

Véronique HAVELANGE\*

**ARTICLE CRITIQUE** : Jean-Pierre DUPUY, *Aux origines des sciences cognitives*. Paris, La Découverte, 1994.

Dans un récent ouvrage<sup>1</sup>, Jean-Pierre DUPUY consacre à la première phase des sciences cognitives une analyse tout à la fois historique et critique d'une remarquable acuité. C'est la phase dite de la "première" cybernétique, amorcée en 1943 par l'article de Rosenblueth, Wiener et Bigelow<sup>2</sup> et celui de McCulloch et Pitts<sup>3</sup>, puis matérialisée par les célèbres Conférences Macy qui se tinrent de 1946 à 1953, qui est au cœur de son analyse. Prenant pour matériau principal les actes de ces conférences<sup>4</sup>, Jean-Pierre Dupuy développe et amplifie ici le travail sur cette période<sup>5</sup> qu'il avait amorcé en 1985, et qui prenait place dans une vaste réflexion menée par des chercheurs du CREA sur la "première" et la "seconde" cybernétiques<sup>6</sup>.

L'élément qui frappe le plus dès les premières pages et qui fait l'originalité de cet ouvrage, c'est l'*esprit* dans lequel se place Dupuy par rapport à ce courant polymorphe, protéiforme qui a donné naissance aux

---

\* Je remercie François Sebbah, John Stewart, Bernard Stiegler et Yves-Marie Visetti pour leurs précieux commentaires et suggestions au cours de l'élaboration de ce travail.

<sup>1</sup> DUPUY J.-P., *Aux origines des sciences cognitives*, La Découverte, 1994.

<sup>2</sup> ROSENBLUETH A., WIENER N., BIGELOW J., "Behavior, Purpose and Teleology", in *Philosophy of Science*, vol. 10, n° 1.

<sup>3</sup> McCULLOCH W., PITTS W., "A Logical Calculus of the Ideas Immanent in Nervous Activity", repris dans *Embodiments of Mind*, MIT Press, 1965 ; rééd. 1979.

<sup>4</sup> Pour les références des documents utilisés, voir DUPUY J.-P., *op. cit.*, p. 14.

<sup>5</sup> "L'essor de la première cybernétique", in *Cahiers du CREA*, n° 7, novembre 1985, p. 7-139.

<sup>6</sup> Respectivement intitulés *Histoires de cybernétique* et *Généalogies de l'auto-organisation*, les numéros 7 et 8 des *Cahiers du CREA* parurent en 1985. Le numéro 9, intitulé *Cognition et complexité* et paru en mars 1986, complétait cette étude historique par une approche des débats contemporains suscités par les théories de l'auto-organisation dans les champs des sciences cognitives et de l'intelligence artificielle, de l'épistémologie et de la philosophie des sciences.

sciences cognitives. De la cybernétique<sup>7</sup> à l'intelligence artificielle aux théories de l'auto-organisation et de la complexité jusqu'à l'épistémologie expérimentale ou naturalisée et aux sciences cognitives, il entend en effet nous livrer une clef d'intelligibilité expressément fondée sur le postulat de la *dynamique autonome des idées*<sup>8</sup> : soucieux de ne pas tomber dans le piège de l'illusion rétrospective, Dupuy se défend de la démarche qui consisterait à proposer une reconstruction rationnelle de l'origine des sciences cognitives biaisée par l'actuel état de l'art. Mais il entend aussi se garder d'une approche qui privilégierait soit l'anecdotique (à la manière de Howard Gardner<sup>9</sup>), soit l'externalisme de la sociologie des sciences (comme le fait Steve Heims dans ses différents travaux<sup>10</sup>). Ce triple refus nous vaut une attention délibérée aux malentendus, aux contradictions, aux méprises interprétatives et aux potentialités inexploitées qui, d'après Dupuy, loin d'être des circonstances périphériques, constituent la trame, l'étoffe même de l'histoire de la cybernétique. C'est en ce sens que l'on peut dire que l'histoire qu'il nous propose n'est pas une "histoire standard".

Fort de ces résolutions, l'auteur fait en effet voler en éclats la "légende" des sciences cognitives<sup>11</sup>, brisant en amont et en aval les bornes de l'historiographie établie. En amont, tout d'abord, il montre que le véritable point de départ n'est pas l'invention de la "métaphore de l'ordinateur"<sup>12</sup>, mais la révolution logique des années 1930 : l'arithmétisation de la logique par Kurt Gödel en 1931<sup>13</sup>, suivie par l'assimilation de la notion de calculabilité effective, entendue comme procédure automatique, à la notion abstraite de machine. C'est en effet en 1936 que Alan Turing,

---

<sup>7</sup> L'auteur commente le fâcheux porte-à-faux de ce terme, forgé par Norbert Wiener en 1947, puisqu'il désigne étymologiquement les sciences de la gouverne, de la commande, alors que le propos de la nouvelle discipline fut — en ce qui concerne la "seconde cybernétique" en tout cas — de rendre compte de l'autonomie de certains systèmes à causalité circulaire (p. 69).

<sup>8</sup> *Op. cit.*, p. 11.

<sup>9</sup> GARDNER Howard, *The Mind's New Science*, Basic Books, New York, 1985 ; traduction française par J.-L. Peytavin, *Histoire de la révolution cognitive*, Payot, 1993.

<sup>10</sup> HEIMS Steve, "Encounter of Behavioral Sciences with new machine-organism Analogies in the 1940's", *Journal of the History of the Behavioral Sciences*, 11, 1975, p. 368-373 ; *John von Neumann and Norbert Wiener. From Mathematics to the Technologies of Life and Death*, MIT Press, Cambridge (Mass.), 1980 ; "An Encounter between neo-mechanists and the Human Sciences", *Cahiers du CREA*, n° 7, octobre 1985 ; *The Cybernetics Group*, MIT Press, 1991.

