

CAHIERS FRANÇOIS VIÈTE

Série I – N°2

2001

Varia

EVELYNE BARBIN - *La courbe comme phénomène technique au XVIIe siècle*

GÉRARD CHAZAL - *De l'automate aux neurosciences*

ARMELLE DEBRU - *La preuve par l'expérimentation dans l'Antiquité*

COLETTE LE LAY - *JulesVerne, vulgarisateur de l'astronomie ?*

MICHEL MORANGE - *Un siècle de Génétique*

OLIVIER SAUZEREAU - *Nantes au temps de ses observatoires*

Centre François Viète
Épistémologie, histoire des sciences et des techniques
Université de Nantes

Achevé d'imprimer sur les presses de l'imprimerie centrale de l'université de Nantes, janvier 2002
Mise en ligne en juillet 2017 sur www.cfv.univ-nantes.fr

SOMMAIRE

- EVELYNE BARBIN 3
La courbe comme phénomène technique au XVIIe siècle
- GÉRARD CHAZAL 29
De l'automate aux neurosciences
- ARMELLE DEBRU 57
La preuve par l'expérimentation dans l'Antiquité
- COLETTE LE LAY 67
JulesVerne, vulgarisateur de l'astronomie ?
- MICHEL MORANGE 79
Un siècle de Génétique
- OLIVIER SAUZEREAU 91
Nantes au temps de ses observatoires

DE L'AUTOMATE AUX NEUROSCIENCES*

Gérard CHAZAL

Résumé

L'idée d'une machine imitant l'homme pour toutes ou partie de ses fonctions biologiques et mentales possède des racines mythiques très anciennes. Depuis le XVII^e siècle, avec le mécanisme, on assiste à une rationalisation de cette idée. Ce texte retrace rapidement l'histoire de cette rationalisation, qui est en même temps celle d'une conquête par la machine d'une part d'autonomie, avec l'introduction successive des notions de mémoires artificielles et de programme. On verra se croiser les lectures mécaniques, biologiques, anatomiques, physiologiques et psychologiques de l'automate jusqu'à la rencontre récente de l'ordinateur et du cerveau.

1. Introduction

L'informatique classique, celle de nos ordinateurs de bureau construits sur le modèle de von Neumann, est étroitement liée à la représentation langagière des informations, des données et des connaissances. Les traitements de l'information dont il est question portent sur des symboles fournis à la machine. Certes, le langage binaire qu'elle utilise est réduit aux plus simples et plus économes des moyens. La pensée philosophique classique qui a tant glosé sur le terme grec de *logos*, à la fois parole et raison, a lié de son côté, bien avant l'informatique, le langage à la pensée et par conséquent à l'intelligence comme à l'émotion et au sentiment qu'il s'agit d'exprimer. Le cartésianisme est à cet égard particulièrement significatif de cette équivalence implicite entre la pensée et le langage. De ce postulat, rarement remis en cause, il ressort deux conséquences manifestes par lesquelles il se confirme.

Premièrement toute expression d'un état mental doit relever du langage. Certes, il faudra pour cela forcer un peu le sens du mot ou plutôt l'étendre suffisamment pour qu'il englobe toutes les manifestations hu-

* Conférence donnée le 17 avril 2001 au Centre François Viète.

maines d'expression ce qui est d'ailleurs tout à fait discutable. Il y aura donc le langage du geste ou du corps, celui de l'image que le peintre exploite, et un langage des sons organisés, la musique. Le langage strictement défini par la parole et sa transcription scripturale ne sera qu'un cas particulier d'une expression langagière plus générale. On postulera même un langage mental. A la limite le langage devient l'équivalent de l'humain ou pour le moins en est la marque. Sur un tel fond on comprend que l'ordinateur, machine dotée d'un langage qu'elle nous emprunte, puisse jeter un trouble. D'où l'effort de tout un courant spiritualiste pour nier que l'ordinateur ait un véritable accès au langage. On accordera à la machine de n'accéder qu'à l'écume du langage, à la manipulation aveugle des signes dans lesquels il s'incarne.

Deuxièmement l'infralinguistique est alors rejeté définitivement et radicalement de l'aire de la pensée comme les animaux machines de la doctrine cartésienne. L'infralinguistique relève donc du machinal, de l'automatisme. On y retrouve le fou, le malade mental, l'aphasique, voire le vagabond dont la pensée ne peut pas se fixer dans des mots tout autant que son corps dans un espace¹. Les conduites répétitives et stéréotypées, les tics, le "machinal" sont les signes d'une perte de l'humain au profit du mécanique sans pensée. Il est alors évident que l'automate qu'est l'ordinateur ne peut relever que de la non-pensée, de la non-intelligence, d'une aveugle froideur minérale. Si l'on a vu dans l'ordinateur et ses programmes une image de l'esprit, comme en d'autres époques d'autres techniques ont pu fournir de telles images, ce n'est que vagues analogies, rhétoriques douteuses dont le philosophe doit se garder. On préférera dans ce cas les *Méditations* au *Traité de l'homme* ; on s'interrogera sur de possibles sous-entendus ontologiques dans les métaphores thermodynamiques, téléphoniques et photographiques des textes de Freud.

Cependant, on peut avancer quelques arguments de poids contre l'idée qu'il n'y aurait de pensée, et partant d'intelligence, de sentiment, d'émotions exprimables que de manière corrélative au langage ou à l'intérieur d'un système de signes. Par ailleurs, les développements récents de la neuroanatomie, de la neurophysiologie, d'une manière générale de la neurobiologie et des neurosciences, les développements concomitants de l'informatique connexionniste apportent des arguments supplémentaires contre ces thèses.

