherchant à éviter le réductionnisme qui prétend expliquer les états mentaux par la description des états neurobiologiques, il met l'accent sur les processus d'acquisition et de modification par rétroaction de l'acquis sur les potentialités innées. Alors que les conceptions innéistes dominent chez les fonctionnalistes, les neurobiologistes du développement insistent sur l'importance des facteurs épigénétiques.

Même si certains tentent de dépasser les débats où risquent de s'enliser les sciences de la cognition en réduisant tour à tour la complexité cognitive à telle ou telle conception de l'esprit, on constate deux modèles de références : l'un, technique, le modèle de l'ordinateur avec, comme disciplines associées l'informatique et l'IA ; et l'autre biologique, celui du cerveau exploré par les neurosciences. Dans le second cas, les conceptualisations des phénomènes cognitifs s'organisent autour des recherches sur le système nerveux central. A des recherches de caractère essentiellement biologique ou physiologique s'associent des études fondées sur l'utilisation de réseaux de neurones pour simuler les phénomènes cognitifs.

A la fin des années 1980, la philosophe Patricia Churchland proposa le réductionnisme éliminativiste, c'est-à-dire la réduction des états mentaux aux phénomènes biologiques sous-jacents et l'élimination du niveau psychologique (Churchland, 1986). Jugeant impossible à appréhender objectivement les états mentaux tels que croyance, désir, espoir, crainte, elle critiquait la psychologie et les positions fonctionnalistes adoptées par Jerry Fodor. Patricia Churchland se rappelle avoir été choquée à la lecture de l'ouvrage de Fodor, *Le langage de la pensée*, qui affirme que la connaissance du cerveau est inutile pour comprendre la cognition (Beauvallet, 1996). Alors professeur débutante à

l'Université de Manitoba, elle commença à étudier la neurobiologie. A l'issue de ses études, elle était persuadée que le niveau psychologique n'était pas pertinent pour étudier le fonctionnement de la pensée et que, seule, une science cognitive enracinée dans la neurophysiologie avait une chance de perdurer.

En France, c'est le neurobiologiste Jean-Pierre Changeux qui a défendu une théorie neuronale de la pensée. A la fois directeur d'un laboratoire CNRS de neurobiologie moléculaire à l'Institut Pasteur et professeur au Collège de France, Jean-Pierre Changeux, membre puissant de la communauté scientifique, jouit à la fois d'un pouvoir institutionnel, lié à l'occupation de positions éminentes dans les institutions, et d'un prestige scientifique qui repose sur la reconnaissance des pairs. Très médiatisé, il est considéré comme un expert et un vulgarisateur scientifique de qualité. Dans son livre, *L'Homme neuronal* (Changeux, 1983) qui a eu un impact important, Changeux présentait une théorie neuronale de la pensée et établissait une relation causale entre structure et fonction. Il décrivait la membrane du neurone comme une « machine à calculer arithmétique » et tentait de traduire les conduites de l'homme en termes d'activités neuronales. Ce réductionnisme, nuancé dix ans plus tard, a déclenché des réactions très vives. Grâce au succès des travaux en biologie moléculaire auxquels il a participé pendant la période héroïque des années soixante et à sa position institutionnelle, il a pu peser sur la structuration du domaine des sciences cognitives en 1989 lorsque les ministères de la Recherche et de l'Éducation Nationale se sont adressés à lui pour présider le comité scientifique chargé de rédiger un rapport pour lancer l'action « Sciences de la Cognition ».

Le rôle que Jean-Pierre Changeux a joué dans le processus institutionnel de l'émergence des sciences cognitives en France est à prendre en compte pour comprendre pourquoi, dans ce pays, le développement des sciences cognitives n'a pas suivi le même schéma que celui des mêmes sciences aux États-Unis. Le chapitre suivant vise à rechercher quelles raisons sociologiques, historiques ou conceptuelles ont favorisé ces différences et dans quel contexte cette institutionnalisation s'est produite.

ÉMERGENCE DES SCIENCES COGNITIVES EN FRANCE : ASPECT INSTITUTIONNEL

L'activité scientifique implique un coût économique et le degré d'autonomie d'une science dépend à la fois de son prestige et de ses besoins financiers. Le combat pour l'autonomie est sans cesse présent, autonomie vis-à-vis des pressions extérieures mais autonomie aussi à l'égard des autres domaines scientifiques. Ce type de combat est illustré ici au travers de la construction institutionnelle des sciences cognitives.

1- Les pionniers et leurs difficultés

En France, le terme « science cognitive » n'a été utilisé qu'à partir de 1981, soit vingt ans après les États-Unis. La première association pour la recherche cognitive (ARC)⁸ a été créée cette année-là. Avant cette création un groupe d'informaticiens, de psychologues et de linguistes se sont rencontrés à plusieurs reprises dans les années soixante-dix dans le but de développer une informatique théorique orientée vers la compréhension du langage naturel. Financées par l'IRIA (Institut de Recherche en Informatique et Automatique)⁹, ces rencontres

⁸ Aujourd'hui ARCo.