<sup>11</sup> *Op. cit.*, p. 44.

<sup>12</sup> *Op. cit.*, p. 21.

<sup>13</sup> *Id.*, p. 23.

proposant une formalisation de la notion abstraite de machine, pose la *thèse* selon laquelle toute fonction calculable mathématiquement est calculable par une "machine de Turing" et démontre que la classe des fonctions calculables par machine de Turing coïncide avec la classe des fonctions récursives<sup>14</sup>. Ainsi la notion de calculabilité effective se trouve-t-elle identifiée à celle de procédure mécanique : la machine de Turing est une calculatrice arithmétique ("*number cruncher*") et, puisque l'on sait depuis Gödel que la logique est arithmétisable, elle est également un calculateur symbolique ("*symbol cruncher*")<sup>15</sup>.

Cependant, de la même manière que Gödel, face au problème de la "décision" posé par Hilbert, avait abouti au théorème d'incomplétude de la logique arithmétisée, Turing arrive lui aussi à une réponse négative concernant le problème dit de "l'arrêt de la machine de Turing" : il n'existe pas de procédure mécanique susceptible de déterminer si le calcul effectué par la machine va s'arrêter ou boucler indéfiniment (p. 27). Tout comme les êtres mathématiques, que nous avons produits, maintiennent par rapport à nous une obscurité irréductible (p. 21), de même, ce que la machine est capable d'engendrer est au-delà du mécanique (p. 27).

Toujours en 1936, dans le même article, Turing démontre l'existence d'une machine de Turing singulière qu'il nomme "universelle" : elle est en effet capable d'imiter, de mimer, de reproduire, ou encore de simuler le comportement de n'importe quelle autre machine de Turing. Dans la mesure où le fonctionnalisme de la science de la cognition définit l'esprit comme le modèle de la faculté de modéliser (p. 22), la machine de Turing universelle est donc le modèle de l'esprit (p. 30).

Tels furent les prémisses sur lesquelles prit appui dans les années 1940 la cybernétique naissante. Ce n'est que bien plus tard, souligne Jean-Pierre Dupuy, que la révolution conceptuelle introduite par la machine de Turing servit de tremplin à une solution formaliste du problème des rapports entre la matière et la pensée et que des philosophes de l'esprit tels Fodor et Pylyshyn, sur base des travaux de Newell et Simon en intelligence artificielle, promurent le "fonctionnalisme computo-représentationnel", qui consiste à prendre au pied de la lettre la métaphore du cerveau comme ordinateur, cette machine de Turing incarnée dans des circuits électroniques (p. 28).

Ayant fait cette mise au point, Jean-Pierre Dupuy montre dans quelle mesure la vulgate des sciences cognitives se fourvoie quand elle met en scène seulement deux protagonistes, l'homme et la machine, ou encore, à

---

<sup>14</sup> *Id.*, p. 25.

<sup>15</sup> Ce point avait déjà été noté entre autres par Daniel ANDLER dans "Le cognitivisme orthodoxe en question", in *Cahiers du CREA*, n° 9, mars 1986, p. 7-106.

la manière des analogies lâches de Wiener, le cerveau et la machine. Il y eut au contraire trois protagonistes : l'organisme dans sa structure (le cerveau), l'organisme dans sa fonction (l'esprit), et la machine, jouant le rôle du modèle et se différenciant elle-même en machine logique (machine de Turing, machine de McCulloch) et en machine artificielle, matérielle (l'ordinateur). L'auteur repère alors trois moments dans la mise en place de la cybernétique :

1) L'assimilation de l'esprit à une machine, logique d'abord. C'est là le fruit de la révolution logique des années 1930, à laquelle il faut ajouter les travaux de Claude Shannon qui élabore en 1938 une théorie logique des circuits électriques : "machine artificielle ou matérielle, d'un côté, logique comme machine, de l'autre, sont liées par un rapport d'identité", commente Jean-Pierre Dupuy (p. 46).

2) Vient alors l'article fondateur de McCulloch et Pitts, publié en 1943, "A Logical Calculus of the Ideas Immanent in Nervous Activity"<sup>16</sup>, qui assimile le cerveau dans sa structure, c'est-à-dire le cerveau naturel, biologique, à une machine de Turing. Dans l'esprit de McCulloch, ce travail résout le problème du rapport entre le corps et l'âme — soit, dans ses propres termes, le problème de l'"*embodiment of mind*" (incarnation de l'esprit) : en effet, écrit Jean-Pierre Dupuy, il démontre en principe l'existence d'une machine logique équivalente à celle de Turing, qui peut être considérée dans sa structure et dans son comportement comme une idéalisation de l'anatomie et de la physiologie du cerveau. "En d'autres termes, ceux du calcul des propositions, chaque neurone idéalisé est un calculateur mathématique élémentaire, qui compute une fonction booléenne à seuil" (p. 48).

3) Ce n'est qu'alors (et non dans le premier temps) qu'entre en scène l'ordinateur, en l'espèce de la "machine de Von Neumann", présentée par ce dernier au symposium Hixon qui se tint en 1948<sup>17</sup> : comme le note Jean-Pierre Dupuy, Von Neumann jette les bases de l'idée constitutive de l'ordinateur moderne, à savoir la mise à plat, au même niveau, des instructions (le "programme") et des données. Cette idée était déjà préfigurée par la machine de Turing, mais il aura fallu, pour "remonter", comme le fit Von Neumann, de la machine matérielle à la machine de Turing, un intermédiaire — et celui-ci fut la machine de McCulloch (p. 62).

---

<sup>16</sup> Voir McCULLOCH W., PITTS W., *art. cit.*, 1943. Voir ci-dessus, note 3.

<sup>17</sup> VON NEUMANN J., "The General and Logical Theory of Automata", in JEFFRESS A. (éd.), *Cerebral Mechanisms in Behavior. The Hixon Symposium* (California Institute of Technology, sept. 1948), John Wiley and Sons, New York, 1951 ; p. 22-24.