Si on refuse le postulat spiritualiste ou dualiste, il reste que toute forme de pensée est liée à l'existence physique de notre cerveau. Mais le cerveau est un organe que nous partageons avec beaucoup d'autres êtres

¹ Cf. J.-C. Beaune, *Le vagabond et la machine*, Champ Vallon, Seyssel, 1983 ainsi que *Les spectres mécaniques*, Champ Vallon, Seyssel, 1988.

vivants. Bien sûr les cerveaux n'ont pas tous atteint le même degré de complexité. Entre le cordon neural des animaux les plus simples et l'encéphale humain il existe des milliers de degrés. Pourtant, tous ces cerveaux sont composés de la même matière et du même type de cellules, les neurones. Ce qui diffère fondamentalement des uns aux autres c'est à la fois le nombre de neurones et le nombre de connexions qu'ils entretiennent entre eux. Si la pensée est liée au cerveau, ce qui est après tout une hypothèse tout aussi vraisemblable que celle d'une âme immatérielle, on doit pouvoir distinguer des degrés de pensée correspondant aux degrés de complexité des structures neuronales et non pas une coupure brutale entre les hommes doués de pensée et les animaux (ou les machines) qui en seraient définitivement et radicalement dépourvus. Dès lors le langage n'est plus le signe unique et univoque de la pensée. Il lui est effectivement lié ; mais s'il est une conséquence de la pensée il n'en est pas la condition nécessaire et suffisante. Il apparaît sous la forme articulée et même doublement articulée, telle que la décrivent les linguistes, à un certain niveau de développement du cerveau, de manière concomitante avec une certaine complexité. D'autres formes d'expression, l'art et le geste par exemple, peuvent alors être reconnues comme des expressions de la pensée, elles aussi liées à un haut de degré d'évolution de l'encéphale, sans pour autant être assimilées de gré ou de force au langage. Ceci du point de vue théorique est intéressant car il est toujours gênant d'élargir le sens des mots. À trop élargir l'extension d'un concept on finit par le vider de sens. En revanche, cela suppose que l'on ne trace plus une frontière absolue et essentielle entre le monde animal et l'homme dont le langage serait la marque. On substitue à cette philosophie de la rupture une philosophie de l'évolution qui devra élucider les mécanismes de complexification qui ont conduit de l'amibe à l'homme. Il reste encore la frontière entre l'inanimé et le vivant, déterminante pour l'acceptation ou non de l'idée d'une machine pensante. Toutefois, cette frontière n'est peut-être pas aussi précise que certains le croient.

Dans l'évolution des êtres vivants à partir de la matière inanimée le cerveau apparaît comme une solution particulièrement favorable aux individus et partant aux espèces qui s'en trouvent dotés. Il permet une meilleure adaptation du comportement à l'environnement. La pensée et corrélativement l'intelligence, l'émotion, le sentiment et l'ensemble des états mentaux sont caractéristiques d'une adaptation particulièrement fine de l'homme à son environnement. Il peut donc exister une pensée infralinguistique que l'on a déjà décelée chez les animaux à différents niveaux de l'évolution mais aussi chez l'enfant humain dans ses premiers mois d'existence alors qu'il ne possède pas encore le langage et que parallèlement dans leur ensemble les connexions entre ses neurones ne sont

pas mises en place. Si l'on ne peut pas nier le caractère d'automatisme de certains comportements infralinguistiques, il faut alors concevoir autrement l'automate. Il devient une forme de pensée, certes d'un degré moindre sur une échelle évolutive, mais une pensée tout de même, une pensée capable, dans une certaine mesure, d'intelligence et d'adaptabilité d'un comportement aux changements de l'environnement.

Cette perspective soulève toutefois quelques questions importantes que j'essaierai d'éclaircir. Premièrement si le cerveau porte la pensée ou si celle-ci dans ces différentes formes relève du fonctionnement cérébral, il faut qu'il représente d'une manière ou d'une autre le monde auquel il adapte le comportement des individus d'une espèce donnée. Comment représente-t-il le monde ? Les neurosciences sont encore loin d'apporter une réponse complète, cependant des avancées importantes ont été faites dont le philosophe ne peut pas manquer de s'instruire. On se demandera aussi, évidemment, dans quelle mesure l'informatique connexionniste qui tente de reproduire la structure matérielle du cerveau et son fonctionnement peut apporter ses propres lumières. Mais il ne suffit pas d'en rester à l'exposé de ce qui va dans le sens d'une thèse, il est encore nécessaire d'examiner ce qui lui fait obstacle et les objections qu'on peut lui apporter. Ainsi en s'appuyant à la fois sur les neurosciences et sur l'informatique connexionniste il faudra tenter de prendre en compte les problèmes soulevés par l'intentionnalité, les *qualia*, c'est-à-dire la qualité subjective de nos états internes, l'introspection et tout ce qui dans notre vie mentale semble échapper à une description en termes neurologiques.

2. Rappels historiques

De l'homme à la machine et inversement de la machine à l'homme, inévitablement l'informaticien est conduit à s'intéresser à la biologie, à l'anatomie, à la physiologie du cerveau et d'une manière générale du système nerveux. Cet intérêt est manifeste dès les débuts de l'informatique comme en témoignent les écrits de von Neumann, en particulier *Théorie générale et logique des automates*². Cependant, les travaux qui tentent de rapprocher l'informatique des neurosciences s'inscrivent dans une longue tradition qui partant au plus près de nous de Descartes parcourt toute la pensée occidentale à travers les automates programmés de Vaucanson et les métaphores technologiques diverses appliquées à l'esprit. Avant de m'arrêter sur les acquis proprement dits

² J. Von Neumann, *Théorie générale et logique des automates*, Champ Val-lon, Seyssel, 1996.

des sciences biologiques je voudrais rappeler cette tradition qui, comme rêve métaphysique de l'homme-machine, justifie encore cette démarche.

Entre mythe et réalité, les automates, les machines imitant le vivant, constituent un intermonde entre le mécanique et l'organique, entre l'inanimé et l'animé, voire entre le matériel et le spirituel. La philosophie ne se lasse pas d'explorer cet intermonde, d'en tenter une rationalisation périlleuse toujours aux limites de la raison et de la déraison. Il est vrai que la machine prolonge toujours quelques-unes de nos fonctions avec une certaine autonomie. Plus loin de nous que l'outil qui se modèle sur notre corps et s'y rattache par ses formes, la machine en même temps nous ressemble et nous échappe. Elle nous ressemble non seulement parce qu'elle vient de nous, de notre activité industrielle, mais aussi parce qu'elle est destinée à assurer des fonctions qui sont les nôtres sur des objets qui dépassent nos forces. Elle nous échappe puisque nous la dotons, contrairement à l'outil qui n'est actif que par notre constante présence, d'une part d'autonomie. Cette autonomie tient d'abord à ce qu'elle possède une source d'énergie qui n'est plus celle de notre propre musculature et d'autre part au fait que l'on en viendra très vite à la programmer. Entendons par cette dernière expression le simple fait que la machine devra être capable d'enchaîner une suite d'actions prévues sans qu'il soit nécessaire qu'un acteur humain les déclenche une à une. Cette ambivalence de la machine est le terrain à la fois des rêveries et des mythes mais aussi de l'entreprise de rationalisation du mythe qui aboutit à la technologie. Il peut paraître étrange de concevoir la technologie et les techniques les plus modernes comme rationalisation du rêve et du mythe ; pourtant un rapide aperçu historique suffit pour s'en convaincre.