⁹ L'IRIA est devenu l'INRIA en 1979.

visaient à développer des systèmes experts censés simuler des activités cognitives.

Des chercheurs en informatique, des mathématiciens (Daniel Kayser, Mario Borillo, André Lentin...) et des psychologues (Jean-François Le Ny, François Bresson, Georges Noizet, Jean-François Richard) sont à l'origine de la création de l'ARC. L'objectif de faire émerger la science cognitive en France sous-tendait l'ambition de reprendre les principaux thèmes de la philosophie et de les traiter avec les outils de l'informatique. Il s'agissait d'automatiser les fonctions intellectuelles telles que la représentation des connaissances ou le raisonnement. À l'époque, les protagonistes de ce projet n'ont pas réussi à faire accepter leurs idées par le milieu académique. En France, l'informatique et l'intelligence artificielle n'étaient pas considérées comme des sciences «nobles» et les informaticiens étaient traités comme des techniciens. Le capital de ces précurseurs en termes de pouvoir dans le milieu scientifique académique était insuffisant pour réussir rapidement à imposer leurs conceptions.

Le délai de vingt ans entre l'utilisation première du terme « science cognitive » en France et aux États-Unis a de multiples sources mais celle qui paraît la plus critique concerne l'essor de l'IA dans ces deux pays. Contrairement aux pionniers de l'IA aux États-Unis, les chercheurs en informatique français n'ont pas bénéficié des financements et du prestige de ces derniers. Si aux États-Unis, les informaticiens souffraient également d'un sentiment d'infériorité vis-à-vis des autres disciplines, le contexte historique, l'importance de la technologie et une conception pragmatique de la science constituaient un terreau propice à la construction des sciences cognitives autour de l'informatique et de l'IA. Le pouvoir et les moyens acquis par les pionniers américains de l'IA ont conduit à une concentration des ressources et des ordinateurs les plus performants leur assurant ainsi une avancée considérable dans ce domaine.

Le rejet par un certain nombre de mathématiciens français de la logique et des mathématiques appliquées avant, pendant et après la seconde guerre mondiale (Andler, 1994) a joué un rôle non négligeable. L'influence dominante du groupe Bourbaki entre 1955 et 1965, hostile à la logique et aux mathématiques appliquées, a probablement constitué un frein au développement de l'IA en France. Seules les mathématiques «pures» étaient considérées comme respectables, les échanges avec les autres disciplines étaient bannis, ainsi que la participation directe des mathématiciens aux applications technologiques. De telles conceptions se trouvaient à l'opposé de celles qui avaient cours aux États-Unis.

À part quelques exceptions, comme René de Possel (1905-1974), Jean Kuntzmann (1912-1992), et Marcel-Paul Schützenberger (1920-1996), très peu de mathématiciens français se sont intéressés au calcul et aux ordinateurs. Jean Kuntzmann joua un rôle pionnier dans le développement de l'informatique à Grenoble (Bolliet, 1993). Les interactions entre l'Université de Grenoble, les industries locales et l'existence d'une école d'ingénieurs, l'Institut Polytechnique de Grenoble, stimulèrent l'essor des mathématiques appliquées et du calcul numérique. Kuntzmann commença à enseigner ces disciplines dès 1945. René de Possel était un membre du groupe Bourbaki mais, en 1941, il quitta ce groupe et devint professeur d'analyse numérique à l'université d'Alger. En 1957, il dirigea le laboratoire de calcul numérique du CNRS, favorisant l'acquisition d'ordinateurs, l'enseignement de l'informatique et la recherche dans de nouvelles directions telles que l'intelligence artificielle. De Possel était

tout particulièrement intéressé par la reconnaissance automatique de l'écriture. Médecin de formation, Schützenberger, l'un des pionniers de l'utilisation des statistiques en médecine, a passé une thèse de doctorat d'État en mathématiques en 1953. Son intérêt pour l'informatique et son idée de l'existence de structures algébriques sous-jacentes à tout problème de modélisation et de traitement de l'information ont séduit nombre de ses élèves¹⁰, qui ont fait à leur tour de nombreux disciples. Autour des premiers ordinateurs, un milieu universitaire restreint mais actif s'organisait.

Entre 1957 et 1964, l'arrivée des transistors et de la troisième génération d'ordinateurs utilisant des circuits intégrés aboutit à une augmentation importante du marché dirigé vers les applications de gestion. Cette période correspond aussi à celle des grands programmes (espace, nucléaire, aéronautique) et les États-Unis prennent la tête du marché mondial. L'hégémonie des États-Unis dans ce domaine et l'utilisation de l'informatique comme d'un instrument de domination par les Américains inquiètent alors l'Europe qui commence à réagir. En 1963, en France, la DGRST désigne l'informatique comme une priorité et crée l'action concertée « Calculateurs électroniques ». En 1966, le Plan Calcul, grand programme technologique de la Cinquième République, est créé dans le but de mettre sur pied une industrie informatique capable d'assurer l'indépendance française vis-à-vis des constructeurs américains (Arsac, 1988).