Ce redressement de l'anamnèse de la cybernétique est pour Jean-Pierre-Dupuy l'occasion de réhabiliter la figure méconnue de Warren McCulloch, qui a eu le mérite, d'un point de vue logique, de prendre au sérieux le caractère de "tout ou rien" de la mise à feu neuronale et a, de ce fait, proposé un modèle de type logique (p. 51), et qui, sur un plan neurologique, a inauguré la démarche fonctionnaliste en tant que reposant sur l'idée de réalisabilité multiple du modèle (p. 54). Jean-Pierre Dupuy souligne l'injustice d'une histoire qui, après les travaux proprement cybernétiques sur le Perceptron (Rosenblatt, 1957-1961), élimine très tôt toute référence à McCulloch dans le connexionnisme contemporain, alors que ce dernier est en fait son héritier direct (p. 61).

Ces diverses rectifications autorisent Jean-Pierre Dupuy à affirmer que, contrairement à l'idée reçue, la cybernétique n'était pas une "science nouvelle" (*scienza nuova*) en rupture avec la physique, qu'elle était au contraire intrinsèquement physicaliste et que son objectif, visant à mécaniser l'humain — et non à humaniser la machine —, relevait d'une entreprise d'extension du physicalisme au domaine de l'esprit. De ce fait, la cybernétique était encore foncièrement behavioriste et, par conséquent, éliminationniste.

C'est ici qu'intervient le second geste de Jean-Pierre Dupuy dans sa démythification de la légende des sciences cognitives. Après avoir fait reculer le jalon fondateur aux innovations logiques des années 1930, il repousse l'étape de l'accomplissement de la cybernétique vers les surgeons particuliers de celle-ci que sont les théories de l'auto-organisation et de la complexité. Cette "seconde cybernétique", méprisée et à peine considérée par les plus grands, tel Herbert Simon, comme une "branche morte" de la cybernétique, réussit en effet, selon Jean-Pierre Dupuy, à opérer le "*Gestalt switch*", le changement global de perspective nécessaire pour rendre compte des processus à causalité circulaire — tâche que s'était en principe assignée la première cybernétique, et que celle-ci réalisa dans la mesure où elle introduisit la notion de *feedback* (boucle de rétroaction), mais qu'elle ne parvint pas à accomplir pleinement parce qu'elle resta prisonnière d'une perspective behavioriste, c'est-à-dire d'un point de vue considérant les réseaux neuronaux sous l'angle de systèmes à entrées et sorties au service d'une tâche prédéfinie. À l'inverse, la seconde cybernétique allait appréhender les réseaux comme des entités autonomes, capables de spontanéité et de "comportements propres" dont l'étude est, en soi, intéressante. Jean-Pierre Dupuy salue ici toute la thématique de l'auto-organisation et de la complexité comme l'accomplissement effectif du projet cybernétique, et relate non sans verve

les discussions enflammées qui eurent lieu à ce propos aux conférences Macy<sup>18</sup>.

Reste donc à sa charge d'expliquer pourquoi et comment les sciences cognitives orthodoxes ont suivi une autre voie que celle de la seconde cybernétique. Nous arrivons ainsi à la thèse majeure de l'ouvrage. Jean-Pierre Dupuy soutient que l'histoire de la cybernétique se solde par un échec, dans la mesure où elle a manqué tous les rendez-vous importants qui se sont offerts à elle : biologie, psychologie, phénoménologie, physique des systèmes désordonnés, thématique de la complexité et, enfin, sciences de l'homme et du social (p. 157-158). Je me pencherai en particulier sur le cas de la "rencontre ratée" avec la phénoménologie et avec les sciences de l'homme. Tout se passe, selon Dupuy, comme si la cybernétique avait échoué à établir ces articulations, et donc à réaliser effectivement son projet interdisciplinaire, "alors même qu'elle avait les moyens *formels* <sup>19</sup> de le faire". La rencontre avec la phénoménologie échoue en raison du contresens commis par Chisholm puis Quine dans la lecture de Brentano et Husserl : l'intentionnalité, loin d'être comprise comme l'*activité* psychique se dépassant elle-même à l'intérieur d'elle-même en direction d'un objet qui lui reste intérieur (le noème), soit comme la "transcendance dans l'immanence" que tente de cerner Husserl, devient, sous la plume de ces philosophes analytiques, un *état* mental doté d'un contenu qui se rapporte à un objet physique dont l'existence n'est pas garantie par le fait que l'état mental, lui, existe ; ce contenu ne peut être qu'intentionnel, donc linguistique (p. 102-106). Cette lecture va donc directement à l'encontre de la démarche phénoménologique par excellence, à savoir "l'épochê". De ce fait, pour Dupuy, la cybernétique ne marque pas le "*cognitive turn*" que d'aucuns pensent voir en elle : bien au contraire, elle reste — et les sciences cognitives resteront avec elles — tributaires d'une philosophie du *langage* (p. 94-101) plutôt que de l'esprit, et la tentative de "naturaliser" cette philosophie psycholinguistique engendrera le dilemme permanent d'une théorie matérialiste de l'esprit oscillant sans fin entre une position "éliminationniste" et une position "non réductionniste" (p. 101). Ironie pour Dupuy, car, paradoxalement, il eût été selon lui plus aisé de mener à bien la naturalisation d'une philosophie transcendantale que celle d'une philosophie psycho-linguistique de l'esprit (p. 101).

---

<sup>18</sup> Notons au passage à quel point le motif formel reste ici prédominant, y compris et jusque dans la reconnaissance de l'incomplétude du formalisme, qui marque sa capacité, hautement louée par Jean-Pierre Dupuy, à formuler ses limites et à affirmer sa réflexivité (thème d'une "cybernétique du second ordre", chez Von Foerster).

