

années où les STC firent leur apparition. Et cette décennie fertile engendra bien d'autres choses encore qui n'apparaissent pas dans mon bref compte rendu. Mais quand vint 1956, les principaux acteurs de la cybernétique, dont l'unité et la vitalité avait été décisive, se trouvèrent dispersés et plusieurs d'entre eux moururent, de sorte que le flambeau de leurs idées, sur l'esprit en tant que mécanisme, allait devoir être porté par d'autres.

3. Deuxième étape. Les symboles : l'hypothèse cognitiviste

L'entrée en scène des cognitivistes

Si la décennie des années quarante vit naître la cybernétique, on peut dire que la seconde phase des STC date de 1956. Cette année-là, lors de deux conférences tenues l'une à Cambridge et l'autre à Dartmouth, de nouvelles voix se firent entendre, comme celle de Herbert Simon, Noam Chomsky, Marvin Minsky et John McCarthy, qui avancèrent des idées dont les sciences cognitives modernes allaient faire les grands axes de leur développement⁹.

La principale intuition qui devait s'imposer au cours de ces conférences était que l'intelligence (y

9. Voir H. Gardner, *op. cit.*, chapitre 5, pour cette période.

compris l'intelligence humaine) est tellement proche de ce qu'est intrinsèquement un ordinateur que la cognition peut être *définie* par la *computation** de représentations symboliques. Il est évident que cette perspective ne pouvait apparaître qu'après la première étape des STC, à laquelle on doit la notion même de traitement computationnel. Ce qui ne semblait encore qu'une orientation possible — l'esprit comme une forme de logique, donc assimilable au comportement d'un ordinateur — est alors promu au statut d'*hypothèse* reconnue qu'on cherche à démarquer de ses antécédents pluridisciplinaires, c'est-à-dire de l'influence des sciences sociales et biologiques et de leurs complexités diverses. Le *cognitivism*¹⁰ est une désignation commode pour cette orientation vaste mais bien délimitée, qui a motivé plusieurs développements scientifiques et technologiques depuis 1956, dans les secteurs de la psychologie, de la linguistique, d'une large part des neurosciences, et, bien sûr, de l'intelligence artificielle. D'autres désignations sont parfois utilisées, comme

* Le terme est employé ici au sens anglo-saxon de traitement par ordinateur. De même, il sera plus loin question du caractère *computationnel* de ce qui réfère à la computation (N.d.T.).

10. Ce terme est justifié dans J. Haugland (ed.), *Mind Design*, MIT Press, 1981.

le *computationnalisme* (préféré par Jerry Fodor) ou le traitement symbolique.

Ébauche de la doctrine

Que veut dire exactement l'idée que la cognition peut être définie par computation? Le traitement computationnel est une opération qui est effectuée sur des *symbols*, c'est-à-dire sur des éléments qui *représentent* ce à quoi ils correspondent. La notion en jeu ici est la représentation, ou l'*intentionnalité*, terme du philosophe pour la qualité de ce qui est « à propos de quelque chose¹¹ ». L'argument cognitiviste est que le comportement intelligent présuppose la faculté de représenter le monde d'une certaine façon. Ainsi, nous ne pouvons pas expliquer le comportement cognitif à moins de présupposer qu'un agent réagisse en représentant les éléments pertinents des situations dans lesquelles il se trouve. Dans la mesure où sa représentation de la situation est fidèle, le comportement de l'agent sera adéquat, toutes choses étant égales par ailleurs.

Cette acception de la représentation est — au moins depuis le rejet du behaviorisme — à peu près unanimement admise. Ce qui l'est moins est l'étape

11. Pour plus d'information sur ce sujet, voir J. Searle, *Intentionnalité*, Paris, Éditions de Minuit, 1985.

suivante : l'hypothèse cognitiviste prétend que la seule façon de rendre compte de l'intelligence et de l'intentionnalité est de postuler que la cognition consiste à agir sur la base de représentations qui ont une *réalité physique sous forme de code symbolique dans un cerveau ou une machine*.