Ces initiatives du gouvernement et les tentatives de quelques chercheurs d'université tentent de favoriser l'informatique en France mais cette discipline est négligée par les grandes écoles, telles que l'École polytechnique. Les chercheurs en informatique sont encore traités comme des techniciens utiles pour les centres de calcul et l'informatique est réduite à la fonction d'aide aux calculs pour les autres disciplines. Tous ces éléments font qu'il est très difficile pour des chercheurs en informatique d'être considérés comme des théoriciens et de faire accepter leur point de vue, d'où les difficultés rencontrées par les pionniers des sciences cognitives.

2- Actions du CNRS et du Ministère de la recherche

En 1984, le CNRS (Centre National de la Recherche Scientifique) commence à s'intéresser aux sciences cognitives. Son directeur général, Pierre Papon, sollicite un sociologue, Dominique Wolton, pour rédiger un rapport sur les sciences de la communication et les sciences cognitives. Un an plus tard, en 1985, Michel Imbert¹¹, neurobiologiste et philosophe, prend l'initiative de contacter Bruxelles pour développer les sciences cognitives en Europe. A la demande de la commission de la Communauté européenne (programme FAST¹²), il écrit un rapport qui va conduire à la création de plusieurs laboratoires, dont celui de Bochum en Allemagne.

Entre 1984 et 1995, plusieurs rapports sont demandés par le CNRS et le ministère de la Recherche pour mettre en œuvre des actions incitatives de développement des sciences cognitives. L'analyse détaillée de ces rapports montre qu'en 1985 l'action dirigée par Dominique Wolton amorce une préfiguration de pôles régionaux autour d'une collaboration entre informaticiens,

10 Ses premiers élèves étaient M. Nivat, J.-F. Perrot, M. Gross, J. Berstel, R. Cori, M. Fliess, D. Perrin.

11 Il a créé à Toulouse en 1993 le Centre de Recherche Cerveau et Cognition.

12 Créé en 1978, le programme FAST (Forecasting and Assessment in the field of Science and Technology) avait une mission de prospective et d'évaluation.

psychologues et linguistes, ceux-là mêmes qui avaient participé à la création de l'ARC. Chercheurs en sciences humaines et informaticiens se trouvaient des intérêts communs autour du thème de l'information et de la communication. Mais, en 1989, le neurobiologiste Jean-Pierre Changeux tente d'orienter les sciences cognitives vers l'étude du cerveau. Président du comité scientifique chargé de rédiger un rapport demandé par les Ministère de la Recherche et de l'Éducation, il assigne aux sciences cognitives l'objectif de mieux comprendre l'organisation et le fonctionnement du cerveau et de s'intéresser tout particulièrement aux facultés humaines les plus élevées. Le rôle de l'informatique est minimisé et l'ARC n'est pas mentionnée. L'importance de l'informatique est pourtant précisée dans le second appel d'offres de l'action concertée et Mario Borillo, chercheur en informatique, fait partie du comité scientifique présidé par Changeux.

Cependant, Jean-Pierre Changeux organise en 1990, avec le neurobiologiste André Holley, un congrès à Lyon fortement orienté vers les neurosciences. La polémique ne tarde guère. La place que Changeux accorde aux sciences du cerveau paraît disproportionnée aux chercheurs en informatique qui se perçoivent comme les pionniers de ce domaine. Deux mois après ce congrès, le CNRS lance le Programme Interdisciplinaire de Recherche (PIR) « Cognisciences » et en confie la direction à André Holley. Le comité scientifique est présidé par Mario Borillo. Ce programme interdisciplinaire participe largement à la structuration en réseaux qui avait été initiée par le programme « Communication » dirigé par Dominique Wolton. Un rapprochement entre les laboratoires de recherche en informatique orientés vers les sciences cognitives se concrétise. Sept réseaux régionaux¹³ sont financés par ce programme et c'est à l'un des responsables du réseau Cogniseine, Alain Berthoz, que le Ministère de la Recherche confie l'organisation d'un grand colloque de prospective sur les sciences de la cognition qui a lieu à Paris en Janvier 1991. Ingénieur de formation, Alain Berthoz¹⁴ est plus proche de l'univers conceptuel des mathématiciens et informaticiens que les neurobiologistes, même si, pour lui aussi, le cerveau constitue l'objet d'étude privilégié. L'accent est mis sur les retombées économiques et sociales des réseaux de neurones, des technologies de l'intelligence artificielle et de la communication.

A la suite du colloque de 1991, Hubert Curien, ministre de la Recherche et de la Technologie, demande un rapport à Bernard Guibert, attaché au service SHS (sciences humaines et sociales) du ministère de la Recherche, pour essayer de répondre à deux questions :

1°) Comment étendre aux sciences sociales les bienfaits de l'interdisciplinarité à l'œuvre de manière spectaculairement féconde dans les sciences de la cognition ?

2°) Comment faire un pas en direction des applications et des transferts de technologie ?