A l'origine même de notre culture occidentale, chez Homère (*Iliade*), Héphaïstos la divinité industrielle construit des trépieds et des jeunes filles capables de se mouvoir de manière autonome. Notons bien ici qu'il ne s'agit pas de la création divine du vivant par un dieu couvrant quelque matière de son souffle. Nous avons bien affaire à une fabrication par le dieu maître du feu et des arts de la forge. Et si ses œuvres ne sont plus simplement de la matière inanimée ce ne sont pas non plus de véritables êtres vivants. L'archéologie nous a livré des jouets articulés de l'antiquité et Aristote rapporte que Dédale avait construit une Vénus mobile en bois ainsi qu'une fausse vache en bois et cuir qui permit à Pasiphaé de recevoir le taureau dont elle était amoureuse. L'ingénieur et sa métis sont partie prenante des mythes originels. A l'aube de la pensée occidentale nous sommes entre mythe et réalité. Est-ce à partir de tels jouets que naquit la légende selon laquelle les forgerons crétois fabriquaient des statues que l'on devait enchaîner la nuit de crainte qu'elles ne sèment le désordre dans la ville ? Le rêve d'une vie artificielle que réalise

à sa façon l'informatique moderne prend ses racines très tôt dans le mythe. Les mythes, comme encore celui de Galatée, sont à l'origine d'une véritable rêverie métaphysique sur la nature humaine proprement dite comme si l'artificiel permettait de pénétrer le naturel.

Le Moyen Age peuple ses horloges d'automates et les jacquemarts se répandent à travers toute l'Europe. Les ingénieurs de la Renaissance élaborent les outils mécaniques qui inspireront l'une des premières philosophies mécanistes comme tentative de penser les frontières de la machine et du vivant : celle de Descartes. Pour sauver le spirituel le dualisme doit rejeter vigoureusement le corps du côté du mécanique et de l'automatisme. Une fois faite la part irréductible à l'esprit le corps et ses mouvements, les plus visibles comme les plus souterrains, peuvent faire l'objet d'une enquête mécaniste. Le *Traité de l'homme* double les *Méditations*. Cet ouvrage commence comme une fiction :

"Ces hommes seront composés, comme nous, d'une Ame et d'un Corps. Et il faut que je vous décrive, premièrement, le corps à part, puis après l'âme aussi à part ; et enfin, que je vous montre comment ces deux natures doivent être jointes et unies, pour composer des hommes qui nous ressemblent. Je suppose que le corps n'est autre chose qu'une statue ou machine de terre, que Dieu forme tout exprès, pour la rendre plus semblable à nous qu'il est possible : en sorte que, non seulement il lui donne au dehors la couleur et la figure de tous nos membres, mais aussi qu'il met au-dedans toutes les pièces qui sont requises pour faire qu'elle marche, qu'elle mange, qu'elle respire, et enfin qu'elle imite toutes celles de nos fonctions qui peuvent être imaginées procéder de la matière, et ne dépendre que de la disposition des organes."³

Descartes entreprend une description anatomique et physiologique de l'homme mais d'emblée cette description est placée sous le signe de l'artificialité. Il sera alors fait appel à toutes les ressources technologiques dont on peut disposer : leviers, cordes, tuyaux, pompes, soufflets, fontaines, etc. Je ne m'arrêterai pas aux détails d'une description aujourd'hui largement dépassée. Peu importe. Mais je retiendrai la démarche, démarche à double sens. Premièrement la compréhension de l'homme passe par l'artificiel, par le reconstruit (même sur un mode fictif). François Dagognet a particulièrement insisté sur cette méthode et sa richesse en commentant l'œuvre d'Etienne Jules Marey. Deuxièmement l'artificiel se comprend à partir d'un retour au biologique, à

³ Descartes, *De l'Homme, Œuvres philosophiques*, Garnier Flammarion, (éd. F. Alquié), Tome I, p. 379.

l'anatomie et à la physiologie. Descartes inaugure une démarche qui est encore celle des sciences cognitives aujourd'hui ou plutôt celle de la cybernétique. La rationalisation du mythe initial est à ce prix. Une lecture attentive de Descartes montre alors clairement que le mécanisme ne consiste pas tant à réduire le vivant à l'inanimé et au physique qu'à effectuer un va et vient enrichissant entre les deux domaines trop souvent donnés comme deux règnes séparés. Le mécanisme cartésien est tout autant un mouvement vers le biologique, ce que l'on a parfois tendance à oublier, qu'une restriction du corps au domaine de la physique. C'est finalement ce même mouvement que reprend aujourd'hui l'informaticien et que le philosophe doit accompagner.

Cependant, la rationalisation cartésienne ne met pas fin au mythe. Elle a probablement trop séparé l'esprit et le corps et leur union devient un problème que la raison ne parvient plus à résoudre. Elle échappe constamment et de toute part à une claire rationalité. Soumis aux questions pressantes de la princesse Elisabeth, Descartes devra dans une lettre célèbre en rejeter la connaissance à une simple intuition existentielle :

“Et les pensées métaphysiques, qui exercent l'entendement pur, servent à nous rendre la notion de l'âme familière ; et l'étude des mathématiques, qui exercent principalement l'imagination en la considération des figures et des mouvements, nous accoutume à former des notions du corps bien distinctes ; et enfin, c'est en usant seulement de la vie et des conversations ordinaires, et en s'abstenant de méditer et d'étudier aux choses qui exercent l'imagination, qu'on apprend à concevoir l'union de l'âme et du corps”⁴.