Selon le cognitiviste, le problème qu'il faut résoudre est le suivant : comment corréler ce qu'on impute aux états intentionnels ou représentationnels (par exemple des croyances, des désirs, des intentions, etc.) avec les changements physiques que subit un agent lorsqu'il agit. En d'autres termes, si nous prétendons que les états intentionnels ont des propriétés *causales*, il nous faut montrer non seulement comment ces états sont physiquement possibles, mais aussi comment ils peuvent déterminer un comportement. C'est ici que la notion de *computation symbolique* intervient : les symboles ont une réalité à la fois physique et sémantique et la computation est conditionnée par cette réalité sémantique. En d'autres termes, la computation est fondamentalement sémantique ou représentationnelle : sans prendre en compte les relations sémantiques entre les différentes expressions symboliques, la notion même de computation — par opposition au traitement aléatoire des symboles — n'a aucun sens. (C'est ce que veut dire l'expression courante « pas de computation sans représentation ».) Un ordinateur, cependant, ne manipule que la forme physique des symboles. Il n'a aucun accès à leur valeur sémantique.

Ses opérations sont néanmoins sémantiquement contraintes car toutes les distinctions sémantiques en jeu dans une computation sont exprimées par le programmeur au moyen de la *syntaxe* du langage utilisé. Car dans un ordinateur, la syntaxe reflète ou est parallèle à la projection sémantique. Le cognitiviste prétend alors que ce parallélisme démontre la réalité physique et mécanique de l'intelligence et de l'intentionnalité (sémantique). L'hypothèse est donc que les ordinateurs offrent un modèle mécanique de la pensée, ou, en d'autres mots, que la pensée s'effectue par une computation physique de symboles. Les sciences cognitives sont donc devenues l'étude de systèmes cognitifs constitués de symboles physiques¹².

Pour bien comprendre cette hypothèse, il est crucial de prendre en compte le *niveau* auquel elle se situe. Le cognitiviste ne prétend pas qu'un regard à l'intérieur d'un cerveau humain nous montrerait de petits symboles en cours de manipulation. Quoique la dimension symbolique comporte une réalité physique, elle n'est pas réductible au niveau physique. (Ceci devient évident quand on se souvient que le

12. Voir Alan Newell, « Physical Symbol Systems », *Cognitive Science*, 4 : 2, 1980, p. 135-183 ; ainsi que Herbert Simon, « Computer Science as Empirical Inquiry : Symbols and Search », réédité dans John Haugeland (ed.), *Mind Design, op.cit.* ; et Zenon Pylyshyn, *Computation and Cognition : Toward a Foundation for Cognitive*, Cambridge, Mass., Bradford Books/MIT Press, 1984.

même symbole peut être associé à de multiples formes physiques.) En conséquence, une expression symbolique comportant une réalité physique peut très bien correspondre à un schéma global d'activité cérébrale d'une grande distributivité. Nous reviendrons plus tard à cette idée. Pour l'instant, ce qu'il faut comprendre, c'est qu'en plus des niveaux de la physique et de la neurobiologie le cognitivisme postule un niveau symbolique distinct irréductible. De plus, puisque les symboles sont des éléments sémantiques, les cognitivistes postulent aussi un troisième niveau proprement sémantique ou représentationnel. (Le fait que la même valeur sémantique puisse être liée à de nombreuses formes symboliques fait aussi apparaître l'irréductibilité de ce niveau¹³.)

Le concept multi-niveaux dans l'explication scientifique est très récent et représente une des principales innovations des sciences cognitives. Les origines et la formulation initiale de cette idée comme notion scientifique d'ordre général remontent à l'époque de la cybernétique, mais les cognitivistes ont largement contribué à développer son articulation philosophique rigoureuse¹⁴. Nous souhaiterions

que le lecteur s'en souvienne plus tard, lorsque il sera question au prochain chapitre de la notion apparentée — bien que toujours controversée — d'*émergence*.