13 Cogniseine, Paris-Centre, Paris-Sud, Cognisud (Aix-en-Provence, Marseille, Nice-Sophia-Antipolis), Grand-Est (Metz, Nancy, Strasbourg), Prescott (Toulouse), Rhône-Alpes (Grenoble, Saint-Etienne, Genève).

14 Après une formation d'ingénieur de l'école des Mines de Nancy, il a fait des études de psychologie, de mathématiques et d'ergonomie. Actuellement, il est professeur au Collège de France et dirige le laboratoire de physiologie de la perception et de l'action. Ces recherches, liées au domaine de l'aérospatial, portent sur le rôle du système vestibulaire, les bases neurales des mouvements d'orientation et la perception visuelle du mouvement.

Le rapport de Bernard Guibert réintroduit une particularité mise en avant dans le rapport de Dominique Wolton : la tentative d'une collaboration entre chercheurs en sciences humaines et chercheurs en informatique.

3- Le programme « Science de la Cognition »

Peu à peu l'informatique se voit attribuer une place plus conséquente jusqu'à acquérir une place de choix en 1995 puisque, pour la première fois, c'est un chercheur en informatique, Jean Gabriel Ganascia¹⁵ qui est désigné pour mettre en place un programme « Sciences de la Cognition », fruit de l'unification des actions du CNRS et des ministères. Ce programme regroupe le ministère de l'Enseignement et de la Recherche, le CNRS, le CEA (Commissariat à l'Énergie Atomique), l'INRIA (Institut National de la Recherche en Informatique et Automatique) et, depuis 1996, l'INRETS (Institut National de Recherche sur les Transports et leur Sécurité).

Les recherches en neurobiologie bénéficiant déjà de financements importants, les chercheurs en informatique qui ont tenté de construire un nouveau domaine focalisé sur la simulation informatique des fonctions cognitives ne pouvaient accepter que des neurobiologistes les privent des moyens qu'ils estimaient leur être dus et d'une légitimité qu'ils pensaient avoir gagnée à force de persévérance dans un milieu culturel et institutionnel peu propice.

Cependant, ce n'est pas à l'un des pionniers à l'origine de la création de l'ARC que la direction du programme « Science de la Cognition » a été confiée mais à Ganascia, chercheur en informatique qui a passé sa thèse sur les systèmes expert en 1982 et qui, en 1988, avait été nommé chargé de mission à la direction du département des sciences physiques pour l'ingénieur. La lecture de son projet, rédigé en 1995, révèle une tentative de réhabilitation du statut de l'ingénieur et des recherches dites appliquées. En 1990, cinq ans avant l'écriture de son projet, Ganascia publiait *L'âme-machine*, un livre centré autour de l'IA, s'attachant à « reconstruire l'esprit sans faire référence aux processus physiologiques ». Il proposait une histoire de l'évolution de l'IA vers les sciences cognitives et présentait un récit héroïque qui faisait de l'IA une science dont l'évolution était comparée à celle d'un être humain : la période embryonnaire, suivie de l'exubérance, l'enthousiasme et les excès de jeunesse, s'achevant par la maturation et la sagesse. En 1996, il publie un ouvrage sur l'IA destiné au grand public, qui reprend cette construction d'une histoire mythique de l'évolution de l'IA (Ganascia, 1996). Le début ressemble à un conte de fée :

*« A sa naissance, en 1956, les fées se sont penchées sur son berceau.
On lui attribuait toutes les qualités, on prophétisait : elle devait transformer le travail, changer la vie et donc changer le monde.
C'était la période des prophètes. »*

Ce récit mêle des arguments d'ordre « scientifique » mettant l'accent sur les fondements théoriques, philosophiques, logiques et mathématiques, et des arguments d'ordre utilitaire et pragmatique, visant les industriels et les politiques. Les pressions de concurrence internationale ayant produit une cristallisation autour de l'intelligence artificielle et les responsables politiques s'étant rendu compte que les sciences cognitives pouvaient se structurer autour

15 Professeur à l'université Pierre et Marie Curie, Jean-Gabriel Ganascia anime l'équipe l'ACASA au Laboratoire d'Informatique de Paris 6 (LIP6), CNRS.

des technologies de l'information, enjeu devenu majeur, les chercheurs en informatique ont pu faire entendre leur voix.

On note une nette orientation vers les recherches finalisées, répondant ainsi à la demande des pouvoirs publics. En effet, depuis 1994, date de la consultation nationale sur les grands objectifs de la recherche française lancée par François Fillon, se précise une volonté toujours plus grande de contrôle de l'État, de programmation et d'orientation de la recherche. Parallèlement à une volonté gouvernementale de rapprocher laboratoires publics et entreprises et de piloter la recherche, de plus en plus de scientifiques sont sensibilisés au caractère utilitaire de leurs travaux. A force d'entendre parler de demande sociale et d'efficacité, ils finissent par se fixer des objectifs à court terme, d'autant que l'obtention de financement pour ce type de recherche est plus aisée.

