

## **Bio-mimétisme : design inspiré par la nature**

Simona MORINI  
*Università Iuav di Venezia*

Dans un roman de Douglas Adams, *Guide intergalactique pour auto-stoppeurs*, une race aliène super intelligente demande au calculateur le plus puissant jamais construit, Deep Thought, de donner « la réponse à la vie, à l'univers, à toute chose ». La machine dotée de toutes les informations imaginables et des meilleures règles de raisonnement, commence à traiter la question et, après des années de travail, répond enfin : « 42 ». Réponse décevante et dépourvue de sens.

Même en dehors de la science-fiction, par exemple dans leur fonctionnement quotidien, les ordinateurs tendent aussi à produire des "bizarreries". Les dossiers disparaissent, les caractères se transforment, les icônes sur l'écran augmentent ou rapetissent, les calculs se bloquent, nous commençons alors à appuyer sur les touches, à secouer la souris, souvent sans résultat. La machine ne réagit pas. A ce moment-là, notre ordinateur super intelligent n'est pour nous qu'un "appareil stupide". La platitude de Deep Thought – de même que l'insensibilité de nos machines – comparée aux réponses ingénieuses produites par l'esprit humain en dépit de toutes les imperfections que nous lui reconnaissons, est assez symptomatique de l'écart qui existe entre les objets produits par la "nature" (systèmes physiques et biologiques) et les objets construits par l'homme (systèmes computationnels et objets technologiques, en général).

Un ordinateur est en mesure d'effectuer des calculs complexes en l'espace de quelques secondes et sans effort, alors que la capacité de calcul "spontanée" du cerveau humain est plutôt de l'ordre de celle d'un pigeon ; pourtant, le cerveau humain sait exploiter ses capacités limitées de façon extraordinaire alors qu'un ordinateur, en quelque sorte, "gaspille" ses compétences (ainsi qu'une quantité d'énergie remarquable, disproportionnée à la tâche) et fournit parfois des réponses insensées. Malgré ses prestations

exceptionnelles, Deep Thought ne parvient pas à élaborer de théorie audacieuse ni à imaginer de solutions anticonformistes. De même, un avion peut voler à très grande vitesse mais il ne fait rien de toutes ces choses utiles et ingénieuses que sait faire un moineau : voler de branche en branche, se poser, construire un nid, migrer quand il fait froid, s'orienter sans boussole ni instruments, chanter, etc.

Aujourd'hui encore, les objets avec lesquels nous avons essayé de "potentialiser" la nature, voire de la plier à nos objectifs et à nos exigences, nous semblent gauches, rigides, limités.

Nous ne savons plus où mettre leurs carcasses, elles encombrant nos maisons et nos villes comme dans les romans de science-fiction les plus noirs. Une stratégie défensive possible consiste à les "camoufler" ou, quand c'est possible, à les recycler ; l'autre, plus intéressante réside dans une modification profonde de nos façons de concevoir les artefacts.

Si nous pensons aux ordinateurs – ou aux avions – et considérons le concept qui les a inspirés, nous constatons que ces machines ne sont pas à proprement parler des "imitations" de la nature, mais plutôt la réalisation d'un modèle abstrait qui ne possède avec la nature que quelques propriétés formelles. C'était, du moins, l'idée qu'en avaient leurs premiers concepteurs. L'idée fondamentale est bien celle de *calculabilité*, sur laquelle la théorie d'Alan Turing se fonde et qu'il énonce dans deux articles datant de 1931 et de 1950. Selon cette théorie, l'esprit n'est pas l'équivalent du cerveau mais il est le résultat du traitement d'informations accumulées, élaborées et transformées au moyen d'opérations réalisées matériellement par le cerveau. Quelque chose d'"immatériel" comme une pensée abstraite peut donc détenir sa base dans les opérations physiques du système nerveux mais aussi dans celles d'un quelconque "mécanisme physique" - ou système "artificiel" – qui soit en mesure de remplir les mêmes fonctions, c'est-à-dire d'exécuter "pas à pas" un calcul, un algorithme apte à simuler tout ce qui pourrait être effectivement *décrit*. Une machine de Turing est donc un modèle formel de la notion intuitive de calcul. Du fait que les processus de pensée se caractérisent par le calcul comme l'entend Turing, en tant que séquences de passages élémentaires constitués d'opérations sur des symboles, il s'ensuit que ce que nous appelons "intelligence" peut être réalisé par des programmes informatiques suffisamment sophistiqués et que l'esprit peut être étudié indépendamment de son substrat neurologique.

