



HAL
open science

La cybernétique à l'aune des sciences cognitives : une erreur scientifique ?

Mathieu Triclot

► **To cite this version:**

Mathieu Triclot. La cybernétique à l'aune des sciences cognitives : une erreur scientifique ? . Journée d'études "qu'est-ce qu'une erreur scientifique?", Université Lyon 3 Jean Moulin, May 2004, Lyon, France. hal-01526649

HAL Id: hal-01526649

<https://hal.science/hal-01526649>

Submitted on 23 May 2017

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

La cybernétique à l'aune des sciences cognitives : une erreur scientifique ?

Journée d'études « Qu'est-ce qu'une erreur scientifique ? », Université Jean Moulin Lyon 3
Lyon, Mai 2004

Mathieu Triclot, Doctorant Philosophie, Lyon 3

0. Introduction – 7
1. Filiation – 13
2. Erreur – 23
3. Jugement – 31

Dans cette intervention je vais aborder le concept d'erreur scientifique en étudiant un cas particulier, dans lequel un jugement d'erreur est porté. Il s'agit du cas de la cybernétique, ou de ce mouvement qui s'est appelé cybernétique, aux Etats-Unis, dans les années 1940-1950.

Ce jugement d'erreur ce n'est pas moi qui en suis porteur (ce n'est pas moi qui définit la cybernétique comme une erreur scientifique). Mais ce jugement est prononcé par un certain nombre de tenants des sciences cognitives.

Je vais m'appuyer en particulier sur le jugement de Marvin Minsky, qui est l'un des pères fondateurs de l'intelligence artificielle. Dans son introduction à *Semantic Information Processing*, en 1968, Minsky livre une analyse critique de la fécondité du mouvement cybernétique. La critique tient en trois thèses principales :

1. les résultats expérimentaux de la première cybernétique sont faibles et au mieux « paradigmatiques » (je cite Minsky).
2. La branche de recherche qui en est issue le plus directement, c'est-à-dire la théorie des systèmes auto organisateurs, n'a jusqu'à ce jour rien produit de valable
3. Pire, si la cybernétique se survit dans les sciences cognitives, et notamment dans l'intelligence artificielle, c'est au prix d'une révision drastique de ses concepts fondamentaux.

L'intérêt de cette critique sévère est qu'on la retrouve citée un peu partout, y compris chez des adversaires de Minsky, comme par exemple Hubert Dreyfus, le grand ennemi de l'intelligence artificielle. Etant donnée l'autorité de Minsky, on peut estimer que ce jugement a beaucoup compté dans l'évaluation de la cybernétique, et en définitive dans sa mise à l'écart. Dreyfus, par exemple, s'autorise dans *What computers can't do ?* du jugement de Minsky pour renvoyer la cybernétique aux oubliettes de l'histoire. Voilà un mouvement qui n'a rien produit de bon.

Mais il y a dans cette affaire un beau paradoxe, et une bonne dose d'ironie. Ironie, parce que les sciences cognitives, dont tout le monde s'accorde à dire qu'elles descendent de la cybernétique, portent un jugement bien désobligeant sur leurs ancêtres. Dans son histoire de la cybernétique, justement intitulée *Aux origines des sciences cognitives*, Jean-Pierre Dupuy insiste sur le fait que « les sciences cognitives ont honte de la cybernétique, lorsqu'elles reconnaissent avoir quelque chose en commun avec elle. »

Ma question est donc la suivante : Pourquoi les sciences cognitives condamnent-elles comme erreur le mouvement de pensée dont elles sont issues ? Quels sont les tenants et les aboutissants de cette histoire là ?

Au-delà de ce paradoxe généalogique, le jugement « la cybernétique est une erreur scientifique » comporte une autre difficulté ou une bizarrerie. Quand on dit la cybernétique est une erreur scientifique, de quelle erreur parle-t-on ?

L'erreur ne réside pas dans tel ou tel résultat particulier ou dans telle ou telle théorie, mais dans le mouvement lui-même.