4- Les actions menées entre 1999 et 2003

A la fin des années 1990, ce sont les sciences humaines et sociales et la linguistique qui ont été mises à l'honneur. Le CNRS a créé le GDR « diversité et évolution des langues » (1999-2002) et le ministère de la Recherche a lancé l'action concertée incitative Cognitive (1999-2003) dont la direction a été confiée à Catherine Fuchs¹⁶. L'exemple du langage a permis d'illustrer l'intérêt des pratiques interdisciplinaires associant linguistique, psycholinguistique, neurologie, informatique mais aussi orthophonie, phonétique et philosophie.

A tour de rôle, chacun des trois secteurs, sciences de l'informatique, sciences du cerveau et sciences humaines et sociales, s'est trouvé représenté de façon privilégiée au cours des différentes actions lancées par les pouvoirs publics avec, à chaque fois un peu plus de recherches en commun. Aujourd'hui, si l'étude de certaines fonctions cognitives arrivent à fédérer des recherches interdisciplinaires, les différentes problématiques théoriques et pratiques, et les rivalités continuent à constituer des obstacles à une réelle interdisciplinarité mais de nouveaux équilibres s'établissent. Les sciences de l'information et les pratiques de modélisation sur ordinateur s'épanouissent au sein d'un nouveau département créé au CNRS en octobre 2000, le département des sciences et technologies de l'information et de la communication (STIC)¹⁷ et un institut de Cognitive a été créé en août 2003 à Bordeaux. Les neurosciences cognitives, qui se présentent comme une alliance entre neurosciences et psychologie cognitive, se pratiquent dans différents laboratoires et, notamment à l'Institut des sciences cognitives de Lyon (Jeannerod, 2004).

Sous l'influence de la psychologie¹⁸, de nouveaux protocoles expérimentaux sont élaborés par des chercheurs en neurosciences pour aborder l'étude des fonctions cognitives et le développement des techniques de neuro-imagerie oriente les recherches vers l'étude de l'activation des régions cérébrales lorsqu'une fonction cognitive particulière est sollicitée. Ainsi des fonctions telles que l'empathie, capacité à partager les émotions avec autrui, qui n'étaient pas auparavant étudiées en neurosciences, le sont à présent¹⁹.

16 Directrice de recherche CNRS, Catherine Fuchs dirige le LaTTICe (Langues, Textes, Traitement Informatique, Cognition), UMR 8094, CNRS, ENS-Ulm, Paris VII (ENS-Montrouge). Alain Berthoz était le président du conseil scientifique de l'action Cognitive.

17 Arrêté du 5 octobre 2000 modifiant l'arrêté du 10 mai 1991 portant création de départements scientifiques au CNRS.

18 Voir R. Plas, Comment la psychologie expérimentale française est-elle devenue cognitive ? La Revue pour l'histoire du CNRS, n°10, 2004, 24-33.

19 Voir J. Decety, Naturaliser l'empathie, L'Encéphale, 2002, 9-20.

CONCLUSION

L'analyse du processus institutionnel des sciences cognitives montre que les frontières de ce domaine sont soumises à négociation entre disciplines et que des facteurs économiques et politiques contribuent à en modifier les contours.

Que ce soit en France ou aux États-Unis, les premiers scientifiques engagés dans la construction des sciences cognitives ont été des chercheurs en informatique, des psychologues et des linguistes. Le besoin de reconnaissance des uns et des autres s'est traduit par la volonté de créer un nouveau domaine avec pour objet d'étude la connaissance. Considérée au début comme une simple technique, l'informatique, intégrée au sein des sciences cognitives a acquis prestige et reconnaissance sociale.

Les neurobiologistes n'ont commencé à s'investir dans ce domaine qu'au milieu des années 1980. A la fin des années 1980, un neurobiologiste français a tenté d'orienter les sciences cognitives vers l'étude du cerveau mais la mobilisation des chercheurs en informatique et une politique dirigée vers les transferts technologiques ont favorisé, dans les années 1990, l'essor de l'IA. A la fin des années 1990, les sciences cognitives adoptaient deux grandes orientations, l'une vers les modèles informatiques et les applications technologiques et l'autre vers les neurosciences cognitives qui explorent un modèle biologique associant pensée et cerveau. Des échanges entre ces deux pôles et des chercheurs en sciences humaines et sociales se concrétisent parfois, notamment autour des techniques d'imagerie cérébrale, ou de l'étude du langage, mais des divergences persistent. L'acquisition de l'autonomie et des financements induit une concurrence qui n'est pas sans effet sur les théories et les pratiques.

Alors même que les chercheurs se proposent de rassembler différentes disciplines pour aborder l'étude des fonctions cognitives, les antagonismes se concrétisent. Les pouvoirs publics, en réclamant toujours plus de travaux présentant un caractère utilitaire et rapidement applicables, exacerbent les rivalités. Ayant besoin de financement, les chercheurs s'orientent vers des projets à court terme favorisant l'automatisation, la production d'outils de communication, de jeux et de mondes virtuels créateurs de marchés.