Cette perspective, selon laquelle le calcul, l'information et les processus de communication jouent un rôle déterminant, a ensuite été développée par Norbert Wiener, le père de la cybernétique. Dans la conception de Wiener, le réel est un pur jeu de différences entre les éléments (les objets) qui le composent. Ces différences ne concernent pas seulement la nature des objets, mais aussi et surtout leur niveau de complexité et les relations qui s'établissent entre eux. Elles concernent l'*intérieur* des choses (l'invisible) et leur extérieur (le visible). L'intérieur des choses – qui comprend leur organisation, leur structure, leurs "propriétés" – en un mot : l'"invisible" – est analysable par une méthode "fonctionnelle". Au contraire, la méthode qu'il

dénomme "comportementale" néglige la structure et l'organisation interne, invisible de l'objet et se concentre sur ses relations ("visibles") avec l'environnement<sup>1</sup>. Ces deux niveaux différents d'analyse permettent d'éclairer de façon tout à fait nouvelle le rapport entre l'homme et la machine – et entre objets naturels et objets artificiels, en général.

Les différences profondes qui existent entre l'homme et la machine et que l'analyse fonctionnelle identifie dans la diversité de leur composition matérielle ne sont pas pertinentes pour l'analyse comportementale. Comme le disait Hilary Putnam en son temps (mais aujourd'hui, il s'en "repent"), l'esprit peut aussi bien être fait de silice que de fromage, car ce sont ses propriétés organisatrices qui en font un cerveau, l'organisation pouvant, en principe, être réalisée à partir du substrat que l'on veut.<sup>2</sup>

De même, dans la conception de Wiener, le monde est peuplé d'entités (naturelles et artificielles) dotées de comportements, leur composition matérielle ne les différenciant qu'à un niveau secondaire. « Machines et êtres humains ont les mêmes comportements. Certains d'entre eux, spécifiques, se trouvent exclusivement dans les unes et non dans les autres. Il n'y a encore aucune machine qui sache écrire un dictionnaire sanscrit-mandarin. De même, il n'y a pas d'organismes vivants qui se déplacent sur roues : imaginez la situation si les ingénieurs s'étaient obstinés à copier les organismes vivants et à munir les locomotives de jambes et de pieds ! »<sup>3</sup>

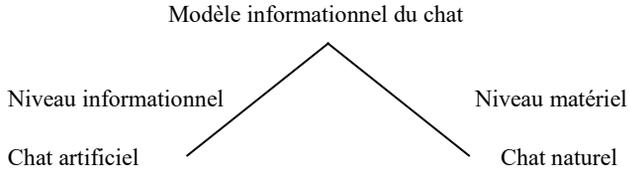
Les relations entre les divers comportements sont assumées par des modèles et, en fonction de ceux-ci, les objets naturels et artificiels sont tout simplement réalisés sous forme de supports matériels différents – naturels ou artificiels, justement. De ce point de vue, un chat artificiel n'est pas l'imitation d'un chat naturel. Les deux objets ne sont pas en confrontation, mais en relation par rapport à un modèle abstrait de comportement qui caractérise l'entité abstraite "chat" à laquelle les deux – le chat naturel et le chat artificiel – doivent se conformer. Donc, les objets artificiels ne sont pas à proprement parler des répliques ou des imitations du vivant, mais bien les "incarnations" d'un modèle d'ordre logique, supérieur, dont le vivant est une autre incarnation. Un chat naturel et un chat artificiel sont, dans ce sens, des objets mixtes, constitués d'un support matériel qui les différencie et d'un modèle informationnel qui les rapproche.