Cette dimension est particulièrement nette dans la critique de Dreyfus. Ce qui est reproché à la cybernétique chez Dreyfus, ce ne sont pas ses résultats erronés, mais le fait qu'elle n'ait produit aucun résultat viable du tout.

Comme le note Dupuy, un tel jugement global a un effet bien précis : celui d'annuler, de mettre à distance son objet. Il n'invite pas à faire la part entre les bons résultats et les mauvais. Dupuy parle d'une « volonté d'oubli » de la part des sciences cognitives. Le jugement « la cybernétique est une erreur scientifique » a une fonction de déni. L'erreur est comme une faute dont il faut se garder.

Il me semble qu'on peut distinguer deux types d'erreurs pour clarifier le jugement porté sur la cybernétique. Le vocabulaire de Kuhn fournit ici un cadre commode. On peut distinguer ce que j'appellerai des erreurs à l'intérieur du paradigme et une erreur portant sur le paradigme. Quand on dit que la cybernétique était une erreur scientifique, on veut dire que le programme de recherche était mal orienté, de telle sorte qu'il n'a rien produit de bon.

Ce qui est intéressant c'est la façon dont un tel jugement est construit. Lorsque nous sommes dans le paradigme, nous apprend Kuhn, les jugements d'erreurs que nous pouvons porter sont normés par le paradigme. Le paradigme nous permet de faire le partage entre la vérité et l'erreur. C'est une de ses fonctions. En revanche, et c'est là une thèse célèbre de Kuhn, il n'y a pas de « preuve » au sens strict qu'un paradigme est meilleur qu'un autre, puisque les modalités de la preuve dépendent toujours d'un paradigme particulier. S'il n'y a pas de preuves, il existe en revanche, précise Kuhn, des manières d'évaluer un paradigme.

Il y a donc deux types d'erreur qui portent sur des objets différents : des objets théoriques dans le paradigme / le paradigme lui-même et qui sont associés à des modes de justification différents : dans un cas l'erreur peut être démontrée, dans l'autre elle ne peut être qu'évaluée. Il est clair que c'est à une telle évaluation que se livre Minsky lorsqu'il conclue à l'erreur de la cybernétique dans son ensemble.

En revanche, lorsque les logiciens Kleene et Arbib démontrent en 1956 et en 1961 que McCulloch et Pitts se sont trompés quant à la portée de leur modèle, ils ont affaire à un tout autre type d'erreur. Une erreur de logique, une erreur de calcul qui porte sur l'étendue de l'ensemble des fonctions calculables par un automate de McCulloch et Pitts. Cette erreur se démontre, elle ne s'évalue pas.

Il y a donc un usage tout à fait spécifique de la notion d'erreur scientifique dans la proposition « la cybernétique est une erreur scientifique à l'aune des sciences cognitives ». Sur quoi cette évaluation est-elle fondée ?

Il faut distinguer deux composantes dans ce jugement d'erreur. L'une est partagée au-delà des sciences cognitives et porte sur le « style scientifique » de la cybernétique. C'est sans doute la source majeure du sentiment de honte dont parle Dupuy. De façon plus ou moins consciente, la cybernétique s'est mise en porte-à-faux vis-à-vis d'un certain nombre de règles de la communauté scientifique.

L'autre composante du jugement d'erreur est en revanche propre aux sciences cognitives. Le jugement de Minsky appartient à cette seconde composante. L'intérêt du jugement de Minsky est qu'il révèle le point de blocage qui fait que la cybernétique a été abandonnée et est restée lettre morte pour les sciences cognitives. Minsky identifie ce point de blocage dans la question de la représentation. Et je pense qu'il a largement raison.

1. Filiation

Avant d'étudier ces deux composantes pour elles-mêmes, je vais commencer par rappeler brièvement en quoi a consisté la cybernétique et pourquoi celle-ci peut être considérée comme l'ancêtre des sciences cognitives.