<sup>19</sup> *Op. cit.*, p. 171. C'est moi qui souligne.

C'est également la question d'une prise en compte de la "transcendance immanente" (pour reprendre l'expression de Husserl) qui signe, selon Dupuy, l'échec de la rencontre de la cybernétique avec les sciences de l'homme : si McCulloch, à l'inverse d'un Wiener, était parfaitement conscient de la contribution de ses travaux à l'élimination du sujet, ce n'est pourtant qu'avec les théories de l'auto-organisation et de la complexité — qui renient toute filiation par rapport à la cybernétique — que seront pensables des *processus sans sujet*. Or, ceux-ci avaient déjà été thématiques par les sciences sociales, en particulier par la tradition libérale et économique du XVIIIème siècle (p. 171 ; "main invisible", "ruse de la Raison", notion reprise sous la forme "d'ordres sociaux spontanés" par Hayek) et par le "nietzschéo-heideggérianisme" qui, en France spécialement, informa les sciences de l'homme de l'après-guerre dans leur réaction contre les philosophies de la conscience et donna carrière au structuralisme, au post-structuralisme et à la "déconstruction" (p. 173). La rencontre de la cybernétique avec ces courants *thématisés comme tels* en sciences humaines et sociales aurait "pu et dû se faire", note Dupuy (p. 171), et autrement que sur le mode de "l'énorme malentendu" (*id.*) qui sévit entre les promoteurs des conférences Macy — qui se voulaient humanistes — et leurs acteurs scientifiques, lesquels, selon Dupuy, étaient sans le savoir en affinité avec l'entreprise de déconstruction de la conception métaphysique du sujet (p. 173)). Cette affinité, cette parenté des sciences sociales avec le projet intrinsèque de la cybernétique, ce n'est pas celle-ci qui les discerna, mais les sciences de l'homme (Lacan, p. 49 et 113, 173, et Lévi-Strauss, p. 112 et 173) et, plus tard, les courants qui succédèrent à la cybernétique : citons en autres la théorie des jeux de Von Neumann et Morgenstern, dès 1943 (p. 55-56), le principe "d'ordre par le bruit" de Von Foerster en 1960, et le principe de la "complexité par le bruit" d'Atlan en 1979 (p. 128).

\*

Déception, déconvenue donc pour Jean-Pierre Dupuy à l'issue de ce bilan de l'histoire de la cybernétique. Mais on peut ici se demander si cette frustration exprimée tout au long de l'ouvrage n'est pas, d'une manière qu'il conviendra de préciser, l'ombre portée d'une ambiguïté inhérente à la pensée de l'auteur, et qui concernerait non pas le passé *comme tel* (finalement, si la communauté scientifique arrive aux concepts d'autonomie, de complexité, de processus sans sujet..., est-il si important que cela se fasse sans la cybernétique, voire contre elle? (p. 144)), mais la position *actuelle* de l'auteur face aux sciences cognitives et à leur articulation avec les sciences du "social". Deux points me paraissent ici

mériter commentaire : les prises de position de Jean-Pierre Dupuy concernant le *formalisme* et la *déconstruction du sujet*.

On l'a compris, l'auteur prend en effet des positions très fermes et très claires en faveur du formalisme. Il identifie ainsi, dans son premier chapitre, consacré à la modélisation, modèle et structure formelle, et on note l'approbation avec laquelle il parle, au chapitre 6, du théorème d'Ashby (dont il remarque la fascination pour Bourbaki). Or, dans le travail scientifique vivant, la modélisation ne se réduit pas à un traitement mathématique abstrait : des intuitions qualitatives et une dimension processuelle y interviennent essentiellement. Comme le montre de manière éloquente Yves-Marie Visetti, la "dimension figurale" et "l'impureté métaphorique" qui la caractérisent confèrent à la modélisation un caractère "d'indiscipline épistémologique"<sup>20</sup>. Notre pratique, même scientifique, n'est pas fondamentalement formelle. De surcroît, il y a lieu de contester le bien-fondé d'une conception de la modélisation qui fait totalement abstraction de la nature, de l'identité et du nombre des éléments concrets dont les relations systémiques sont censées être abstraites par le modèle (p. 18). Jean-Pierre Dupuy procède ici à une assimilation graduelle entre connaissance, formalisation, modélisation et simulation. C'est dans la mesure où il définit la connaissance comme la capacité même de modéliser, et où il en infère que ce qu'il considère comme le geste scientifique par excellence s'applique à toutes les formes de cognition — y compris élémentaires (dans l'ordre du vivant) et ordinaires (dans le contexte de la vie quotidienne des agents humains) —, que l'objectif des sciences cognitives se voit dès lors défini, comme nous l'avons noté plus haut, en tant que "modélisation de la faculté de modéliser".

Il convient ici de rappeler les bases d'une position alternative au formalisme. Pour remettre celui-ci à sa place, pour montrer en quoi il n'est pas à prendre comme fondement absolu, il suffit de considérer sa genèse : le formalisme est certes issu des mathématiques, mais de mathématiques où les symboles étaient au départ pourvus de sens ; le geste fondateur du formalisme abouti, tel qu'il a été parachevé par Hilbert et Bourbaki, consiste à épurer ces symboles de leur signification. Ce que l'on pourrait appeler "l'histoire naturelle" du formalisme montre donc que les symboles, à l'origine, ne viennent à être que parce qu'ils font sens. Les symboles *formels*, quant à eux, sont dérivés et constituent une forme particulière de symbolisme qui peut, certes, être utile dans certains

---

<sup>20</sup> Voir notamment VISETTI Y.-M., "Notes sur la question "Comment naturaliser une phénoménologie" envisagée sous le double point de vue de la modélisation et de la provocation husserlienne", texte inédit, p. 2-3.

contextes particuliers ; cependant, les placer au fondement d'une tentative d'élucidation scientifique de la cognition ne peut que mener à une impasse.