Un autre aspect de l'hypothèse cognitiviste que le lecteur doit noter est qu'elle implique une forte pré-supposition quant aux relations entre syntaxe et sémantique. Comme nous l'avons dit, dans un programme informatique la syntaxe du code symbolique reflète sa sémantique. Mais dans le cas du langage humain, il est loin d'être évident que toutes les distinctions pertinentes pour l'explication scientifique d'un comportement puissent correspondre à une structure syntaxique. En fait, plusieurs arguments philosophiques s'opposent à cette idée¹⁵. De plus, bien que nous sachions d'où vient le niveau sémantique

Science Considered as a Working Hypothesis », et « Computation and Reduction », réédité dans son *Representations : Philosophical Essays on the Foundations of Cognitive Science*, Cambridge, Mass., Bradford Books/MIT Press, 1981.

15. Pour une argumentation venant de la philosophie analytique, voir Hilary Putnam, « Computational Psychology and Interpretation Theory », réédité dans son *Realism and Reason : Philosophical Papers*, vol 3, Cambridge University Press, 1983. Pour une critique de ce point de vue depuis celui de l'approche de l'enaction, voir Terry Winograd et Fernando Flores, *Understanding Computers and Cognition : A New Foundation for Design*, New Jersey, Ablex Press, 1986. Ce problème est aussi la base de l'ingénierie et maintenant célèbre expérience imaginaire dite de la « chambre chinoise » de John Searle, dans « Minds, Brains, and Programs », réédité dans John Haugeland (ed.), *Mind Design*, op. cit.

13. L'irréductibilité du niveau sémantique est en fait l'objet d'une controverse entre cognitivistes. Voir Stephen Stich, *From Folk Psychology to Cognitive Science : The Case Against Belief*, Cambridge, Mass., Bradford Books/MIT Press, 1983 ; et Jerry Fodor, *Psychosemantics : The Problem of Meaning in the Philosophy of Mind*, Cambridge, Mass., Bradford Books/MIT Press, 1987.

14. Voir Jerry Fodor, « Special Sciences, or the Disunity of

tique d'une computation par ordinateur (le programmeur), nous n'avons aucune idée de la provenance du sens des expressions symboliques dont le cognitiviste suppose le cerveau habité.

Le programme de recherche du cognitiviste peut maintenant être résumé par les réponses aux questions suivantes :

Question 1 : Qu'est-ce que la cognition ?

Réponse : Le traitement de l'information : la manipulation de symboles à partir de règles.

Question 2 : Comment cela fonctionne-t-il ?

Réponse : Par n'importe quel dispositif pouvant représenter et manipuler des éléments physiques discontinus : des symboles. Le système n'interagit qu'avec la forme des symboles (leurs attributs physiques), et non leur sens.

Question 3 : Comment savoir qu'un système cognitif fonctionne de manière appropriée ?

Réponse : Quand les symboles représentent adéquatement quelque aspect du monde réel, et que le traitement de l'information aboutit à une solution efficace du problème soumis au système.

Il est évident que le programme cognitiviste que nous décrivons ici n'est pas apparu tel quel, comme Athéna sortant de la tête de Zeus. Nous le voyons maintenant avec un recul de trente ans. Cepen-