La période cybernétique s'étend grosso modo de 1943 à 1956. On a du mal à imaginer aujourd'hui combien ce mouvement a été actif et influent. Il a rassemblé des dizaines de membres éminents appartenant à toutes les disciplines : des mathématiques et de la physique jusqu'aux sciences humaines, en passant par les sciences de l'ingénierie, la physiologie et la neurologie.

Les figures clés de la cybernétique sont les mathématiciens Norbert Wiener (qui forge le terme cybernétique), John Von Neumann (un des pères de l'ordinateur moderne), les physiologistes Warren McCulloch et Arturo Rosenblueth, le premier associé au jeune logicien prodige Walter Pitts. Il faut aussi rajouter Claude Shannon, l'inventeur de la théorie de l'information. Du côté des sciences humaines, il y a un fort contingent de psychologues, et quelques anthropologues dont le couple célèbre Bateson / Mead.

Outre son influence dans la communauté scientifique, le mouvement a connu une large diffusion publique. L'année 1950 marque le summum de la popularité cybernétique avec la publication des articles de Times, Newsweek et Life qui annoncent triomphalement la nouvelle ère des machines pensantes !

1943 est choisie comme date de naissance, parce que, en 1943, paraissent deux articles fondamentaux.

Il s'agit d'abord de l'article de Wiener, Bigelow et Rosenblueth, intitulé « Behavior, Purpose and Teleology ». L'article développe une description novatrice des dispositifs techniques en termes d'information et de feedback, c'est-à-dire d'action rétroactive. Le grand exemple cybernétique est le canon automatique de DCA. L'idée est que ce canon, lorsqu'il s'ajuste à sa cible au moyen de l'information dont il dispose, manifeste un comportement finalisé. La finalité n'est donc plus une propriété mystérieuse qui serait l'apanage des seuls êtres vivants, mais c'est un effet de certains dispositifs mécaniques complexes. C'est la bonne nouvelle cybernétique.

La seconde partie de l'article développe une analogie célèbre entre certains troubles du comportement humain et des défaillances du canon de DCA. L'action finalisée peut échouer, chez l'homme ou la machine, soit parce que le feedback est trop faible ou que les capacités d'ajustement feedback ont été détruites, soit parce que le feedback est trop fort et mène à un comportement oscillatoire catastrophique. Dans ce dernier cas, les ajustements succèdent aux ajustements sans jamais atteindre un point d'équilibre.

Fin 1943, paraît un autre article fondamental pour la cybernétique. Il s'agit de l'article de McCulloch et Pitts intitulé : « Un calcul logique des idées, immanent à l'activité nerveuse ». Les auteurs présentent un modèle simplifié du fonctionnement de l'activité cérébrale, dans les termes de la logique des propositions. Le modèle repose sur une description simplifiée des neurones comme des unités tout-ou-rien, organisées en réseau, et activées par l'action conjointe des autres neurones.

Pour prendre un exemple très simple : si deux neurones d'entrée sont nécessaires pour déclencher l'action d'un neurone de sortie, on est alors en présence d'un ET logique. Si un

seul des neurones d'entrée suffit à activer le neurone de sortie, c'est un OU... et ainsi de suite. On peut construire des réseaux de plus en plus complexes pour implémenter d'autres fonctions. N'importe quelle fonction logique affirment McCulloch et Pitts, qui croient à tort, comme je l'ai déjà signalé, que leur modèle est équivalent à une machine de Turing universelle, c'est-à-dire peut calculer la même classe de fonctions. Ce qui est faux.

L'article de 1943 est cité à plusieurs reprises par Von Neumann dans le célèbre *First Draft of a Report on the EDVAC*, de juin 1945, le document qui énonce les principes de l'ordinateur moderne. A partir de là peut se construire la seconde grande invention cybernétique, après la description mécaniste des processus finalisés : c'est-à-dire l'analogie entre l'ordinateur et le cerveau. La cybernétique est la première à faire du cerveau une sorte d'ordinateur naturel extrêmement complexe, qui est le siège de processus de traitement de l'information, analogues à ceux que réalisent nos ordinateurs.