À la question de savoir en quoi un système formel peut être cognitif, il existe en effet deux réponses possibles. La première consiste à faire fond sur le *computationnalisme*, c'est-à-dire sur la logique devenue écriture mécanique du calcul : le computationnalisme consiste à voir dans le calcul une manière d'opérer non seulement sur les nombres (*number crunching*), mais aussi sur les symboles (*symbol crunching*). Telle est la voie prise par la cybernétique. McCulloch et Pitts écrivent ainsi que "toute idée est incarnable dans un réseau de neurones" (p. 158). Ils forgent à cet effet le concept de *psychon*, tout petit réseau composé de deux neurones en entrée connectés à un neurone de sortie et permettant de réaliser (c'est-à-dire d'incarner dans la machine) les opérations logiques de base. Cette voie computationnaliste est donc celle qu'adopte McCulloch et qui conduira vers le cognitivisme fonctionnaliste de Fodor et Pylyshyn (p. 54), en conformité avec toute la tradition de la philosophie analytique. Comme le note Jean-Pierre Dupuy, ce n'est pas par hasard que, dans ses travaux tardifs, orientés vers la logique, McCulloch s'est entouré de ceux qui allaient fonder l'intelligence artificielle, en particulier Newell et Simon.

Point intéressant cependant, Jean-Pierre Dupuy se démarque du computationnalisme en faisant siennes toutes les critiques (qui commencent actuellement à être bien rodées) portant sur le problème de l'ancrage des symboles (*symbol grounding problem*) (p. 29). L'auteur préfère explorer une autre voie, à savoir la propriété qu'exhibent certains systèmes formels de posséder une dynamique en termes d'attracteur et de point fixe endogène. Dupuy découvre dans ces systèmes dynamiques une propriété de "transcendance immanente" pouvant, d'après lui, être rattachée à Husserl plutôt qu'à la philosophie analytique. Il semblerait donc que Jean-Pierre Dupuy voie en Husserl le précurseur de sciences cognitives formalistes mais non computationnalistes, c'est-à-dire n'interprétant pas nécessairement les opérations de calcul comme des opérateurs logiques ; Jean-Pierre Dupuy évoque effectivement ici, entre autres, les travaux de Jean Petitot, qui, comme on sait, se réclame expressément de la phénoménologie transcendantale<sup>21</sup> et resitue dans cette perspective les concepts de morphogenèse et de catastrophe.

---

<sup>21</sup> Voir notamment PETITOT J., *La philosophie transcendantale et le problème de l'objectivité*, Entretiens du Centre Sèvres (père F. Marty éditeur), Editions Osiris, Paris, 1991. Cité par J.-P. DUPUY, *op. cit.*, p. 128.

Cependant, la question posée tout à l'heure rebondit et devient à présent : en quoi un point fixe dans un système dynamique est-il cognitif ? — ou, en termes légèrement différents, comment peut-on penser la cognition avec ces outils mathématiques ? En fait, le point fixe peut avoir une valeur cognitive s'il est interprété dans le cadre d'un organisme autopoïétique. Remarquons-le toutefois, le concept d'autopoïèse n'est pas lui-même formel : comme en témoigne l'entretien d'Isabelle Stengers avec Francisco Varela<sup>22</sup>, Maturana et Varela définissent comme autopoïétique une entité qui se produit en permanence matériellement. On s'achemine donc ici vers une biologie qui, loin d'être découplée des propriétés de la matière, s'enracine au contraire dans leurs particularités pour comprendre les processus de morphogenèse<sup>23</sup>. On dissout ainsi le vitalisme (dans la mesure où l'on ne recourt pas à un principe vital abstrait), mais on ne verse pas pour autant dans le fonctionnalisme, puisque l'on prend en compte les singularités de l'organisation biochimique de la matière. On est donc ainsi très proche de la physique — mais une pause est ici de mise : proche de *quelle* physique ? Le lecteur ne peut qu'être déconcerté par les propos de Jean-Pierre Dupuy quand celui-ci fustige (à juste titre) les partisans cognitivistes, computationnalistes d'une Dame physique rationalisée, conceptualisée, abstraite, pour (prétendument) se réclamer de la physique "vivante, concrète et plurielle des laboratoires" (p. 169)... alors même qu'il s'érige lui-même en champion du formalisme.

Dans la perspective de l'autopoïèse, ce n'est pas le point fixe en tant que tel (c'est-à-dire en tant qu'entité formelle) qui compte, c'est le point fixe en tant qu'il est *interprété* par des scientifiques comme faisant partie des mécanismes de production matérielle de soi. En fait, il n'y a pas de réponse à la question de savoir en quoi un point fixe peut être cognitif. Ou, pour le dire de manière plus précise, la pseudo-réponse qu'apporte Jean-Pierre Dupuy à cette question consiste à *ontologiser*<sup>24</sup> implicitement le formalisme, et ce par le truchement de l'assimilation, posée dès le premier chapitre de l'ouvrage, entre le sens scientifique (abstraction des relations systémiques entre éléments) et le sens ordinaire (simulacre, feinte, imitation fallacieuse) des notions de modèle et de simulation (p. 15-33). Les systèmes dynamiques à points fixes endogènes ou bassins d'attracteurs seraient ainsi "naturellement" ou "spontanément" cognitifs,

---

<sup>22</sup> Voir *Cahiers du CREA*, n° 8, novembre 1985, p. 283-297.

<sup>23</sup> Voir SAUNDERS, "L'évolution des formes biologiques", in STEWART J. (éd.), *Biologie et cognition, Intellectica*, n° 16, p. 61-83.