Le rôle de la cybernétique dans cette orientation n'est pas négligeable. Cette tentative de rassembler des chercheurs de disciplines différentes pour comprendre le fonctionnement de la pensée a donné lieu au développement de courants riches et parfois divergents. Intelligence artificielle, cognitivisme, fonctionnalisme, connexionnisme mais aussi une conception auto-constructive du savoir semblent s'enraciner dans les efforts des cybernéticiens. Les théories de la communication et du contrôle, la théorie de l'information, le principe de rétroaction, une certaine vision mécaniste de la pensée, la conception de l'homme, machine auto-régulée, sont les fruits de la première cybernétique dont a émergé l'intelligence artificielle, domaine qui s'en est détaché pour adopter une approche plus radicale. Ce qui est connu sous le nom de « seconde cybernétique » a produit la « métascience » des systèmes qui a laissé son empreinte sur l'ensemble des sciences humaines et sociales (thérapie familiale, anthropologie culturelle, économie, sociologie, science politique...). Les idées et les outils produits par la première et la seconde cybernétique ont été utilisés par les chercheurs en sciences cognitives qui se sont nourris de courants divers et qui ont adopté, en fonction de leurs inclinations, philosophie analytique ou phénoménologie.

Cet article met en évidence certains enjeux sous-jacents à la construction de ce domaine : enjeux disciplinaires, conceptuels, économiques, politiques et sociaux. En ce qui concerne les aspects économiques et politiques : si la cybernétique est née dans un contexte de guerre et que l'essor de l'intelligence artificielle aux États-Unis a largement bénéficié des financements militaires pendant la période de la guerre froide, les sciences cognitives se nourrissent de financements fournis dans un contexte de guerre économique où, d'une part, les technologies de l'information et de la communication apparaissent comme les armes à maîtriser et, d'autre, part, l'étude du fonctionnement du cerveau constitue un enjeu important.

En France, lorsque des hommes politiques se sont rendu compte que les sciences cognitives pouvaient se structurer autour des technologies de l'information, les financements ont été trouvés. Les pressions de concurrence internationale et l'économie de marché ont favorisé un domaine orienté vers la production de machines toujours plus performantes visant à remplacer l'homme dans divers tâches ou à fabriquer des jeux et des mondes virtuels qui peuvent se vendre à des millions d'exemplaires. Ces produits modifient en retour les habitudes, les modes de pensée ou les réactions des individus rendus de plus en plus dépendants des objets techniques.

Une intrication extrême est observée entre l'impact des pressions socio-économiques et politiques sur le développement de la science et, inversement, de la science sur le mode de vie. Science et technologie sont intégrées dans une stratégie politique et se trouvent au centre des choix nationaux et internationaux. Science et politique sont liées par le biais de la fonction dévolue à la science de résoudre des problèmes. Pierre Bourdieu (Bourdieu, 1997: 48) explique que :

« l'un des grands paradoxes des champs scientifiques, c'est qu'ils doivent en grande partie leur autonomie au fait qu'ils sont financés par l'État, donc placés dans une relation de dépendance d'un type particulier, à l'égard d'une instance capable de soutenir et de rendre possible une production qui ne soit pas soumise à la sanction immédiate du marché [...]. Cette dépendance dans l'indépendance (ou l'inverse) ne va pas sans ambiguïtés, puisque l'État qui assure les conditions minimales de l'autonomie est aussi en mesure d'imposer des contraintes génératrices d'hétéronomie et de se faire l'expression ou le relais de la contrainte de forces économiques dont il est censé libérer. »

Pour acquérir une certaine autonomie vis-à-vis des autres domaines scientifiques, les artisans de la construction des sciences cognitives ont besoin de financement mais pour en obtenir, ils doivent répondre à une demande qui les place dans une position de dépendance vis-à-vis des pressions du marché économique.

Si la composante intelligence artificielle est importante, elle ne constitue qu'un des aspects de la construction des sciences cognitives. Un autre versant concerne les neurosciences et, dans ce cas, la pensée est perçue comme étant liée au fonctionnement cérébral et l'individu est parfois assimilé à un ensemble de molécules. Nombre de neuroscientifiques laissent espérer à terme une explication complète de l'esprit par la biologie, son entière naturalisation (Ehrenberg, 2004).

Les extrêmes ainsi posés : réductionnisme logique versus réductionnisme neurobiologique, des conceptions intermédiaires se multiplient, souvent en relation directe avec la formation initiale des chercheurs. Leurs pratiques influent sur leurs conceptions : biologistes, ils ont tendance à raisonner en termes de molécules ; informaticiens, en termes de simulation sur ordinateur. Pratiques et théories se conjuguent pour produire des schémas qui tentent de redéfinir ce qu'est un être humain, la pensée, la vie.

J'ai passé en revue brièvement, et très incomplètement, les différentes théories élaborées. J'en donne probablement une image déformée parce que partielle mais je ne cherchais ici qu'à dégager ce que ces conceptions admettent, ou impliquent, en ce qui concerne le fonctionnement de la pensée.