Les deux grands modèles cybernétiques, la conduite de tir et l'ordinateur-cerveau, ont partie liée. L'ordinateur-cerveau est une sorte de cas particulier des nouvelles descriptions mécanistes de l'émergence de comportements organisés et finalisés.

De 1946 à 1953, le groupe cybernétique se retrouve de façon régulière dans une série de conférences organisées par la fondation Macy. Mais, à partir de 1952, la situation se dégrade. Wiener rompt avec le groupe cybernétique. Von Neumann qui est occupé par ses responsabilités militaires est de moins en moins présent. McCulloch continue à porter seul le flambeau de la cybernétique.

On pourrait choisir 1953, c'est-à-dire la fin des conférences Macy, pour clore la période cybernétique. 1956 me paraît une date plus appropriée.

D'une part, parce que les cybernéticiens sont restés productifs. En 1956, Von Neumann rédige un texte fondamental, qui constitue son testament philosophique, *l'ordinateur et le cerveau*. D'autre part, 1956 c'est la date de naissance officielle des sciences cognitives avec deux événements majeurs : la rencontre d'été de Dartmouth College entre les pères de l'intelligence artificielle (John McCarthy, Marvin Minsky, Allen Newell et Herbert Simon) et le symposium sur la théorie de l'information en septembre au MIT. A partir de là une nouvelle génération prend la relève.

S'il y a filiation c'est donc à deux titres. Les cognitivistes ont d'abord souvent été les élèves des cybernéticiens. C'est le cas par exemple de Minsky qui est l'élève de McCulloch. D'autre part, il y a une série de thèmes : l'information, les ordinateurs, le cerveau, etc, qui sont clairement l'invention de la cybernétique et qui constituent le lexique des sciences cognitives. Pour autant, la cybernétique ne joue plus aucun rôle déclaré dans les sciences cognitives. La filiation est méconnue ou déniée.

2. Erreur

Le jugement d'erreur qui frappe la cybernétique vise d'abord le « style scientifique » du mouvement.

Steve Heims, le grand historien américain de la cybernétique, rapporte une anecdote qui est tout à fait révélatrice du rejet que les travaux du groupe ont pu susciter chez certains scientifiques.

A la 5^e conférence Macy, Von Neumann avait décidé d'inviter le physicien Max Delbrück pour ses recherches en biologie sur les bactériophages. L'idée de Von Neumann était de faire

de Delbrück un membre permanent du groupe cybernétique. Le rapprochement était justifié du point de vue des cybernéticiens parce qu'ils travaillaient à l'époque sur des modèles d'automates autoreproducteurs. Ils pensaient partager avec Delbrück le projet d'une explication mécaniste des phénomènes de la vie.

La rencontre fut un échec sur toute la ligne. Delbrück est reparti fâché en jurant qu'on ne le reprendrait plus dans ce genre de discussions. Dans une lettre à Heims, dans les années 1970, il explique les raisons pour lesquelles il a refusé de participer au groupe et prononce un jugement sans appel sur la qualité des discussions : « vide au plus haut point et totalement inepte » (ce sont ses mots).

Je crois que l'impression de Delbrück est un bon exemple de ce que bon nombre de scientifiques « sérieux » ont ressenti face à la cybernétique et ses prétentions.

A l'intérieur du groupe cybernétique lui-même, certains ont cherché à prendre leurs distances. En 1956 par exemple, Shannon signe un éditorial virulent dans les *IRE Transactions in Information Theory*, dans lequel il dénonce la fanfare qui a accompagné la cybernétique et l'usage immodéré de mots à la mode comme information, entropie, redondance...

Ce rejet tient, à mon avis, principalement à deux éléments du « style cybernétique » :

1. La cybernétique est une entreprise interdisciplinaire. Est-ce qu'on peut avoir une pratique scientifique authentique dans un groupe interdisciplinaire ? C'est une question qui est posée à la cybernétique.