42

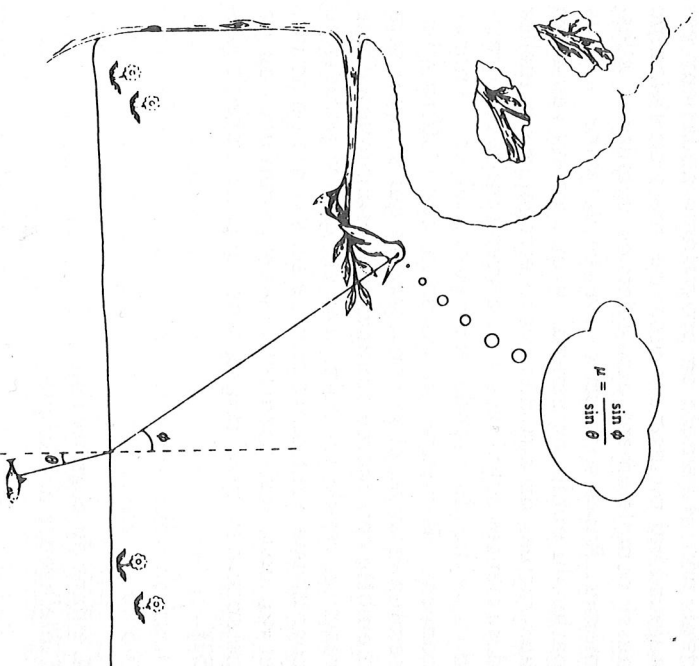


Figure 3 : Un dessin de *Pynch* illustrant succinctement l'hypothèse cognitiviste. Pour capturer sa proie, un martin-pêcheur doit avoir dans son cerveau la représentation de la loi de la réfraction de Snell.

dant, non seulement ce programme audacieux est aujourd'hui pleinement reconnu, mais on l'*identifie* aussi complètement aux sciences cognitives elles-mêmes. Rares sont ceux dans ce domaine — sans parler du public en général — qui restent sensibles aux racines ou aux défis et interrogations actuelles des sciences cognitives. « Le cerveau traite l'information venue du monde extérieur », tel est l'énoncé convenu. Remettre en question cette affirmation semble étrange, et la conversation qui s'ensuit sera d'emblée considérée comme « philosophique ». Cette réaction révèle un point aveugle que le paradigme cognitiviste a introduit dans le sens commun actuel, et qui peut sérieusement fausser l'appréhension de perspectives plus larges ainsi que de l'avenir des STC.

L'enfant du cognitivisme : l'intelligence artificielle

Les manifestations du cognitivisme ne sont nulle part plus apparentes qu'en intelligence artificielle (ci-après IA), qui est la *projection littérale* de l'hypothèse cognitiviste. Pendant des années, plusieurs percées théoriques et technologiques ont été faites dans ce cadre de recherche : les systèmes experts, la robotique, le traitement de l'image, etc. Ces résultats ont été largement commentés sur la place publique

44

et nous n'allons pas en reprendre ici les débats¹⁶.

L'IA et ses fondations cognitivistes ont atteint leur apogée sociale au Japon dans le programme ICOT de la « cinquième génération ». Pour la première fois, un programme de recherche à l'échelle nationale a été mis en place pour faire converger les énergies de l'industrie, de l'État et des universités vers un même but : faire du Japon le leader de l'IA. Le noyau de ce programme — la « fusée qui doit atteindre la Lune » d'ici 1992 — est l'ordinateur dit de la cinquième génération : un dispositif cognitif capable de comprendre le langage humain et d'écrire ses propres programmes pour résoudre des problèmes soumis par des interlocuteurs inexpérimentés. Le projet ICOT s'inscrit incontestablement dans la perspective cognitiviste : en son centre, c'est en PROLOG — un langage de programmation de haut niveau basé sur la logique des prédicats — qu'on traite la résolution de problèmes et les interfaces entre schèmes de représentation. Le programme ICOT a suscité une réponse immédiate en Europe (le programme Esprit) et aux États-Unis. Il y a évidemment là un secteur compétitif d'importance pour la technologie de l'information (ci-après TI). Cependant, ce qui nous intéresse ici n'est pas de savoir si la fusée sera construite ou non, mais si elle pointe effectivement vers la Lune. Nous y reviendrons.

16. Voir par exemple *La Recherche en intelligence artificielle*, Paris, Editions du Seuil, coll. « Points Sciences », 1987.