En même temps que les chercheurs élaborent ces théories, ils dessinent le portrait d'une société qu'ils s'essaient à construire. Certains considèrent parfois les hommes comme des machines qu'on peut contrôler, programmer, manipuler ; avec des données qui entrent et sortent comme dans un ordinateur. Steven Rose (Rose, 1973) a déjà évoqué le piège qui consiste à penser les cerveaux comme des ordinateurs, piège du même ordre que celui qui consiste à prendre les hommes pour des singes programmés génétiquement. Les dangers potentiels de la psychopharmacologie ont déjà été largement débattus (Rose, 1973). Les excès engendrés par la création de mondes virtuels ne sont pas moins préoccupants (Chamak, 1996a, 1996b).

Pour la plupart des chercheurs en intelligence artificielle des années soixante dix et quatre-vingt, les fonctions cognitives étaient assimilées à des expressions logiques, le sujet était perçu comme un ensemble de fonctions et les fonctions comme des processus de traitement de l'information. Un objet était choisi : la connaissance, l'esprit ou la vie et une définition en était donnée de telle façon que la meilleure approche pour aborder le problème soit celle des logiciens-informaticiens. Ce processus contribuait à rehausser le prestige social de l'intelligence artificielle. Intégrée dans les sciences cognitives, elle devenait une force majeure qui cherchait à imposer aux autres partenaires les logiques propres à ses chercheurs.

Formalisation et modélisation mathématique étant considérées comme le stade ultime de la scientificité, c'était ce vers quoi devaient tendre les sciences cognitives. L'approche qui consiste à redéfinir un objet d'étude pour qu'il puisse être traité par le biais de la simulation sur ordinateur y était privilégiée. Certains ont parlé de révolution scientifique, ce n'en est pourtant pas une au sens que lui attribue Thomas Kuhn : ce n'est pas une crise déclenchée par des anomalies, mais plutôt un combat pour imposer une approche et une nouvelle vision où la tentative de découverte des mécanismes de la nature fait place à la volonté d'inventer la nature. En raisonnant en terme d'échanges d'informations, biologie moléculaire et sciences cognitives, deux domaines en concurrence, ont distillé une nouvelle façon de concevoir la vie et se sont orientés vers l'assimilation du vivant à l'artificiel. Aujourd'hui, à cause des revers subis, des critiques et surtout d'une connaissance accrue des difficultés rencontrées pour simuler les fonctions cognitives, une place plus importante est accordée à d'autres types de modélisation que le strict logico-mathématique et laisse entrevoir une remise en question de la prétention à un savoir global sur le monde. Se forge ainsi une prise de conscience qu'une approche, un niveau d'observation ou d'analyse conduit à certaines conclusions, qui ne sont que partielles. Cette problématique laisse la place à la question de l'interprétation

des sciences et à la responsabilité du citoyen, le scientifique ne pouvant plus prétendre à détenir une vérité absolue.

Bibliographie

- Andler, M. (1994). Les mathématiques à l'École normale supérieure au XX^e siècle : une esquisse. In *École normale supérieure : le livre du bicentenaire*, dirigé par Jean-François Sirinelli, PUF : 351-404.
- Andreevsky, E. (éd.). (1991). *Systémique et cognition*. Paris : Dunod.
- Arsac, J. (1988). Des ordinateurs à l'informatique. *Histoire de l'informatique en France*, colloque, Grenoble, vol. 2, P. Chatelain (éd.) : 31-44.
- Bateson, G. (1951). *Communication : The Social Matrix of Psychiatry*.
- Bateson, G. et al., (1956). Toward a Theory of Schizophrenia. *Behavioral Science*, 1, 251-264.
- Beauvallet, G. (1996). *Un voyage d'exploration en sciences cognitives*. Paris : L'Harmattan.
- Bolliet, L. (1993). Jean Kuntzmann (1912-1992): un extraordinaire pionnier, troisième colloque : *Histoire de l'informatique*, Sophia Antipolis, France, INRIA.
- Bourdieu, P. (1997). *Les usages sociaux de la science : Pour une sociologie clinique du champ scientifique*. INRA éditions.
- Chamak, B. (1996a). La « réalité virtuelle » et le jeu, ou comment changer l'homme ? *Alliage*, n°26, 55-65.
- Chamak, B. (1996b). Ceux qui construisent la « réalité virtuelle... » *Alliage*, n°27, 44-52.
- Changeux, J.-P. (1983). *L'homme neuronal*. Fayard.
- Churchland, P. (1986). *Neurophilosophy*. MIT Press/Bradford books.
- Dennett, D. (1994). *La conscience expliquée*. Paris : Odile Jacob.
- Dreyfus, H. (1972). *What Computer Can't Do: A Critique of Artificial Intelligence*. New York, Harper & Row. (Trad. franç. 1984. *Intelligence artificielle : mythes et limites*. Flammarion).
- Dupuy, J.-P. (1994). *Aux origines des sciences cognitives*. Paris : La Découverte.
- Edelman, G. (1987). *Neuronal Darwinism: the Theory of Neuronal Group Selection*. New York: Basic Books.
- Edelman, G. (1988). *Topobiology: an Introduction to Molecular Embryology*. New York: Basic Books.
- Edelman, Gerald (1989). *The Remembered Present: a Biological Theory of Consciousness*. New York: Basic Books.
- Edelman, G. (1992). *Biologie de la conscience*. Paris : Odile Jacob.
- Ehrenberg, A. (2004). Le sujet cérébral. *Esprit*, nov 2004.
- Fodor, J. (1975). *The language of Thought*. Brighton: The Harvester Press.
- Fodor, J. (1983). *The Modularity of Mind*. Cambridge (MA): MIT Press. (Trad. franç. 1986, *La modularité de l'esprit*, Paris, Minuit).
- Ganascia, J.-G. (1996). *L'intelligence artificielle*. Flammarion, Dominos.
- Gardner, H. (1993). *Histoire de la révolution cognitive : la nouvelle science de l'esprit*. Paris : Payot.
- Heims, S. (1991). *Constructing a Social Science for Postwar America. The Cybernetics Group: 1946-1953*. MIT press.
- Helmreich, S. (1998). Recombination, Rationality, Reductionism and Romantic Reactions: Culture, Computers, and the Genetic Algorithm. *Social Studies of Science*, 28/1: 39-71.
- Holland, J. (1992). Les algorithmes génétiques. *Pour la Science*, n°179.
- Jeannerod, M. (2004). La création de l'Institut des sciences cognitives du CNRS (1992-1998). *La Revue pour l'histoire du CNRS*, n°10, 16-22.
- Johnson-Laird, P. (1988). *The computer and the Mind: An Introduction to Cognitive Science*. Harvard University Press. (Trad. franç. 1994, *L'ordinateur et l'esprit*. Odile Jacob).