Dans son projet initial, Wiener mettait un certain nombre de gardes fous. Je le cite : « Le mathématicien n'est pas obligé de savoir conduire des expériences en biologie, mais il doit être capable de les comprendre, de les critiquer et d'en suggérer de nouvelles. A l'inverse, on ne va pas demander aux physiologistes de prouver des théorèmes mathématiques, mais ils doivent être capables d'en comprendre la signification physiologique et de dire aux mathématiciens dans quelles directions chercher. »

En pratique cet idéal de la discussion cybernétique a été rarement atteint. Dans les conférences Macy, le groupe est au prise avec une sorte de dilemme : soit, comme c'est le plus souvent le cas, les discussions portent sur des sujets extrêmement pointus, mais ce sont alors les généralistes en sciences humaines, comme par exemple Gregory Bateson, qui sont perdus et s'en plaignent, soit les discussions s'égarer dans des considérations générales.

Il s'agit là d'une difficulté qui est propre à toute entreprise interdisciplinaire.

2. Mais elle se redouble dans la cybernétique de l'usage systématique des analogies. Comme nous l'avons vu, elles apparaissent dès les deux articles de 1943. La pratique des analogies est très liée à l'interdisciplinarité, puisque c'est via l'analogie que l'on peut passer d'un domaine à un autre.

Toutes les analogies cybernétiques ne sont source d'erreurs. Mais il est souvent, dans les textes cybernétiques, difficile de distinguer entre des analogies heuristiques légitimes et des analogies dogmatiques qui deviennent des sortes d'images toutes faites qui masquent plus de problèmes qu'elles n'en découvrent. C'est typiquement le cas de la fameuse analogie entre les neurones et les tubes à vide qui devient une sorte de leitmotiv. Et, pendant toute la période cybernétique, on voit Von Neumann notamment, batailler contre une réification de l'analogie neurone-tube à vide. Von Neumann rappelle constamment qu'il y a d'autres déterminants du fonctionnement des neurones qui sont largement aussi intéressants que le caractère tout ou rien de la mise à feu.

Autant à l'intérieur de la cybernétique, il y avait sans doute une certaine prudence

quant à l'usage des analogies, autant à l'extérieur, ces analogies ont été prises pour argent comptant. C'est manifeste dans les articles de presse enthousiastes des années 1950. Ces articles ont irrités beaucoup de monde, y compris chez les cybernéticiens.

Au final, il me paraît évident que ces deux composantes du style cybernétique, l'analogie et l'interdisciplinarité, ont beaucoup fait pour la disqualification de la cybernétique. Les normes de la discussion scientifique étaient forcément beaucoup plus relâchées que dans des domaines disciplinaires reconnus et bien constitués.

Mais on peut aller plus loin dans la caractérisation de l'erreur.

Tout à l'heure, j'ai fait appel à la terminologie de Kuhn pour essayer de caractériser le type d'erreur qui est reproché à la cybernétique. Ici, il me semble qu'on peut mobiliser un autre modèle épistémologique. Je pense au concept d'idéologie scientifique, tel qu'il a été défini par Canguilhem dans l'article « idéologie et rationalité dans les sciences de la vie ». Ce sont les deux premières déterminations de l'idéologie scientifique qui m'intéressent. Canguilhem dit :

1. « les idéologies scientifiques sont des systèmes explicatifs dont l'objet est hyperbolique, relativement à la norme de scientificité qui leur est appliquée par emprunt ».
2. « Il y a toujours une science dans une idéologie, dans un champ latéral que cette idéologie vise obliquement. »

Une bonne part des erreurs cybernétiques relève de l'idéologie scientifique. La cybernétique s'appuie sur des domaines de savoir locaux, reconnus et validés, que ce soit la théorie de l'information, des éléments de théorie physique, de logique formelle, d'informatique théorique, etc. Mais elle les utilise pour constituer, au moyen des analogies, des objets hyperboliques qui transcendent chaque domaine particulier : l'information, l'organisme, l'automate... On a donc un ensemble de grands concepts qui proviennent de différentes disciplines reconnues, mais qui échappent pour partie aux normes de ces disciplines.