---

<sup>1</sup> Mais aujourd'hui, ce qu'on appelle « fonctionnalisme » est lié à la conception d'un état mental qui est tel s'il joue une fonction donnée ou recouvre un certain rôle dans le système - cerveau considéré dans sa globalité. Pour une discussion claire du fonctionnement computationnel, voir Paternoster A. (2002).

<sup>2</sup> Putnam (1960).

<sup>3</sup> Rosenbluth, Wiener, Bigelow (1943, p. 22).



Au niveau informationnel, enfin, les objets se distinguent sur la base de la complexité de leurs comportements. Certaines actions partent de l'objet pour modifier l'environnement, d'autres proviennent de l'environnement et modifient l'objet. De plus, certaines actions sont "rétroactivement" intentionnelles, cela signifie qu'elles sont guidées par une finalité qui commande leur comportement; ceci différencie les objets quant à la complexité de leurs comportements. En principe, il se peut donc qu'un objet artificiel doté de mécanismes sophistiqués de rétroactivité intentionnelle "surpasse", sur le plan "visible" des comportements, un objet naturel. Tous les êtres (objets) sont équivalents, à condition qu'ils se trouvent au même niveau de complexité informationnelle. Dans cette perspective, tout est visible et transparent, et le rapport intérieur /extérieur est renversé. Du moment que tout est connaissable à travers le comportement, l'intérieur (l'invisible) se retourne comme un gant et devient "visible".

C'est dans le cadre de cette conception que le test de Turing acquiert un sens, tout comme le premier ordinateur conçu par John von Neumann<sup>4</sup>. Il est assez significatif que le projet de l'EDVAC ne soit pas un projet "d'ingénieur". Le texte de von Neumann ne contient aucun dessin ni aucune représentation graphique. C'est une suite de définitions et de descriptions. Les analogies entre cerveau et machine sont, exactement comme pour Wiener, des analogies de comportement. Le modèle fonctionne indépendamment de la nature du support. Ainsi, le premier ordinateur est le résultat d'une tentative abstraite pour comprendre le fonctionnement de l'intelligence en tant qu'élaboration des informations (c'est-à-dire, en tant que comportement, résultat (visible) en sortie (*output*)), avec transfert de cet "algorithme" sur un support matériel. Il s'agit de créer une "incarnation" du même modèle sur un support différent. La machine se compose donc de parties (l'unité de calcul, l'unité de contrôle logique et la mémoire) qui doivent répondre au modèle abstrait sur la base duquel fonctionnent les neurones. Alors, il s'agit tout simplement de trouver des matériaux qui soient en mesure d'assumer la même fonction que les matériaux biologiques, tout en étant différents : meilleurs sur certains points, pires sur d'autres.

Une fois que la machine est réalisée, on peut l'interpréter, certes, comme une imitation de la nature, mais cette terminologie est impropre et fourvoie. Bien sûr, on peut penser à un avion comme à un grand oiseau, mais l'avion a été construit selon un schéma dont seul le plan du comportement visible

---

<sup>4</sup> Pour une reconstruction historique de ces thèmes et une discussion théorique, voir Licata (2008).

l'apparente à l'oiseau et par rapport auquel la structure invisible n'est pas pertinente. De même, une fois que le comportement "somme de deux nombres" est décrit, l'opération peut être exécutée soit par notre cerveau, soit par un réseau neuronal artificiel, un boulier ou un programme LISP. Dans les deux cas, nous sommes en mesure de reproduire certaines opérations sans qu'aucune forme d'imitation n'ait été produite. Nous pouvons, éventuellement, parler d'*émulation* d'un modèle<sup>5</sup>. Pour un mathématicien, « quoi qu'il fasse, le cerveau le fait "en vertu" de sa structure, en tant que système logique, mais pas parce qu'il se trouve dans la tête d'une personne (...), cette même structure logique pouvant être représentée, tout aussi bien, par d'autres moyens »<sup>6</sup>. Comme l'écrivait Wiener : « Le fait que nous ne puissions pas télégraphier le modèle d'un homme d'un lieu dans un autre provient plutôt, sans doute, d'obstacles techniques [...] que du résultat d'une impossibilité de l'idée en soi »<sup>7</sup>.