Descartes renvoie alors au vécu laissant un espace libre pour que se perpétuent les mythes de l'automate. Le terrain est ouvert à de multiples tentatives pour réaliser des machines de plus en plus automatiques et mimant de plus en plus fidèlement le vivant. Le cartésianisme ne l'interdit pas, bien au contraire, finalement le *Traité de l'Homme* l'autorise et y invite.

Le XVIII^e siècle sera celui des automates. Evidemment on pense d'abord à Vaucanson (1709-1782) construisant entre autres un joueur de flûte et un canard célèbre capable d'absorber et digérer de la nourriture. Mais les mécaniciens suisses Pierre Jaquet-Droz (1721-1782) et son fils Henri-Louis construiront un dessinateur et une musicienne automates. On n'aura garde d'oublier Kempelen et Mailzel et le célèbre joueur d'échec qui, s'il cache une supercherie, ne révèle pas moins cette idée dominante qu'une machine peut reproduire des fonctions du vivant y

⁴ Ouvrage cité, Tome III, pp. 44-45.

compris les facultés habituellement attribuées à l'esprit humain (dessiner, jouer de la musique, jouer aux échecs). Les développements de l'automatisme suivent une route parallèle à celle des avancées de l'anatomie même si après le *Traité de l'Homme* de Descartes les deux disciplines ne se croisent plus autant. En fait, les XVII^e et XVIII^e siècles mettent en avant plusieurs idées que nous retrouvons encore dans l'informatique et l'intelligence artificielle, idées qui suffisent à justifier un détour de l'informaticien par les neurosciences.

Plusieurs perspectives accompagnent l'attention particulière que le siècle des lumières accorde aux automates. Premièrement, accroître l'autonomie de la machine est une finalité technique donnant lieu à une réalisation progressive. On peut relever deux étapes importantes dans cette autonomisation des dispositifs mécaniques. La première est celle de la mise en œuvre de la notion de programmation. L'autonomie de la machine, c'est-à-dire le fait qu'elle puisse accomplir une suite de mouvements de manière déterminée sans qu'intervienne un agent humain, si ce n'est au moment initial de la mise en route, tient à l'existence d'un programme. Chez l'automate de Vaucanson ou ceux des Jacquet-Droz le programme est immanent à la structure mécanique, il est dans la disposition des rouages, des arbres à cames, des poulies et des leviers. De ce fait il garde une fixité quasi absolue. On ne peut le modifier qu'au prix d'une reconstruction presque totale du mécanisme. On gagnera en souplesse à partir du moment où le programme sera séparé du dispositif physique qu'il anime. Cela suppose une mémoire artificielle sur laquelle ce programme peut être stocké et à partir de laquelle la machine est guidée. L'invention paradigmatique en ce domaine est celle de la carte perforée que le métier à tisser de Jacquard va utiliser. Joseph Marie Jacquard, mécanicien lyonnais, s'inspirant des travaux antérieurs de Vaucanson, de Bouchon et de Falcon met définitivement au point un système programmé de sélection des fils sur les métiers. La machine réalise alors un motif au tissage entièrement commandé par un programme inscrit sur la mémoire artificielle des cartes perforées. On met en avant le métier Jacquard par l'importance industrielle et sociale que son introduction aura (on sait que le chômage provoqué par cette introduction conduira les canuts lyonnais lors de leur révolte de 1831 à briser ces machines), mais les dispositifs à cartes perforées se retrouveront dans de nombreuses machines : automates androïdes, orgues de barbarie et boîtes à musique de toutes sortes jusqu'à la *machine* de Charles Babbage qui préfigure le moderne ordinateur. Le comportement de la machine est modifié lorsqu'on change son programme. On peut ainsi, par simple échange des cartes perforées, modifier les motifs du tissu ou la mélodie que joue l'orgue de barbarie. Cette séparation qui recouvre l'actuelle distinction entre mat-

ériel et logiciel ouvre la voie à la souplesse et à l'adaptabilité de la machine à des tâches diverses.

La seconde étape dans la conquête de l'autonomie par des dispositifs mécaniques tient à la possibilité que l'action de la machine agisse en retour sur elle-même. Très tôt de tels dispositifs ont été conçus. J'en citerai deux particulièrement significatifs. L'histoire raconte — peu importe d'ailleurs la véracité des faits rapportés —, que Humphry Potter inventa la distribution automatique de la vapeur pour les machines de Newcomen alors qu'il était affecté à la surveillance d'une de ces machines. Son travail consistait à manœuvrer successivement deux robinets mettant le cylindre en communication alternativement avec la chaudière et avec le bac d'eau, c'est-à-dire la source chaude et la source froide. L'idée de Potter fut de faire opérer l'ouverture et la fermeture de ces robinets par des leviers commandés par les mouvements mêmes du piston. Il crée ainsi une boucle par laquelle la machine intervient sur son propre fonctionnement. De ce fait la tâche qu'il devait assumer est transférée à la machine. Le second exemple d'un tel dispositif est celui du régulateur à boules dont s'équipe la machine de Watt. La soupape tourne en fonction de la pression à l'intérieur de la chaudière. Elle tourne d'autant plus vite que la pression est élevée. Cependant plus elle tourne vite plus les poids qu'elle entraîne s'écartent. Le parallélogramme articulé qui porte ces poids se déforme et ouvre d'autant la soupape. Mais aussi plus le dispositif tourne vite plus la soupape s'ouvre et par suite la pression dans la chaudière décroît. La baisse de cette pression provoque un ralentissement ; les poids, en vertu de la pesanteur, redescendent ; la soupape tourne plus lentement et se referme. La pression croît à nouveau et un nouveau cycle s'amorce de sorte que la pression se maintient autour d'une valeur déterminée. De cette manière la pression dans la chaudière acquiert une certaine constance moyennant de petites variations dans les deux sens autour d'une valeur fixe. Ainsi, la régulation est commandée par la pression elle-même. Nous avons là une des premières machines cybernétiques.