Les risques d'une telle pratique sont illustrées dans ce qui est sans doute un des plus beaux exemples d'idéologie scientifique dans la cybernétique. Je veux parler de l'analogie entre l'information et l'entropie.

Il se trouve que, pour des raisons mathématiques liées à l'usage des probabilités, la définition de l'information donnée par Shannon, dans sa théorie mathématique, est formellement analogue à la définition de l'entropie dans le théorème de Boltzmann.

La cybernétique, à la suite de Brillouin et Gabor, a cru que l'on pouvait faire de l'information une quantité physique opposée à l'entropie, une mesure de la quantité d'ordre et de structure dans un système physique.

L'idée est illustrée dans la situation du démon de Maxwell. Le raisonnement est le suivant : l'information prélevée par le démon sur le mouvement des particules lui permet de diminuer l'entropie du système dans lequel il opère. Ceci a pour conséquence de contredire la seconde loi de la thermodynamique. Mais, en réalité, disent les cybernéticiens, l'acquisition d'information a un coût thermodynamique, ce qui sauve la loi.

Il y a deux thèses :

- l'information est une quantité inverse de l'entropie. Posséder de l'information sur un système permet d'en diminuer l'entropie.
- mais prélever de l'information suppose de produire en retour de l'entropie.

Cette construction intellectuelle appartient au passé. Aujourd'hui, les travaux de Ralf Landauer, notamment, ont montré que ce n'est pas le processus d'acquisition de l'information qui a un coût entropique, mais le processus de destruction d'information avec la remise à zéro des mémoires.

Je ne veux pas entrer plus en détail dans cette affaire-là. Ce qui m'intéresse c'est de voir comment l'erreur des cybernéticiens a été produite. Il faut prendre en compte deux éléments :

1. Comme le dit Canguilhem, ils ont appliqué de façon hyperbolique les concepts d'un champ (la théorie des télécoms) dans un autre (la physique).
2. Mais, et c'est ce qu'on peut ajouter par rapport à Canguilhem dans un esprit bachelardien, cette construction théorique a été rendue possible dans la cybernétique par une valorisation induite de la notion d'information. Pour les cybernéticiens, l'information avait la valeur de l'ordre opposé au désordre, au bruit... y compris avec des interprétations éthiques et politiques.

Dans le monde de Wiener, celui de l'après-guerre et des débuts de la guerre froide, le monde d'Hiroshima, ça ne pouvait pas être la destruction mais la création d'information qui avait valeur, qui avait un coût. Il y a une impossibilité à imaginer que ce soit la remise à zéro des mémoires qui ait un coût ou une valeur. Derrière l'analogie entre information et entropie, il y a tout un réseau d'équivalence d'arrière-plan entre l'éthique, la politique, la physique, la théorie des télécoms, qui est extrêmement prégnant dans la cybernétique.

Non seulement on a le passage entre différents champs de savoir, mais ce passage est réalisé, dans le cas information/entropie, au moyen de considérations axiologiques sur la valeur de l'information.

3. Jugement

C'est typiquement de ce style de pensée, qui se laisse bien décrire par le concept d'idéologie scientifique, que les sciences cognitives peuvent avoir honte, et dont elles cherchent à se démarquer.

Mais, si on regarde chez Minsky, ce n'est pas du tout ça qui le gêne. Ce n'est pas ça qu'il condamne dans la cybernétique. Il faut donc bien prendre garde de distinguer entre une critique de la cybernétique comme idéologie scientifique si l'on veut, et une certaine critique de la cybernétique qui provient des sciences cognitives, de l'IA notamment, et qui lui reproche tout à fait autre chose que son style scientifique.