Le modèle abstrait assume donc certaines propriétés du comportement naturel (qu'il soit humain ou pas), communes à des supports matériels très différents mais indépendantes de ceux-ci. Le terme *émulation* que nous avons employé précédemment correspond, dans la terminologie qu'utilise John Searle dans un célèbre article<sup>8</sup>, à l'Intelligence Artificielle "faible" ou "prudente", selon laquelle la principale valeur d'un ordinateur (d'un objet artificiel) est de nous aider à « formuler et contrôler des hypothèses (sur le cerveau) de façon rigoureuse et précise » et d'en reproduire certaines fonctions circonscrites. Mais, soutenir que l'ordinateur puisse *imiter* l'intelligence humaine signifie que l'on considère un programme d'Intelligence Artificielle "fort" ; c'est dire que les ordinateurs "dotés de programmes appropriés" pourraient se trouver dans des états cognitifs comparables aux états naturels du cerveau humain et ne se distingueraient pas de celui-ci (ce qui revient à soutenir qu'une machine pourrait réussir le test de Turing) : « Dans l'Intelligence Artificielle forte, étant donné que l'ordinateur a des états cognitifs, les programmes ne sont pas de simples instruments pour mettre à l'épreuve nos explications psychologiques, mais ils sont eux-mêmes des explications »<sup>9</sup>.

A la question fatidique « Une machine peut-elle penser ? », Searle répond que seuls les cerveaux (ou les machines munies "de pouvoirs logiques internes" équivalents) le peuvent et que l'Intelligence Artificielle "forte" « n'a pas grand chose à nous dire sur la pensée car elle concerne les programmes, pas les machines, et qu'aucun programme à lui seul n'est suffisant pour penser »<sup>10</sup>.

---

<sup>5</sup> Pour une discussion intéressante sur ce point, voir Boncinelli, Giorello (2008, p. 53 et suivants)

<sup>6</sup> Hodges (1983, trad., it. p. 380).

<sup>7</sup> Wiener (1951, p. 141).

<sup>8</sup> Searle (1980).

<sup>9</sup> *Ibid.*, p.2.

<sup>10</sup> *Ibid.*, p.1.

L'un des points forts de la critique de Searle – avec d'autres du même genre – porte sur la "cécité" des computations quant au sens<sup>11</sup>. En effet, chez les humains, l'introspection et l'observation des comportements montrent que les processus cognitifs sont "chargés" de sens. Ce sont des acteurs intentionnels guidés par des états représentationnels (intentions). Dans sa célèbre pièce chinoise, Searle prend comme exemple de système computationnel un programme d'élaboration du langage naturel et le compare à un homme enfermé dans une pièce, ignorant tout du chinois, et qui parvient pourtant à donner, en chinois, des réponses écrites et appropriées à des questions écrites en chinois, grâce à des instructions rédigées dans sa propre langue où il est expliqué comment faire correspondre les symboles chinois en entrée avec les symboles chinois en sortie. L'homme ignore que ce qu'il reçoit sont des questions et qu'il renvoie des réponses : il exécute une simple tâche mécanique qui consiste à associer des symboles entre eux. Donc, selon Searle, les systèmes artificiels ne comprennent pas le langage parce qu'ils n'ont pas accès aux propriétés *sémantiques* des symboles qu'ils manipulent. Comme l'homme dans la pièce chinoise, les ordinateurs exécutent des instructions sans vraiment comprendre ce qu'ils sont en train de faire.