- Kuhn, T. (1962). *The Structure of Scientific Revolution*. The University of Chicago Press. (Trad. franç. 1983. *La structure des révolution scientifique*. Paris, Flammarion).
- Langton, Ch. (ed.). (1989). *Artificial Life*. Addison-Wesley.
- Le Moigne, J.-L. (1994). *Le constructivisme*. ESF éditeurs.
- Lighthill, J. et al. (1973). *Artificial Intelligence: a Paper Symposium*. London: Science Research Council.
- Luria, S. E. (1953). *General Virology*. New York: John Wiley.
- Luria, S. E., et al. (1978). *General Virology*. Third edition, New York: John Wiley.
- McCorduck, P. (1979). *Machines who think*. San Francisco: Freeman.
- McCulloch, W. & Pitts, W. (1943). Logical Calculus of the Ideas Immanent in Nervous Activity. *Philosophy of Science*, vol. 10, n°1, 18-24.
- Minsky, M. & Papert, S. (1969). *Perceptrons*. Cambridge: The MIT Press.
- Minsky, M. (1975). A framework for representing knowledge. In P.H. Winston (ed.), *The psychology of Computer Vision*. New York: McGraw-Hill.
- Minsky, M. (1985). *The Society of Mind*. (Trad. franç. 1988. *La société de l'esprit*, InterEdtions).
- Piaget, J. (1959). *La naissance de l'intelligence chez l'enfant*. Neuchâtel : Delachaux et Niestlé.
- Polyshyn, Z. (1984). *Computation and Cognition: Toward a foundation for Cognitive Science*, Mass., MIT Press.
- Rose, S. (1973). *The conscious brain*. (Trad. franç. 1975. *Le cerveau conscient*, Seuil).
- Rosenblatt, F. (1962). *Principles of Neurodynamics. Perceptrons and the Theory of Brain Mechanisms*. Spartan Books, 1962. Cf. Anderson, J.A. et Rosenfeld E. 1988. *Neurocomputing. Foundations of Research*, The MIT Press.
- Rosenblueth, A., Wiener, N. & Bigelow, J. (1943). Behavior, Purpose and Teleology. *Philosophy of Science*, vol. 10, n°1, 18-24. (Trad. franç. in *Les Études Philosophiques*. 1961. n°2).
- Searle, J. (1980). Minds, brains and programs. *Behavioral and Brain Sciences*. (*Du cerveau au savoir*, 1985, Editions Hermann).
- Simon, H. (1969). *Sciences of the Artificial*. The MIT Press.
- Tabary, J.-C. (1991). Cognition, systémique et connaissance. in Evelyne Andreewsky et coll. (ed.), *Systémique et Cognition*, Paris : Dunod.
- Turing, A. (1936). On Computable Numbers, with an application to the entscheidungs problem. *Proceedings of the London Mathematical Society*, 42.
- Varela, F. (1989a). *Autonomie et connaissance*. Paris : Seuil.
- Varela, F. (1989b). *Connaitre les sciences cognitives : tendances et perspectives*. Paris, Seuil.
- Vendryès, P. (1973). *Vers la théorie de l'homme*. Paris : PUF.
- von Bertalanffy, Ludwig (1968). *General System Theory*. New York. (trad.franç. 1973, Dunod).
- Weizenbaum, J. (1976). *Computer Power and Human Reason*. San Francisco: Freeman (Trad. franç. 1981. *Puissance de l'ordinateur, Raison de l'homme*, Paris, Ed. Informatique).