Dans un petit paragraphe de *Perceptrons* en 1969, intitulé « cybernétique et romantisme », Minsky et Pappert font la mise au point suivante : « Nous avons le sentiment que ces experts solennels qui ont le plus reproché leurs « prétentions exagérées » aux cybernéticiens ont, tout compte fait, commis une erreur beaucoup plus grande. »

Minsky, Pappert sont engagés dans une opération d'évaluation beaucoup plus complexe de la cybernétique, qui porte sur les concepts d'arrière-plan qu'elle a développée (et non sur son enthousiasme excessif).

La situation de Minsky est intéressante parce qu'il se retrouve coincé entre deux exigences opposées :

- Spécifier l'erreur de la cybernétique
- tout en sauvant Warren McCulloch, son maître et son mentor, celui qu'il appelle encore son héros, en 1985 dans *Society of Minds*.

L'exposé le plus clair de l'erreur cybernétique se trouve dans l'introduction à *Semantic Information Processing*. Minsky y distingue trois champs de recherche issus de la cybernétique :

1. la théorie des systèmes auto-organiseurs, qui est le descendant direct de la cybernétique : elle consiste à étudier l'émergence de comportements finalisés à partir de l'interaction entre des composants mécaniques.
2. Il y a ce que Minsky appelle la simulation cognitive, branche de recherche rattachée aux travaux de Newell et Simon. Leur but est de simuler des capacités cognitives humaines, au plus près de la façon dont elles sont effectuées par les êtres humains.
3. Enfin, il y a la branche de recherche de Minsky lui-même, qui consiste à créer des programmes qui manifestent un comportement intelligent, sans que celui-ci soit nécessairement le reflet des comportements humains.

Dans la simulation cognitive on imite l'intelligence humaine. Dans l'intelligence artificielle à la Minsky on construit une intelligence des machines sans référence anthropologique.

Ces deux derniers programmes de recherches font partie des sciences cognitives. Qu'est-ce qui les distingue de la cybernétique ? L'erreur de la cybernétique nous dit Minsky est de ne pas avoir cherché à construire une théorie de l'esprit. La simulation cognitive comme l'IA de Minsky supposent d'isoler un niveau autonome, celui des représentations, sur lequel on effectue des manipulations réglées. La cybernétique n'a pas cherché à définir un niveau du traitement de l'information indépendant des processus matériels.

Dans son introduction aux sciences cognitives, Daniel Andler résume en 3 thèses le paradigme cognitiviste : «

1. Le complexe esprit/cerveau est susceptible d'une double descriptions matérielle ou physique et informationnelle ou fonctionnelle ; ces deux niveaux sont largement indépendants, et le rapport qui s'établit entre eux est à l'image de celui qui lie un ordinateur en tant que système physique à la description du même appareil en tant que système de traitement de l'information.
2. Au niveau informationnel le système cognitif de l'homme est caractérisé par ses états internes ou mentaux et par les processus qui conduisent d'un état au suivant. Ces états sont représentationnels : ils sont dotés d'un contenu renvoyant à des entités externes.
3. Les états ou représentations internes sont des formules d'un langage interne ou « mentalais » proche des langages formels de la logique. »

La cybernétique n'a pas isolé ce niveau autonome, analogue au software, de traitement de l'information. Pour une raison que nous avons rencontré lors de l'analyse de la première dimension de l'erreur : la cybernétique a cherché à faire de l'information une propriété des états physiques, une expression de la quantité d'ordre ou de structure dans les agencements matériels.

Dans le paradigme cognitiviste, l'information représente un niveau formel autonome qui peut s'implémenter comme un programme dans différentes architectures matérielles, cerveau ou ordinateur par exemple, sans que cela change rien à la structure de haut niveau du programme. C'est la métaphore de ce qu'on appelle en informatique la compilation. En revanche, dans le paradigme cybernétique, l'information est un effet ou une propriété de l'organisation matérielle.

Pour illustrer ce point de divergence, je vais prendre deux exemples.