On peut penser, toutefois, que ce manque d'appréhension du sens dépend du fait que cet homme applique des règles sans encourir la moindre conséquence et qu'il est "enfermé dans la pièce" sans pouvoir établir de relations avec son interlocuteur, et que ce manque dépendrait du fait que l'échange linguistique n'a pas lieu dans le contexte social qui est indispensable pour son bon fonctionnement. Mais si on suppose, par exemple, que l'homme sera puni ou récompensé en fonction de certaines réponses ou, encore, qu'il sera en mesure de distinguer, chez son interlocuteur, des réactions positives ou négatives, cela donnerait déjà un sens (même élémentaire, "positif" ou "négatif") à certains signes qu'il manipule. Ceci a lieu parce que l'homme n'est pas une machine (eu égard au matériau qui le compose) et qu'une réponse négative peut "lui porter atteinte" tandis qu'une récompense peut "l'encourager". C'est cette préoccupation "biologique" qui l'induit à "apprendre", à "comprendre" et à attribuer une "signification", un "sens" à ses propres opérations, c'est-à-dire à activer les mécanismes qui consentent de *choisir* certaines actions (délibérément, intentionnellement) et d'en écarter d'autres. Afin que le "*sens*" surgisse, il faut doter la machine de "signaux d'orientation" : désirs, peurs, projets qui l'allèchent ou l'épouvantent ; d'une manière générale : d'émotions positives ou négatives. Il faut que la machine (ou l'homme dans la pièce chinoise) "se préoccupe".<sup>12</sup>

---

<sup>11</sup> Paternoster (2002, p. 6).

<sup>12</sup> L'importance de la composante émotive dans la décision a été mise en lumière par Damasio (1994). Il a miné, de façon convaincante, l'idée que rationalité et émotivité sont des facultés mentales opposées. Au contraire, c'est par ces constituants complémentaires et indispensables qu'il est possible – au niveau neurophysiologique – de parvenir à une décision.

Les composants électroniques étaient, du temps de Wiener, le matériau le plus adapté pour réaliser un modèle abstrait de raisonnement ; Wiener n'excluait pas, toutefois, la possibilité d'en trouver de meilleurs : « Si un ingénieur devait construire un robot dont le comportement ressemblerait *grosso modo* à celui d'un organisme animal, il ne le construirait pas aujourd'hui avec des protéines et des matériaux colloïdaux. Il le ferait, probablement, à partir d'éléments métalliques, de composants électriques et électroniques. Les mouvements du robot pourraient être plus rapides et plus puissants que ceux de l'organisme original. Mais ses capacités d'apprentissage et de mémoire seraient plus rudimentaires. Dans l'avenir, en raison des progrès de nos connaissances sur les colloïdes et les protéines, les ingénieurs chercheront à construire des robots qui, non seulement, auront un comportement analogue à celle d'un mammifère mais aussi la structure. Le modèle ultime pour un chat est, certes, un autre chat, qu'il provienne d'un semblable de son espèce ou qu'il ait été construit en laboratoire ». <sup>13</sup>

L'avenir dont parle Wiener est assez proche de *notre* présent ; le support matériel, quant à lui, prend sa revanche. En effet, si les machines accomplissent le travail des hommes de façon remarquable, il n'en reste pas moins que l'importance du support matériel de ces machines s'est accrue. Si, en ce qui concerne les phénomènes, notre cerveau trouve des explications ingénieuses et sensées (même erronées) alors qu'une machine peut tranquillement donner une réponse insensée, telle que le "42" formulé par Deep Thought, c'est parce que, à la différence des machines, nos cerveaux "tournent" sur un support périssable, vulnérable, qui doit économiser forces et énergies, supporter la fatigue pour apprendre, pour se procurer des informations et pour les conserver et les organiser. En somme, les objets naturels doivent lutter pour survivre tandis que les objets artificiels sont "dénaturés", "indifférents" ; mais c'est précisément pour cela qu'ils ne savent pas encore orienter leur comportement, s'adapter, développer la flexibilité, la capacité de transformation et de mutation qui constitue la richesse de la nature par rapport à l'artificiel, de l'oiseau par rapport à l'avion, du cerveau par rapport à l'ordinateur.