<sup>24</sup> Ce qu'éclaire de manière particulièrement suggestive son adhésion à la démarche de McCulloch, qui confère une "réalité ontologique" au modèle (p. 43).

sans qu'il soit besoin d'assumer, de thématiser autrement ce problème<sup>25</sup>. Cette ontologisation du formalisme débouche directement sur une confusion délibérée — d'ailleurs préparée d'entrée de jeu (chapitre 1) — entre la notion de réel et celle de modèle, et sur leur appréhension comme interchangeables. Cette identification culmine à son tour avec les théories de la complexité, qui permettent "l'émancipation du modèle" : "Les modèles de la cybernétique sont déjà poststructuralistes, ils ne sont modèles que d'eux-mêmes, ou bien d'autres modèles, miroirs de miroirs, spéculums ne réfléchissant aucune réalité"<sup>26</sup>. Et plus loin : "Le modèle, qui était hiérarchiquement subordonné au réel qu'il ne faisait que mimer, s'émancipe et devient l'égal de son référent. (...) Mais alors, il sera non seulement modèle de son objet, mais aussi modèle de lui-même, ou plutôt de son comportement"<sup>27</sup>. Jean-Pierre Dupuy va même jusqu'à se réclamer de Vico<sup>28</sup>. Il rejoint en tout cas tacitement les physiciens qui affirment que "l'univers compute"<sup>29</sup> ou les philosophes qui évoquent une "machine univers"<sup>30</sup>. Quoiqu'il s'abstienne — par sobriété ou par élégance — d'aller

---

<sup>25</sup> Soulignons au passage le clivage conceptuel que Jean-Pierre Dupuy établit constamment entre, d'une part, le logique, identifié au naturel, et, d'autre part, l'artificiel, assimilé au matériel.

<sup>26</sup> *Aux origines des sciences cognitives*, *op. cit.*, p. 152. Et l'auteur d'évoquer ici Jacques Derrida, l'un des maîtres de la déconstruction. Nous reviendrons sur ce point.

<sup>27</sup> *Id.*, p. 156.

<sup>28</sup> Auteur notamment d'un ouvrage intitulé *La scienza nuova* (rééd. par P. Rossi, Rizzoli Editore, Milan, 1963), Giambattista Vico (1661-1744) s'employa avec force à réfuter les doctrines de Gassendi et de Descartes et à combattre l'idée d'une mathématique universelle, c'est-à-dire la prétention de tout ramener aux mathématiques, considérées comme la science parfaite sur laquelle les autres devraient prendre modèle. Il développa dans cette optique les fondements d'une science humaine autonome (une "*scienza nuova*") basée sur le primat de l'action humaine concrète dans l'histoire et sur une réhabilitation du statut cognitif de l'imaginaire. Certes — et c'est sur ce point que Jean-Pierre Dupuy fait fond —, cette démarche le conduisit à reconnaître dans les mathématiques elles-mêmes un produit de l'action humaine. Mais on ne peut s'empêcher de songer à une distorsion de la démarche de Vico — lequel affirmait "*Verum ipsum factum*" ("le vrai est cela même qui est fabriqué" ; je traduis), et ce dans le sens de l'action humaine, concrète dans l'histoire — quand on voit Jean-Pierre Dupuy interpréter la même sentence (*op. cit.*, p. 16) dans un sens où la production *de modèles formels* serait l'activité scientifique par excellence : "Le modèle est comme une forme abstraite qui vient à s'incarner dans les phénomènes (...). Si (...) le scientifique est un *Homo faber*, c'est avant tout parce qu'il est concepteur et fabricant de modèles" (*op. cit.*, p. 18). Voir POMPA L., *Vico. A Study of the "New Science"*, Cambridge, Cambridge University Press, 1990.

<sup>29</sup> Communication personnelle de Francis Bailly.

<sup>30</sup> Voir LEVY P., *La Machine univers*, La Découverte, 1990.

jusque là, Jean-Pierre Dupuy légitime de la sorte les diverses tentatives qui visent à dégager des isomorphismes entre domaines et pensent ainsi faire oeuvre d'explication scientifique<sup>31</sup>. En tout état de cause, Jean-Pierre Dupuy s'emploie donc à ne pas interpréter le formalisme, ou à l'interpréter subrepticement en mêlant son acception scientifique et son acception ordinaire. Cette prise de position lui sied pleinement, car elle prépare à la perfection son entreprise de déconstruction du sujet.

Car c'est bien là que Jean-Pierre Dupuy veut en venir : selon lui, nous l'avons vu, "l'intérêt de la cybernétique, c'est qu'elle avait les moyens formels de penser les processus sans sujet" (p. 171). Dès le premier chapitre, il en a formulé l'annonce prémonitoire : "La révolution copernicienne introduite par la "science de l'esprit" aura réalisé une *déconstruction* de la métaphysique de la subjectivité qui va beaucoup plus loin que la démarche philosophique qui s'est donné ce nom" (p. 33). Et Jean-Pierre Dupuy de faire allégeance à Jacques Derrida, maître incontesté de ce courant de pensée. C'est dans la même veine qu'il commente une citation du philosophe de l'esprit Daniel Dennett : "Ne croirait-on pas lire du Derrida ?"<sup>32</sup>. Jean-Pierre Dupuy est extrêmement clair sur ce point : la déconstruction du sujet est l'objectif commun de la cybernétique, de la philosophie et des sciences sociales. Son argumentation tient en quatre points que l'on peut expliciter de la manière suivante :

1) Il existe des entités qui ne possèdent pas les attributs de sujet. La "règle d'or de l'individualisme méthodologique" alléguée à ce propos par Dupuy ne dit en effet rien d'autre : il est des entités, soit supra-humaines (sociétés, organisations...), soit infra-humaines (gènes et enzymes par excellence, à propos desquels Dupuy pourfend l'attribution de "surpuissance subjective" que leur fait la biologie moléculaire), auxquelles on ne peut prêter de qualités subjectives sauf à violer la règle de l'individualisme méthodologique, c'est-à-dire "faire une faute sur les catégories" (p. 174).