La troisième étape dans la conquête de l'autonomie machinique est plus récente et étroitement liée à l'informatique elle-même, même si elle découle en droite ligne des systèmes d'autorégulation que je viens d'évoquer. Il s'agit de rendre la machine capable d'adapter son comportement aux changements du milieu dans lequel elle agit. Cela suppose qu'elle soit dotée de capteurs de toutes sortes qui lui fournissent une information sur son environnement. Toute la robotique moderne a développé ce type de dispositifs. Caméras, capteurs de température, de pression et de position, l'idée fondamentale est que l'autonomie de la machine passe par le fait qu'elle est dotée d'un dispositif sensoriel à l'image

de notre corps avec ses yeux, ses oreilles, sa peau, l'ensemble des sens. L'idée que pour agir de manière autonome une machine doit être équipée d'organes sensoriels est dans le fond assez triviale. Sa réalisation s'est pourtant avérée beaucoup plus difficile qu'il n'y paraît au premier abord. Elle implique que nous ayons au préalable une connaissance déjà suffisante de la manière dont notre propre système nerveux transforme en actes les messages que les sens lui fournissent sur le monde extérieur. Et bien que cette connaissance soit encore extrêmement partielle nous retrouvons ici la démarche cartésienne telle qu'elle se déploie dans le *Traité de l'Homme*. Cependant les connaissances anatomiques et physiologiques du XVII^e siècle ne pouvaient pas permettre d'aller techniquement très loin même si un cadre philosophique à une telle entreprise pouvait d'ores et déjà s'élaborer.

La seconde idée qui accompagne les développements de l'automate depuis le XVII^e siècle est celle d'une lecture mécaniste du corps qui peu à peu contamine celle de l'esprit. Au départ cette lecture mécaniste est fondée chez Descartes sur la séparation entre la substance étendue et la substance pensante. On connaît les difficultés liées à un tel dualisme dans la mesure où l'on tente de le radicaliser. La principale de ces difficultés est celle qui consiste à expliquer comment deux substances radicalement différentes peuvent interagir. Or nous avons une expérience quotidienne d'une telle interaction. Déjà Descartes doit fixer les limites de ce qui tient de l'étendue et s'explique en termes mécaniques de ce qui tient de l'esprit. La tâche est difficile. Les émotions, par exemple, relèvent en partie d'une véritable réflexologie fondée sur l'hypothèse des *esprits animaux*. On le voit, la frontière est particulièrement difficile à définir. Les humeurs et les inclinations tiennent leurs caractéristiques de la nature physique des esprits animaux. La lecture mécaniste du corps appuyée sur l'anatomie et la physiologie perpétuellement éclairées par des métaphores techniques (leviers, poulies, tubes, circulation de fluides...) glisse sans arrêt vers une même lecture de l'esprit. La pensée cartésienne est une pensée sur le fil du rasoir; la métaphysique de l'automate, le va et vient constant entre l'anatomie et la métaphore machinique maintiennent une tension constante. C'est de cette tension que se libère La Mettrie en concevant l'*Homme machine*. Evacuée l'hypothèse d'une substance pensante séparée, le mécanisme règne de plein droit. Mais il serait faux de croire qu'il s'agit là du seul matérialisme possible. On peut réduire le vivant et partant la pensée au mécanisme mais on peut tout aussi bien opérer la réduction inverse. Alors l'inanimée n'est plus qu'un cas limite du vivant. C'est ce vitalisme que développera une philosophie comme celle de Leibniz où toute matérialité se trouve réduite au dénominateur commun de la monade et où la

moindre parcelle de matière recèle une infinité de substances. L'inanimé n'existe plus. Il n'est que le réceptacle d'une vie latente, infime certes, mais de nature identique à celle que nous accordons aux organismes supérieurs. Diderot généralisera cette pensée du vivant en un autre matérialisme, tel que l'inerte doit être pensé avec les catégories du vivant. On le voit, la pensée cartésienne est une pensée en équilibre — c'est peut-être là sa grandeur.

Cette problématique autour de l'automate donne lieu à diverses manifestations.

Premièrement, les lectures biologiques, anatomiques et physiologiques d'une part, la lecture psychologique d'autre part, se croisent sans cesse autour de la notion d'automate. L'artefact physique du corps débouche sur l'artefact de l'esprit.

Deuxièmement, le flou de la frontière entre le corps et l'esprit que l'automate instaure implique que le mécanisme se voile, se cache comme pour préserver une part de mystère. Roentgen présenta à Louis XVI une joueuse de tympanon. La poupée cachait son mécanisme sous des vêtements, mais la rumeur voulait que son corps fût merveilleusement modelé sous la robe brodée. D'un côté une entreprise d'explication, de l'autre le voile qui ménage le rêve ! Mécanismes et fantasmes font bon ménage. Ainsi on est toujours entre doter la machine des pouvoirs peu ou prou mystérieux de la vie et expliquer l'homme par la machine donc accepter de nous percevoir comme première machine.

La troisième idée, liée aux automates mais aussi aux développements de la logique inaugurés essentiellement par Leibniz, est que toute pensée peut se ramener à une forme de calcul. *Penser c'est calculer* — déjà Hobbes l'affirmait —, or le calcul est mécanisable. Les développements de la logique dans la première moitié du XX^e siècle le montreront à travers les travaux de Turing et de Church. Schikard, Pascal et enfin Leibniz ont réalisé des machines à calculer mécaniques. Faut-il en conclure que la puissance de la pensée est transposable à une machine ? Pour que cette question soit explicitement formulée et discutée il faudra attendre l'invention de l'ordinateur et le célèbre article de Turing. Cependant d'autres questions précèdent celle que l'informatique posera de manière radicale, questions liées à l'assimilation de la pensée à un calcul. De quelle sorte de calcul peuvent bien relever les pensées non opératoires au sens mathématique ordinaire du terme opération ? Comment l'esprit ou le cerveau peuvent-ils effectuer ces calculs ? La connaissance sera-t-elle en tout domaine, comme elle peut l'être en mathématique, le résultat de procédures de calcul ?