C'est, notamment, la thèse que soutient Read Montague dans un ouvrage où il recherche la nature des processus décisionnels à l'aide d'un modèle computationnel qui relève de l'informatique, de la théorie du contrôle et de la théorie statistique de l'apprentissage <sup>14</sup>. Selon Montague, le cerveau est capable de trouver des réponses *sensées* aux problèmes qu'il se pose parce qu'il est "lent, bruyant et imprécis", c'est-à-dire qu'il est tout à fait différent de la séquence élégante, rapide et précise de symboles qui composent un algorithme opérant sur un matériau moins fragile – plus "indifférent" – que l'organique. Le calcul qui n'est que manipulation de chaînes et de symboles, n'a pas besoin du sens. Même les machines qui "pensent", comme les ordinateurs, "ne se préoccupent pas". Mais les hommes, oui.

---

<sup>13</sup> Rosenbluth, Wiener, Bigelow (1943).

<sup>14</sup> Montague (2006).

Au début des années 1990, un groupe de chercheurs du MIT guidés par Tomaso Poggio, chef de file de l'interprétation computationnelle de l'intelligence biologique, conjointement au Centre de Neurosciences computationnelles du Caltech et au Computational Neurobiology Lab (CNL) où Montague travaillait alors, ont proposé d'interpréter les opérations mentales en termes computationnels – faisant passer le modèle "classique" du champ purement théorique au champ opérationnel et concret de la décision – à partir des "choix" élémentaires opérés par des molécules et des bactéries comme l'*Escherichia coli* jusqu'aux choix beaucoup plus complexes qu'effectuent les êtres humains et les groupes sociaux. L'idée de base demeure celle d'Alan Turing, de Norbert Wiener et de John von Neumann : "la matière pensante" est faite de *modèles* d'informations emmagasinées, élaborées et transformées dans notre cerveau ou encore dans la reformulation qu'en donne Montague : « ton esprit n'est pas l'équivalent de ton cerveau ; il est le résultat de l'élaboration des informations qui est prise en charge par ton cerveau »<sup>15</sup>. Mais la nouvelle théorie inclut, dans le modèle classique, les mécanismes d'*évaluation* qui distinguent le calcul biologique du travail des ordinateurs. L'étude quantitative des processus d'évaluation en termes d'*objectifs* et de *signaux d'orientation* est un point fondamental pour la théorie de l'*apprentissage de renfort* qui constitue le cadre dans lequel on étudie comment, généralement, une machine peut apprendre de l'expérience.

Partant de la constatation que nous sommes des « configurations 'pleines de sens' qui, non seulement exécutent des opérations de traitement d'informations desquelles dépendent la perception et le mouvement mais sont aptes, aussi, à les évaluer »<sup>16</sup>, Montague soutient que les valeurs (et les buts qui les engendrent) ne nous arrivent pas de l'extérieur (voire d'en haut, de quelque "divinité" ou autorité externe), mais qu'elles font partie intégrante de chacune de nos représentations et sont finement "prises dans l'engrenage" des mécanismes de fonctionnement du cerveau. Dans notre machine, il n'y a pas seulement un flux de symboles qui "tourne", il y a aussi un flux d'associations "symbole - valeur" accompagnées de toute une série de signaux d'orientation corrective (signaux de récompense et d'erreur) qui sont générés par un système complexe neurophysiologique d'"apprentissage de renfort" orienté vers un but. En bref : l'esprit identifie des buts qu'il cherche à atteindre par tentatives et erreurs, "évaluant" chaque résultat, dans un grand souci "d'économie" ce qui implique : un moindre effort pour un rendement maximum !