Premièrement, avec le recul, on a tendance, je crois à mésinterpréter l'article de 1943 de McCulloch et Pitts. On y voit le premier énoncé programmatique d'une théorie de l'esprit. Or, si on lit le texte en s'attachant au lexique, on se rend compte qu'il n'y a dans le texte qu'une

seule occurrence du terme esprit... et encore pour dire « je n'emploierai pas ce terme » ! Cette réticence à employer le vocabulaire de l'esprit et à isoler un niveau de représentation qui deviendrait l'objet d'étude par excellence tient à la fois au contexte behavioriste, mais aussi au projet de McCulloch qui me paraît, au moins en 1943, franchement réductionniste, faisant l'impasse sur ce niveau autonome des représentations qui constitue le cœur des sciences cognitives.

Deuxième exemple pour illustrer la divergence entre le projet des sciences cognitives et celui de la cybernétique : le texte de 1956, *l'ordinateur et le cerveau* de Von Neumann. Ici encore, je ne compte aucune occurrence du terme esprit. Quand Von Neumann veut désigner avec la plus grande généralité son objet d'étude, il emploie le terme automate (une trentaine d'occurrences). La thèse massive développée par Von Neumann est que nos langages de haut niveau, nos processus de traitement de l'information, dépendent profondément des langages de bas niveau et de l'organisation matérielle du système nerveux. On ne peut pas isoler un niveau fonctionnel autonome en faisant l'impasse sur l'architecture matérielle.

On ne peut pas mieux dire l'erreur de la cybernétique. Celle-ci s'est refusée à isoler un niveau de calcul et de représentation comme les sciences cognitives l'ont fait. Ce qui est tout à fait fascinant quand on compare les sciences cognitives et la cybernétique c'est qu'on retrouve sans mal un lexique commun : il est question d'information, d'ordinateurs, de cerveau, de machines de Turing, etc. Mais les termes ne sont pas du tout distribués de la même façon et les significations attachées aux termes diffèrent. Je pense notamment que le terme information a un usage différent en cybernétique et en sciences cognitives. Cela, je crois que Minsky l'a bien vu.

L'erreur cybernétique signifie donc ici : ils n'ont pas cru bon de dégager un niveau formellement autonome des représentations.

Mais il y a une deuxième composante dans le jugement d'erreur de Minsky. Minsky cherche à exonérer McCulloch de l'erreur cybernétique.

La tactique consiste à soutenir que McCulloch faisait en réalité œuvre de logicien. Le système nerveux en tant que tel n'était pas son objet. Là aussi, rétrospectivement, cette évaluation apparaît pleine d'ironie involontaire. Faire de McCulloch un logicien c'est certes le sauver de l'erreur cybernétique, mais nous avons vu que l'œuvre de logicien de McCulloch est plutôt maigre et qu'il s'est largement trompé quant à la portée logique de son modèle.

L'évaluation de Minsky repose donc sur deux actes théoriques :

- elle isole, à juste titre, l'absence de la question de la représentation chez les cybernéticiens en général
- elle conduit à réinterpréter l'œuvre de McCulloch dans un sens cognitiviste, pour l'exempter de l'erreur cybernétique.

Pour conclure, nous avons croisé trois niveaux d'erreur dans la cybernétique. De bas en haut :

1. l'erreur sur la portée logique du modèle de McCulloch et Pitts
2. l'erreur de l'analogie entre information et entropie qui est un condensé du style cybernétique
3. l'erreur portant sur la question des représentations.

Ce qui est intéressant dans ce troisième type d'erreur c'est qu'elle est le résultat d'un processus d'évaluation, qui engage des valeurs sur ce qui compte ou pas dans un programme de recherche. Ce qui caractérise l'évaluation, contrairement à la preuve, c'est qu'elle n'est jamais close.

Dans le cas qui nous intéresse, le programme de Minsky a été sérieusement battu en brèche dans les années 1980 et on a vu revenir, sous le nom de néo-connexionnisme, des programmes dont l'inspiration était beaucoup plus cybernétique... mais entre temps la référence à la cybernétique s'était perdue.