On voit parfaitement que ces questions sont déjà au cœur de la philosophie leibnizienne, mais elles se retrouveront du XVII^e siècle jus-

qu'à nos jours chez beaucoup de penseurs et, bien sûr, à commencer par les logiciens. Supposer que penser équivaut à calculer implique que l'on se donne une définition de la notion de calcul plus vaste que ce que recouvre le sens arithmétique du terme. Les auteurs du XVII^e (Descartes et Leibniz au premier chef) portaient déjà l'idée d'étendre la notion de calcul. Cette idée s'enracine dans le développement du calcul symbolique et de l'algèbre qui substituent aux nombres des caractères capables de représenter soit un nombre particulier, mais non précisé, soit une variable, c'est-à-dire un nombre pouvant varier suivant certaines règles. Elle repose aussi sur la possibilité instaurée principalement par Descartes de donner une formulation algébrique des formes géométriques et donc d'introduire le calcul dans le domaine de l'espace. Le développement de la géométrie analytique montrera le pouvoir des procédures de calcul appliquées à des objets autres que numériques. Enfin, la physique quantitative, après Galilée, offre le spectacle d'un pouvoir prédictif de la science s'appuyant sur le calcul suite à une activité de mesure qui ramène les phénomènes les plus divers à des valeurs numériques. D'une certaine manière les mathématiques appliquées à partir du XVII^e siècle retrouvent une conception algorithmique qui vient doubler la conception intuitive et contemplative des formes qu'avait promue la naissance grecque de la géométrie. Or l'algorithme n'est rien d'autre que l'organisation rationnelle et efficace d'une suite de calculs.

Avec les deux grands logiciens du XIX^e siècle, Georges Boole (1815-1864) et Augustus De Morgan (1806-1871), la logique devient explicitement un calcul et ce calcul est présenté comme décrivant les lois de la pensée. Ainsi Boole écrit *L'analyse mathématique de la logique* (1847) et *Recherches sur les lois de la pensée* (1854) qui présente les termes ET, OU, SI... ALORS, empruntés au langage naturel, comme des opérateurs et le calcul des propositions, tel que nous l'avait légué l'antique logique stoïcienne, comme une algèbre; Augustus De Morgan (1806-1871) dans *Formal Logic or the Calculus of Inference* (1847) contribue pour sa part à l'intégration de la logique dans le corps des mathématiques. Cette œuvre d'extension de la notion de calcul sera poursuivie aussi bien par Charles Sanders Peirce (1839-1914) que par Gottlob Frege (1848-1925), considérés à juste titre comme les pères de la logique moderne.

Pour partie l'ordinateur classique, le modèle de von Neumann, organisé autour d'une unité arithmétique et logique est l'héritier direct de cette conception de la pensée comme calcul. La question demeure cependant de savoir si notre esprit pense comme l'ordinateur calcule. L'apparition de l'ordinateur soulève d'ailleurs deux sortes de problèmes. Premièrement, si nous pensons comme l'ordinateur calcule, cela peut

vouloir dire que la pensée est un processus qui peut être implanté sur n'importe quelle structure physique adéquatement organisée ; on rejoint par-là les thèses fonctionnalistes. Cela peut vouloir dire aussi, en se projetant à la limite, que la pensée relève d'un automatisme qui peut être dégagé de tout support matériel et dans ce cas c'est la notion d'automate intellectuel tel que Spinoza ou Leibniz la présentent qui revient au devant de la scène philosophique.

3. L'ordinateur et le cerveau

Cependant, très vite il apparut que ce que faisait l'ordinateur classique était très différent de ce que nous réalisons nous-mêmes, tant dans la façon de faire que dans le résultat. L'ordinateur s'avère très performant en matière de calcul arithmétique, beaucoup plus que nous. Il peut effectuer des millions de calculs élémentaires à la seconde et, par combinaisons de ces opérations élémentaires, obtenir rapidement le résultat d'opérations très complexes. Il peut, par exemple, calculer des milliers de décimales de π là où il faudrait un temps considérable à un opérateur humain. De plus, sous réserves que l'on prenne certaines précautions, les calculs de la machine sont à l'abri des erreurs d'inattention et de mémoire dont les hommes souffrent si facilement. En revanche, l'ordinateur s'avère particulièrement inefficace pour des tâches que nous effectuons rapidement et aisément comme, par exemple, reconnaître un visage parmi des milliers d'autres. Si penser c'est calculer, il faut reconnaître que le calcul qui constitue notre pensée est différent de celui que met en œuvre l'unité arithmétique et logique d'un ordinateur. D'où les réserves que l'on peut formuler quant à la thèse fonctionnaliste selon laquelle le support serait indifférent pourvu que soit conservée la fonction. (Cf. Pitrat "L'intelligence dépend des caractéristiques de son support matériel" dans *Intellectica* n° 21, 1995, pp. 265-268) On retrouve là une raison fondamentale qui pousse l'informaticien, à commencer par le premier d'entre eux, c'est-à-dire John von Neumann, à se tourner vers les neurosciences. Si on met entre parenthèses l'hypothèse d'un pur esprit qui effectuerait des calculs d'une nature tout à fait inconnue et si on suppose que notre cerveau est l'organe de la pensée, il devient évident que l'informaticien, s'il veut étendre la pensée mécanique à autre chose que les calculs arithmétiques et logiques, doit demander au neurobiologiste des informations sur la structure et le fonctionnement du cerveau.

On comprend ainsi que tout conduit le philosophe qui tente de confronter sa réflexion aux techniques informatiques à interroger le

vivant et en particulier le système nerveux. Que l'on considère la conquête de l'autonomie par la machine, la lecture mécaniste du corps parallèle au développement des automates de toutes sortes ou encore l'assimilation de la pensée à une forme de calcul, on est conduit à comparer l'artefact et le vivant, à faire en sorte que la réflexion effectuée un va et vient de l'un à l'autre et entre dans le jeu de miroir qui s'instaure. En développant l'artefact au-delà de tout ce qui avait été fait jusque-là l'informatique nous oblige à reposer en termes nouveaux la question de la nature de la pensée.