2) Néanmoins, il est difficile de ne pas prêter à ces entités les qualités de sujet. "Cette faute, nous sommes constamment tentés de la commettre" (p. 174). Et plus loin, Dupuy radicalise cette affirmation en alléguant à nouveau les positions de Dennett : "Comme l'explique Dennett, c'est inévitable" (p. 175).

---

<sup>31</sup> Voir, à titre d'exemple, VIROLE B., *Sciences cognitives et psychanalyse*, Nancy, Presses Universitaires de Nancy, 1995.

<sup>32</sup> *Aux origines des sciences ccognitives, op. cit.*, p. 176. L'auteur se réfère ici au livre de D. DENNETT, *The Intentional Stance*, Cambridge, MIT Press, 1987 ; trad. fr. par P. Engel, *La stratégie de l'interprète*, Gallimard, 1990.

3) L'auteur aboutit ainsi à l'idée qu'il peut y avoir apparence de subjectivité là où il n'y a pas de sujet.

4) On peut dès lors affirmer que les êtres humains ont l'apparence de sujets, alors même qu'ils ne sont pas des sujets. CQFD.

Cette argumentation ne peut être admise comme telle, en raison des glissements subreptices et des assimilations indues qui la sous-tendent. Analysons tout d'abord, d'un point de vue biologique ou, plus largement, scientifique, les glissements qu'elle opère. Certes, il existe des entités infra-humaines et supra-humaines possédant certains attributs relatifs à la subjectivité, mais elles ne les exhibent pas tous à la fois : l'argument de Jean-Pierre Dupuy ne paraît crédible que parce qu'il ne fait aucune différence entre les attributs de la subjectivité et les appréhende *en bloc*, sur le mode du tout ou rien. Ainsi, par exemple, il existe des entités infra-humaines douées d'une dimension cognitive : les cellules (par contraste avec les gènes et les enzymes), régies par une clôture opérationnelle, sont de ce fait susceptibles d'avoir une histoire et de développer une subjectivité ; mais elles ne sont pas des sujets doués d'une présence à soi consciente. De même, il existe bien des entités supra-humaines, telles les structures sociales, que l'on peut en droit constituer en objets scientifiques pourvus de certaines qualités subjectives ; elles peuvent, notamment, dépérir. Mais cela n'oblige pas pour autant à leur attribuer *toutes* les propriétés subjectives des êtres humains (volonté, responsabilité, etc.). Comme le montre la phylogenèse, la constitution de la pleine subjectivité est un processus qui consiste en de nombreuses étapes différenciées. Contrairement à ce qu'affirme Dennett, il est donc possible de considérer la capacité d'élaborer des intentions — c'est-à-dire d'exécuter une suite d'actions en vue de réaliser un but prédéfini — comme une qualité cognitive parfaitement identifiable scientifiquement sur la base d'un substrat physicochimique déterminé<sup>33</sup>.

Venons-en à présent aux assimilations contestables opérées par l'argumentation de Jean-Pierre Dupuy. Celui-ci, on l'a vu, identifie ce qu'il appelle la "déconstruction philosophique du sujet" avec l'éliminationnisme de philosophes de l'esprit tels que Dennett. Nous nous devons ici d'examiner d'un point de vue critique la référence de l'auteur à Jacques

---

<sup>33</sup> Jean-Pierre Dupuy souligne à juste titre (p. 41) que, quand les cybernéticiens, mus par leur physicalisme militant, rejettent la distinction entre comportements réflexes et comportements volontaires, ils se privent d'une différenciation cruciale pour l'étude de la cognition et de l'esprit. Mais Jean-Pierre Dupuy ne fait rien d'autre que les cybernéticiens dans les phases décisives de son argumentation.

Derrida. En effet, il importe de rappeler ici que Derrida, déployant une problématique de la "déconstruction", n'a jamais voulu poser par là un geste destructeur ou — dans le langage des sciences cognitives contemporaines — éliminationniste. La déconstruction telle qu'il l'entend *n'annule pas, ne congédie pas* ce à quoi elle s'applique ; encore moins vise-t-elle à le *dominer* en le "démontant". *Déconstruire*, sous la plume de Jacques Derrida, signifie au contraire "inventer", "se laisser surprendre", ou encore "déverrouiller" la subjectivité afin d'en libérer l'efficace<sup>34</sup>. En outre, la déconstruction derridienne ne porte pas seulement sur la conception métaphysique de la subjectivité, mais sur l'activité de *lecture*, et ce dans la mesure même où la *différance* (à savoir la différence ontologique, comme écart et comme retardement) est reconnue à l'oeuvre dans la parole, l'action et le discours. A l'instar de la déconstruction, la *dissémination* est donc pour Derrida une stratégie pratique d'interprétation visant à garantir que l'épaisseur du texte, comme celle du sujet, ne soit pas sous-estimée<sup>35</sup>. On peut dire en ce sens que Derrida *radicalise l'herméneutique* : déconstruire, c'est faire proliférer les interprétations, qui constituent elles-mêmes le sujet dans sa compréhension de lui-même. Or, nous l'avons vu, Jean-Pierre Dupuy, quant à lui, *radicalise le formalisme* : là où Derrida pose que le sujet n'existe pas avant les interprétations qu'il effectue ni indépendamment d'elles, Dupuy — à la suite de Dennett — voit des "processus sans sujet". La perspective, on l'aura remarqué, est totalement différente. Ce que Derrida critique, c'est l'idée d'un sujet métaphysique, transcendantal, c'est-à-dire *a priori*. À l'opposé, Jean-Pierre Dupuy soutient que la cybernétique, ainsi que l'intelligence artificielle et les sciences cognitives, perpétuent le transcendantalisme kantien, en "inventant un type d'enquête transcendantale qui (fait) l'économie du sujet"<sup>36</sup>. "De l'esprit (au sens de *mind* ), oui, du sujet, non. Un esprit sans sujet, voilà peut-être l'apport le plus significatif de la cybernétique à la philosophie"<sup>37</sup>. Pour

---

<sup>34</sup> Voir DERRIDA J., "L'invention de l'autre", in *Psyché*, Galilée, 1987, p. 11-63.