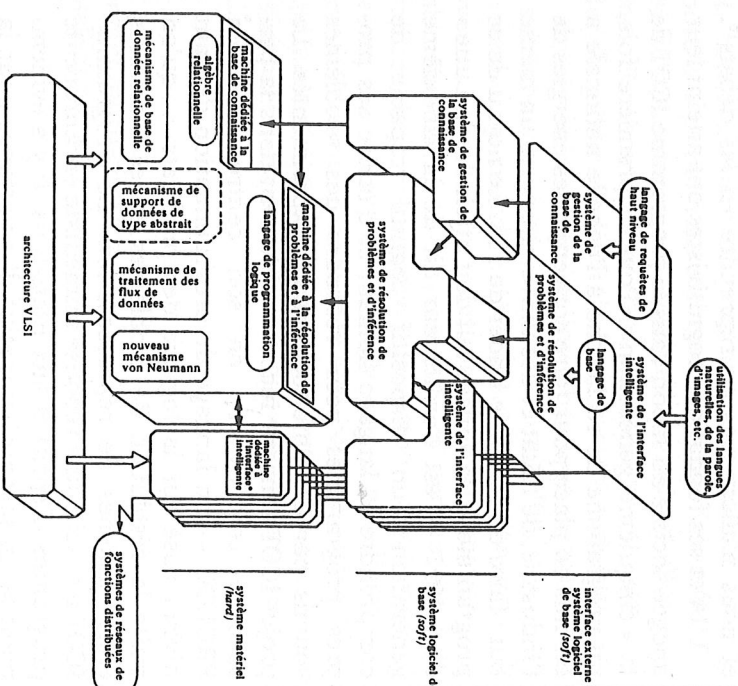


Figure 4 : Le schéma de base d'un système de la cinquième génération, le fer de lance de l'approche cognitive en STC du Japon. (Source : *Japan Information Processing Development Center*.)

Les sciences cognitives

Si l'hypothèse cognitive trouve en l'IA sa projection la plus littérale, sa quête complémentaire est l'étude de systèmes cognitifs biologiques et naturels, en particulier l'homme. Le principal véhicule d'exploration utilisé est, ici aussi, la formulation computationnelle de la représentation. Les représentations mentales sont assimilées aux éléments d'un système formel auxquelles la dynamique de l'esprit attribue le relief interprétatif : les croyances, les désirs, la planification, etc. Contrairement à l'objectif que poursuit l'IA, on s'intéresse ici à la véritable substance des systèmes cognitifs *naturels*, et l'on présume que leurs représentations ont une signification pour le système (c'est-à-dire qu'elles sont intentionnelles).

Par exemple, on présente des formes géométriques à des sujets auxquels on demande ensuite d'en visualiser mentalement la rotation. Il a été systématiquement remarqué que la difficulté de la tâche, évaluée en temps, dépend du nombre de degrés de liberté avec lesquels la forme doit pivoter — en deux ou trois dimensions. C'est-à-dire que tout se passe exactement comme si sous avions un « écran » mental sur lequel les formes tournent comme sur un écran de télévision¹⁷. Ces expériences finirent par aboutir à

17. R. Shepard et J. Metzler, *Science*, 171, 1971, p. 701-703.

une théorie explicite, postulant des règles de conditionnement de l'espace mental similaires à celles qui régissent l'affichage de données par un ordinateur. Les chercheurs impliqués suggérèrent alors que les opérations langagières et les opérations visuelles interagissent pour composer notre œil interne¹⁸. Cette approche a donné lieu à une abondante littérature, pour et contre, et des variantes furent proposées pour chaque niveau d'observation. Ainsi, les cognitivistes du noyau dur, comme Zenon Pylyshyn, ont avancé que les images sont simplement les épi-phénomènes subjectifs d'une computation symbolique plus profonde¹⁹. Quoi qu'il en soit, l'étude de la vision reste un bon exemple représentatif de l'approche cognitiviste dans le domaine de la phénoménologie mentale.