Si l'on compare la pensée et le cerveau à l'ordinateur classique, on est inmanquablement conduit à dissocier cerveau et pensée, car si l'ordinateur accomplit bien certaines fonctions qui appartiennent chez l'homme au domaine de la pensée, l'organisation matérielle cérébrale est totalement différente de celle de la machine. De ce fait ou bien on fait le postulat que l'ordinateur est une machine capable — ou qui deviendra capable — de penser et alors on doit négliger le substrat matériel des fonctions mentales, ou bien en mettant en avant ce substrat matériel on refusera toute possibilité à la machine de penser. Ce problème est parfaitement perçu par John von Neumann lui-même dès les débuts de l'informatique comme en témoigne son article *Théorie générale et logique des automates* paru en 1951. Partant des travaux de Rosenblatt (le *Perceptron*)⁵, de McCulloch et Pitts⁶, de Hebb, il comprend que l'informaticien doit se tourner vers la biologie et la neurologie s'il veut construire une machine capable de penser comme pense le cerveau. On voit bien qu'il y a dans cette démarche qui va de l'informatique naissante vers les neurosciences le postulat implicite que nous sommes des automates. Certes, nous sommes beaucoup plus complexes que tous les automates que notre industrie a jusqu'ici pu construire, mais des automates tout de même. Cela signifie qu'il doit être possible d'expliquer la pensée, les états mentaux et à terme la conscience elle-même ainsi que toute signification sous-jacente en termes d'automatisme. Il s'agit bien dans le fond de revenir au mécanisme cartésien et de lui donner une extension nouvelle. L'informaticien reprend une démarche déjà ancienne mais dans un contexte technique et technologique très différent. Il va se développer ce que nous pourrions appeler un mécanisme algorithmique.

Beaucoup d'arguments ont été avancés pour contester l'idée que la pensée sous ses diverses formes serait explicable par l'anatomie et la

⁵ F. Rosenblatt, *Principles of Neurodynamics*, Spartan Books, New York, 1959.

⁶ W. McCulloch et W. Pitts, "A logical calculus of the ideas immanent in nervous system", *Bulletin of Mathematical Biophysics*, 5, 1943, pp. 115-133.

physiologie du système nerveux central. Or, il est évident que si ces arguments ont une valeur probante il est vain de chercher dans l'ordinateur un modèle de la pensée, de vouloir expliquer celle-ci par celui-là ou encore de vouloir construire une machine pensante. J'aborderai trois arguments parmi les plus fréquemment proposés contre l'idée d'une équivalence entre les états cérébraux et les états mentaux, autrement dit entre le corps et l'esprit.

4. Les arguments contre l'équivalence entre états cérébraux et états mentaux

Le premier argument est lié au fait que nous sommes conscients non seulement du monde extérieur mais aussi de nous-mêmes et que nous disposons de la faculté d'introspection qui nous rend présents à nous-mêmes. Or, ce que nous livre l'introspection ce sont des états que nous avons appris à caractériser en termes mentaux et psychologiques ou en termes d'intentionnalité. Nous pouvons dire que nous éprouvons une douleur ou un chagrin, une joie ou un plaisir, que nous croyons en telle ou telle chose, que nous en désirons telle autre, que nous pensons à telle personne, et ainsi de tous nos états mentaux et de toutes les attitudes propositionnelles. Mais en aucun cas nous n'avons conscience de l'état de nos circuits neuronaux. Cela ne veut pas dire que nous n'ayons pas conscience de manière plus ou moins nette de notre corps. La douleur est bien liée au corps. De même nous avons parfaitement conscience d'un certain nombre de sensations kinesthésiques. Je sais très bien comment sont placés mes membres, si mon corps est assis ou allongé. J'ai une sensation globale de mon corps, de sa position dans l'espace ou encore des dommages qu'il peut subir mais non des mécanismes qui me fournissent cette sensation ou cette connaissance. Si je peux ressentir le froid ou le chaud, la douleur d'une piqûre ou le plaisir d'une caresse, les mécanismes de transmission de l'influx nerveux des récepteurs de la température ou de la douleur aux différents centres cérébraux demeurent hors d'accès de ma conscience et de ma connaissance immédiate. Sans nier le rôle que le système nerveux joue dans ces différents phénomènes liés au corps, certains ont pu tirer de ce qui précède la conséquence qu'il y aurait une différence radicale de nature entre les phénomènes neuronaux et ce que nous livre l'introspection. De là la nécessité de postuler l'existence de quelque chose d'autre, au-delà ou au-dessus du fonctionnement neuronal, que l'on appellera l'esprit ou l'âme ou encore de manière plus vague le mental. Cet argument est renforcé par le fait que s'il est assez aisé dans l'état actuel de la science de faire correspondre à certaines choses, comme les

diverses sensations, des états du système nerveux, la science ne nous a pas révélé de correspondance entre le système nerveux et certains états d'âme comme la joie ou la mélancolie. Autrement dit, le jeu de correspondances entre les états internes tels qu'ils apparaissent à la conscience et à l'introspection est encore loin d'être parfaitement élucidé. La science se découpe suivant des niveaux (neurosciences, psychologie cognitive, psychiatrie, psychanalyse, théorie de la connaissance, sciences cognitives, intelligence artificielle), parfois concurrentiels, que l'épistémologue a bien du mal à articuler entre eux.

Réciproquement, et c'est là le second argument, l'examen du cerveau — qui a pourtant beaucoup progressé avec les techniques modernes de l'électroencéphalogramme, du scanner ou de l'imagerie par résonance magnétique —, ne nous révèle pas la nature qualitative et subjective de nos états mentaux, de nos états d'âme ou de nos intentions. On s'est appuyé sur ce fait pour nier qu'il y ait équivalence entre les états mentaux en général et des états physiologiques ou encore que les premiers puissent se réduire au second⁷. Frank Jackson illustre cet argument en imaginant une neurologue, Mary, qui connaîtrait parfaitement tout ce qui peut être connu du fonctionnement du cerveau mais qui n'aurait aucune expérience des couleurs soit parce qu'elle serait privée des cellules rétinienne sensibles aux couleurs soit parce qu'elle souffrirait d'une quelconque maladie ayant le même effet, soit encore parce qu'elle aurait toujours vécu dans un univers monochrome. Mary sait tout du fonctionnement neuronal, mais ignore tout de ce qu'est éprouver une sensation de rouge ou de vert. Il existe donc une connaissance qui relève du vécu, complètement irréductible à celle que nous fournissent les neurosciences.