Tout le livre de Montague gravite autour de l'analyse du lien fondamental entre la vie (à tous les niveaux et sous toutes ses formes) et les mécanismes d'évaluation. Ce sont les évaluations (et les objectifs qui lui sont associés) qui font entrer la signification dans les calculs biologiques. Si, dans la pièce chinoise, l'homme pensait que sa survie ou quelque chose d'important pour lui dépend des messages qui lui parviennent (et cela comporte, évidemment,

---

<sup>15</sup> *Ibid.* p. 11

<sup>16</sup> *Ibid.* p. 231

une forme d'interaction avec ce qui est à l'extérieur de la pièce), il ferait ce qu'il est nécessaire pour leur trouver une signification. Alors que les évaluations des êtres plus simples biologiquement (comme l'*Escherichia coli* ou la majeure partie des animaux) sont guidés par des objectifs primaires liés à la survie et à la reproduction (nourriture et sexe), Montague démontre de façon convaincante comment le cerveau humain s'est développé afin de pouvoir conférer, même aux pensées les plus abstraites, un pouvoir directif sur nos actions. Ce "super pouvoir" qui confère aux idées la capacité d'obtenir une récompense de valeur comparable à celle des récompenses primaires – sexe, nourriture et autres mécanismes qui règlent habituellement la survie biologique – expliquerait pourquoi les élaborations à caractère social peuvent avoir un vrai "impact neural" et pourquoi les êtres humains – à l'opposé de la plupart des animaux<sup>17</sup> et des machines – sont prêts à se sacrifier, à prendre des risques et à mourir non seulement pour leur survivance mais aussi pour des idées telles que l'égalité sociale, la justice ou la religion.

Cela signifie que ce dont les objets artificiels et les programmes informatiques actuels ont besoin, afin de pouvoir *imiter* les êtres vivants – et non seulement de les stimuler – ce sont des *objectifs*. Un objectif, en effet, ne peut se réduire à une instruction fixée à l'avance qui déterminerait une fois pour toute le comportement d'un système. Les instructions s'avèrent une méthode excellente si la solution du problème est déjà connue (comme dans un logiciel de jeu d'échecs qui indique le coup correspondant à chaque nouvelle situation). Mais souvent, la plupart des actions et des choix adviennent dans un contexte en changement perpétuel, incertain et imprévisible ; dans ces cas l'orientation "l'emporte" sur la prescription. Elle ressemble, en effet, à une consultation avisée, dote le système d'un "signal" qui est consultable en cas de nécessité, d'ambiguïté et d'incertitude, l'empêchant de tourner à vide et de gâcher de l'énergie tout simplement parce qu'elle ignore quel sera le coup suivant (ce qui n'est pas déterminable a priori).

Donc, si une machine se préoccupe<sup>18</sup> pour des objectifs, il faut alors comprendre comment on peut représenter physiquement le signal d'un objectif dans un calculateur efficient, qu'il soit mécanique ou biologique. Pour Montague, la réponse réside dans l'évaluation, qui lie buts et normes d'orientation. Un philosophe dirait : le *jugement*. "42" est la réponse d'une machine qui ne sait pas juger, qui manque de critères d'évaluation pour juger ses prestations pendant son fonctionnement. Le modèle que Deep Thought incarne est, de plus, le modèle de fonctionnement d'un cerveau adulte ; né achevé – telle Minerve naissant de la tête de Zeus. Sans histoire et sans liberté. Sa réponse insensée est due tout simplement au fait que les questions

---

<sup>17</sup> De nombreux animaux sont aptes au sacrifice, mais ici, nous nous référons au terme sacrifices en tant que idéaux abstraits qui vont au-delà de la survie de l'individu ou du groupe.

<sup>18</sup> *Ibid.* p. 4, mais aussi p. 19 et 47-56.