Le traitement de l'information dans le cerveau

L'influence du cognitivisme sur la manière actuelle d'envisager le cerveau est aussi très importante. Bien qu'en *principe* les idées du cognitivisme

soient compatibles avec un vaste éventail de théories sur le cerveau, en *pratique*, presque toute la neurobiologie (y compris son immense corpus de données empiriques) a été infiltrée par le corollaire cognitiviste du traitement de l'information. Voici la phrase d'introduction d'un manuel de neurosciences répandu : « Le cerveau est un ensemble de cellules constamment actives qui reçoit de l'information, la développe et la perçoit, et prend des décisions²⁰. » Le plus souvent, la justification et les implications de ce point de vue ne sont pas remises en question.

Ce point est particulièrement bien illustré par les deux décennies d'études sur le cortex visuel, une aire du cerveau où il est facile de détecter la réaction électrique des neurones lorsqu'une image est présentée au sujet. Il apparut très tôt que les neurones corticaux pouvaient être classés comme *détecteurs de traits* correspondant à certains attributs de l'objet présenté : l'orientation, le contraste, la vitesse, la couleur, ainsi de suite. En concordance avec l'hypothèse cognitiviste, ces résultats sont habituellement interprétés comme donnant une réalité biologique à l'idée que le cerveau reçoit l'information visuelle de la rétine par le biais des neurones détecteurs de traits du cortex visuel, l'information étant ensuite acheminée plus avant dans la chaîne du traitement

18. S. Kosslyn, *Psychol. Rev.*, 88, 1981, p. 46-66.

19. Voir *Beh. Brain, Sci.*, 2, 1979, p. 535-581, et H. Gardner, *The Mind's New Science*, op. cit.

20. Traduit de S. Kuffler et J. Nicholls, *From Neuron to Brain*, Boston, Sinauer Associates, 1976, p. 3.

cérébral (catégorisation conceptuelle, associations mnémoniques, et éventuellement action)²¹.

Cette approche est exprimée dans sa forme la plus radicale par la doctrine de Barlow de la « cellule grand-mère », qui fait correspondre des concepts (comme celui qu'on a de sa grand-mère) ou des percepts avec des neurones spécifiques²² — il s'agit là d'un équivalent IA des détecteurs et des *labeled lines* (« canaux étiquetés »). Cette position extrême est maintenant moins répandue²³, mais l'idée générale que le cerveau est un dispositif de traitement d'information, réagissant de façon sélective aux aspects discriminants de l'environnement, persiste au centre des neurosciences modernes et de l'idée que s'en fait le public. Nous reviendrons plus loin sur cette question.

L'ébauche de la dissension

Parmi les STC, le cognitivisme constitue un programme de recherche bien défini et complet, comportant des institutions prestigieuses, des revues spécialisées, une technologie appliquée et des préoccupations commerciales d'ordre international. Dans l'ensemble, la plupart de ceux qui travaillent en IA (et en technologie de l'information) adhèreraient — sciemment ou non — au cognitivisme. Après tout, comment pourrait-il en être autrement pour qui la poursuite de comportements bien définis — parmi les neurones, ou à l'aide de programmes de simulation LISP — est le pain quotidien ? Dans cet essai, nous voudrions attirer l'attention sur l'importance de cet *engagement social* d'un vaste secteur de la communauté de recherche en STC. Notre démarche est d'examiner les bases du cognitivisme au sein des STC, que les objets de dissension apparaissent aussi. Aujourd'hui, les principales divergences par rapport aux courants établis prennent les formes suivantes :

- la critique de la computation symbolique en tant que support approprié pour les représentations ;
- la critique de l'adéquation de la notion de représentation comme élément primitif des STC.

Ces deux positions seront analysées dans les deux prochains chapitres.

21. Le livre de P. Buser et M. Imbert, *Vision*, Paris, Hermann, 1986, illustre bien cette approche en France.

22. H. Barlow, « Single Units and Sensation : A Neuron Doctrine for Perceptual Psychology », *Perception*, 1, 1972, p. 371-394.

23. Voir, par exemple, la critique que fait David Marr de Barlow dans son ouvrage, *Vision : A Computational Investigation into the Human Representation and Processing of Visual Information*, New York, W.H. Freeman and Company, 1982.