<sup>35</sup> Voir HOY D., "Jacques Derrida", in SKINNER Q. (éd.), *The Return of Grand Theory in the Human Sciences*, Cambridge University Press, 1986, p. 41-64 (p. 64).

<sup>36</sup> DUPUY J.-P., *Aux origines des sciences cognitives*, *op. cit.*, p. 112. Notons que le structuralisme a été, en son temps, qualifié de kantisme sans sujet transcendantal.

<sup>37</sup> DUPUY J.-P., *op. cit.*, p. 112. Jean-Pierre Dupuy reprend à son compte l'analyse de Joëlle PROUST (dans "L'intelligence artificielle comme philosophie", *Le Débat*, 1987, n° 47, p. 88-102) selon laquelle l'intelligence artificielle serait une philosophie de type transcendantal où le "système symbolique physique" de Newell et la machine de Turing tiendraient respectivement la place du sujet transcendantal et de la synthèse *a priori*, perpétuant ainsi la distinction kantienne entre questions de droit (la genèse transcendantale de la connaissance) et questions de fait (la genèse factuelle, les

Jean-Pierre Dupuy en effet, comme nous l'avons vu, le formalisme est "naturellement", immédiatement cognitif, sans que cette dimension d'intentionnalité requière une quelconque thématization. C'est dans cette exacte mesure que Dupuy peut s'identifier à la "posture intentionnelle" d'un Dennett, pour lequel le constat d'une dimension intentionnelle n'est jamais que l'affaire d'un "comme si" anthropomorphique et non scientifique de la part de l'observateur. Derrida, quant à lui, dirige sa critique contre la postulation du transcendantal en tant que tel : pour lui, la radicalisation de l'interprétation herméneutique permet de saisir la subjectivité dans son jaillissement. A l'inverse, Jean-Pierre Dupuy ambitionne de réduire le sujet à une simple apparence, et donc de l'éliminer au profit d'un '*comportement sans sujet*'<sup>38</sup>. En fait, entre la projection d'intentionnalité par un observateur — à la manière de Dennett et Dupuy — et l'interprétation herméneutique, qui transforme indissociablement ce sur quoi elle porte et celui qui la pratique, il y a un abîme<sup>39</sup>.

Nous l'avons dit, l'ouvrage de Jean-Pierre Dupuy dégage une forte charge affective dans laquelle, étrangement, se mêlent gratitude et déception, sympathie et tristesse à l'égard de la cybernétique. Mais un tel constat d'échec<sup>40</sup> est-il en fait justifié? Après avoir identifié dans la progression du formalisme et l'élimination du sujet les ambitions de l'aventure cybernétique telle que la voit l'auteur, on s'aperçoit bien au contraire que toutes deux se portent fort bien ! En biologie, le formalisme

---

limites techniques de l'intelligence artificielle (DUPUY J.-P., *op. cit.*, p. 94-96). Pour une critique serrée de J. Proust, voir Bernard STIEGLER, "Six thèses sur l'artifice techno-logique et l'effondrement du temps", in *Semiotica*, 1989, n° 77, 1/3, p. 339-362 (p. 347-348).

<sup>38</sup> *Ibid.*, p. 116.

<sup>39</sup> Je renvoie ici à la distinction que j'ai évoquée plus haut entre le point de vue formaliste et celui de l'autopoïèse (cf. supra, p. 10). Il est à noter également que Jean-Pierre Dupuy identifie le structuralisme de Lévi-Strauss et de Lacan au poststructuralisme (*ibid.*, p. 112-118), ce dont se démarque explicitement Jacques Derrida (HOY D., *art. cit.*, p. 45-46). Plus profondément, l'opposition que l'on voit se profiler ici entre la projection intentionnelle de Dupuy et Dennett, d'une part, et l'interprétation herméneutique radicalisée de Derrida, de l'autre, renvoie à deux conceptions antagoniques des sciences sociales et de leur épistémologie : d'un côté, un observateur indépendant et objectif sans aucun rapport avec l'agent social (cas de la projection intentionnelle), de l'autre, un observateur qui sait que ses propres théories et interprétations peuvent être l'objet d'une réappropriation consciente de la part des agents (cas de l'interprétation herméneutique). Voir à ce propos HAVELANGE V., "Sciences cognitives et tradition sociologique", in *Revue Internationale de Systémique*, vol. 8, n° 1, 1994, p. 79- 89.

<sup>40</sup> *Op. cit.*, p. 13 et p. 157 et suivantes.

règne presque sans partage sous les traits de la génétique formelle. Il habite les sciences sociales en l'espèce notamment de la théorie des jeux et de l'économisme ambiant. En sciences humaines, le structuralisme de Lacan et Lévi-Strauss se trouve reconduit, comme le note très justement Jean-Pierre Dupuy, dans l'élimination systématique de la subjectivité opérée par les sciences cognitives. En d'autres termes, formalisme et élimination du sujet sont quasiment hégémoniques. Pourquoi, dès lors, un tel pessimisme sous la plume de l'auteur ? Je souhaite suggérer pour conclure une contre-hypothèse : cette déception et cette amertume ne tiennent pas tant à ce que le formalisme et l'élimination du sujet n'ont pas triomphé qu'au fait que, tout simplement, ces deux orientations de pensée mènent inéluctablement au désespoir et au nihilisme. Les regrets de Jean-Pierre Dupuy en sont peut-être, à un second degré, l'aveu involontaire.

Véronique HAVELANGE  
UTC - PHITECO