sur la vie, sur l'univers et sur toute chose "ne l'intéressent pas". La théorie de Montague ne parvient pas encore à expliquer pourquoi nous, les hommes, au contraire, nous sommes intéressés. De même qu'elle ne parvient pas à expliquer le pouvoir "supérieur" que certains objectifs – même irraisonnables et extravagants – parviennent à dominer l'esprit et à agir eux-mêmes comme récompense. Elle n'explique pas non plus pourquoi, par exemple, chez les disciples de la secte religieuse américaine de Heaven's Gate, leur conviction absurde – qu'après leur mort, ils seraient transportés à bord d'un astronef pour atteindre "un niveau supérieur" d'existence – ait pu devenir une récompense primaire ; qu'un si fort "surdosage d'idées" l'ait emporté sur l'instinct de survivance chez ces disciples qui se suicidèrent en masse, dans une villa californienne, en 1977. Le thème du "pouvoir supérieur" fait appel à une explication physique sur la façon dont certaines idées, ou certaines problématiques, obtiennent une telle "force" qui déroutent les commandes de l'esprit humain, le portant à dépasser ses propres objectifs primaires et jusqu'à l'instinct de survie lui-même. Mais elle ne clarifie pas les mécanismes qui font que nous préférons une idée à une autre, « elle ne dit rien de ce qui fait que les idées sont en concurrence pour obtenir le statut de récompense »<sup>19</sup>. Il pourrait se faire que quand nous réussirons à dévoiler aussi ce mécanisme, nous serons en mesure également de formaliser l'idée même de liberté.

Certes, comme le reconnaît Montague lui-même, les modèles computationnels du fonctionnement de l'esprit et des mécanismes décisionnels en sont encore à leurs premières armes, mais ils présentent les avantages typiques des modèles mathématiques face aux descriptions "verbales" habituelles : ils permettent d'observer, en les quantifiant, les relations entre activités neurales et comportement, d'analyser la façon dont les systèmes d'évaluation "se détériorent" et les types de dysfonctionnement qui en découlent et consentent de fournir, par la dotation de comportements typiques des systèmes "biologiques", un nouveau cadre théorique capable de révolutionner la conception des machines et des objets et, donc, d'imiter la nature<sup>20</sup>. Ils ont peut-être aussi un avantage sur le plan philosophique : ils consentent, enfin, de donner un contenu empirique et contrôlable à des notions philosophiques fumeuses, comme "sens" et "âme", termes commodes que, dans bien des cas, nous donnons à notre ignorance.

(Traduction de l'Italien par Yolande THIERRY-ROSSETTI)

---

<sup>19</sup> *Ibid.*, p. 143.

<sup>20</sup> Je me réfère à la conception d'objets qui auraient, par exemple, un cycle de vie ou qui réagissent à l'environnement. Pour une discussion intéressante sur ce genre de conception et de design, voir Benyus (1997).

## **Bibliographie**

- Janine M. Benyus, *Biomimicry. Innovation inspired by nature*, Harper Perennial, New-York, 2007.
- Edoardo Boncinelli e Giulio Giorello(2008), *Lo scimmione intelligente*, Rizzoli, Milano, 2008.
- Antonio Damasio (1994) *Descartes' Error: Emotion, Reason, and the Human Brain*, Putnam, New York, 1994.
- Andrew Hodges (1983), *Alan Turing. The Enigma*, Simon&Schuster, New-York, 1983.
- Ignazio Licata (2008), *La logica aperta della mente*, Codice edizioni, Torino, 2008.
- Read Montague (2006), *Why choose this book? How we make decisions*, Dutton, New York, 2006.
- Alfredo Paternoster (2002), “La mente tra scienza e filosofia: il funzionalismo computazionale”, en P. Gentilini (s.d.), *Scienza cognitiva e rappresentazioni della razionalità*, IRRE-DPS, Genova, 2002.
- Hilary Putnam, «Minds and machines» (1960), en *Mind, Language and Reality*, vol. 2, Cambridge University Press, Cambridge, 1975, pp. 362-385.
- Rosenbluth A., Wiener N., Bigelow J., «Behavior, purpose and teleology», in *Philosophy of Science*, 10 (1) (1943), pp.18-24 (trad. Fr. «Comportement, intention, téléologie », *Les études philosophiques*, 2, pp. 147-196).
- Searle J. R. «Minds, brains, and programs», *Behavioral and Brain Sciences*, 3 (3) (1980), pp 417-457 (maintenant en ligne sur <http://www.cogsci.soton.ac.uk/bbs/Archive/bbs.searle2.html>)
- Norbert Wiener, *Cybernetics and Society*, Executive Techniques, New-York, 1951 (trad. Fr. *Cybernetique et société*, Deux Rives, Paris, 1971)