Le troisième argument met en avant le caractère spatio-temporel du cerveau et du système nerveux. Les ondes cérébrales révélées par l'électroencéphalogramme se développent dans le temps et sur des chemins repérables dans la matière nerveuse. On peut associer à tous les phénomènes nerveux une localisation dans l'espace et dans le temps. Si

⁷ Ce débat est particulièrement présent dans la philosophie anglo-saxonne. Nous pouvons citer entre autres : Thomas Nagel, "What is it like to be a bat?", *Philosophical Review*, 83, n°4, 1974, pp. 435-450. Ce texte a été traduit dans D. Hofstadter et D. Dennett, *Vues de l'Esprit*, InterEditions, Paris, 1987, pp. 391-403. ; Frank Jackson, "Epiphenomenal Qualia", *Philosophical Quarterly*, 32, n°127, 1982, pp. 127-136 ainsi que "What Mary didn't Know", *Journal of Philosophy*, 83, n°5, 1986, pp. 291-295. ; Paul Churchland, "Reduction, qualia, and the Direct Introspection of Brain States", *Journal of Philosophy*, 82, n°1, 1985, repris dans *A Neurocomputational Perspective*, MIT Press, 1985, pp. 47-66. ; on lira dans le même ouvrage le chapitre 4, "Knowing Qualia : a Reply to Jackson", pp. 67-76.

on peut bien encore situer dans le temps une émotion ou un sentiment, il semble absurde de vouloir leur donner une situation spatiale. Ainsi que le remarquait déjà Descartes, l'esprit comme substance pensante s'oppose à la substance étendue. L'âme n'occupe pas une place dans l'espace, elle n'a ni masse ni aucune des caractéristiques de la matière. Mon chagrin n'est ni ici ni là, ni long ni large et si je peux le qualifier de profond ce n'est que par métaphore. On pourra effectuer des mesures de l'influx nerveux, relever des fréquences, les traduire par des graphiques comme le montrent les ouvrages de physiologie, on pourra suivre sa propagation dans le réseau neuronal. Comment pourrait-on mesurer un sentiment, une croyance, un quelconque état mental ? On pourra bien parler de l'intensité de ces états, mais là encore c'est le plus souvent une façon métaphorique de parler. Les cartes cérébrales que dressent les neurosciences en tentant de circonscrire dans le cortex certaines zones comme celle du langage ou celle de la motricité volontaire fournissent plutôt une topologie fonctionnelle qu'une localisation d'émotions, de sentiments, de désirs, de craintes déterminés. De plus il n'est pas évident que les différences d'intensité dans un sentiment ou une sensation correspondent à des différences d'amplitude ou de fréquence dans le message électrique qui parcourt les neurones ou à des différences de concentration des neurotransmetteurs dans les fentes synaptiques, même si on peut par des apports médicamenteux influencer sur les sentiments internes.

Ces différents arguments tendent tous non seulement à s'opposer à une philosophie réductionniste qui voudrait ramener la psychologie à une description physico-chimique du système nerveux mais aussi à instaurer une infranchissable barrière entre le physique et le mental, entre le corps et l'esprit. Leur enjeu philosophique est donc important puisqu'ils conduisent à une conception dualiste et éradiquent définitivement tout espoir de construire une machine pensante, invalidant au passage tout modèle computationnel de l'esprit qui demeure la source immatérielle et unique du sens. Ils méritent donc d'être minutieusement considérés. Il faut, en effet, se garder de quelque paralogisme.

Un paralogisme ?

Derrière tous ces arguments il y a le postulat implicite de l'identité des indiscernables de Leibniz. Selon ce principe deux objets ayant exactement les mêmes attributs sont identiques, indiscernables et en fait ne constituent qu'un même objet. Deux objets identiques ont les mêmes prédicats et ne font qu'un. Inversement deux objets distincts diffèrent nécessairement par au moins un prédicat. Autrement dit du fait que l'on ne peut pas attribuer exactement les mêmes prédicats à deux objets on doit conclure qu'ils sont différents.

Si les états mentaux étaient équivalents à des états neuronaux, si les premiers se réduisaient aux seconds, selon ce principe on pourrait dire la même chose des uns et des autres. Par suite si l'on trouve un attribut s'appliquant à x et ne s'appliquant pas à y alors x est différent de y et x et y ne sont pas un seul et même objet. Or ce que les arguments ci-dessus veulent mettre en évidence c'est que certaines choses peuvent être dites d'un état mental et non de l'état cérébral que l'on suppose lui être équivalent. Ainsi l'état cérébral correspondant à la sensation de rouge est connu de Mary alors que l'état mental correspondant à cette sensation n'est pas connu de Mary. Il en résulte que la sensation de rouge saisie par l'introspection est de nature différente de l'état neuronal qui lui correspond. Il y a quelque chose de plus dans la sensation de rouge que dans le mécanisme de transmission de l'excitation de certaines cellules de la rétine aux différentes aires visuelles du cortex. Par rapport au premier argument les états mentaux sont connus ou connaissables par introspection alors que les états cérébraux ne le sont pas, il s'agit donc de deux choses différentes. Et par rapport au troisième argument nous pouvons attribuer une fréquence, une amplitude, etc. aux phénomènes nerveux, ce qui n'est pas le cas des phénomènes psychologiques (une croyance ou un désir n'ont pas d'amplitude et de fréquence mesurables); donc on ne peut pas assimiler les seconds aux premiers.

Or un examen attentif montre qu'en l'occurrence on fait un usage tout à fait abusif de l'argument de Leibniz. En effet on utilise le prédicat "être connu" ou "être connaissable" et par-là on confond le plan objectif et le plan subjectif puisque ces prédicats renvoient à un sujet de la connaissance. En fait "être connu" n'est pas un prédicat de même nature que "être rouge" ou "être grand". Le fait que Mary connaisse tel état cérébral mais ne connaisse pas la sensation de rouge ne signifie pas que l'état cérébral et la sensation de rouge ne soient pas la même chose. Paul Churchland dans "Reduction, qualia, and the Direct Introspection of Brain States", (*Journal of Philosophy*, 82, n°1, 1985), repris dans *A Neurocomputational Perspective*, MIT Press, 1985, pp. 47-66. ainsi que dans le même ouvrage au chapitre 4, "Knowing Qualia : a Reply to Jackson", donne de nombreux exemples où un même objet peut être connu sous un nom et non sous un autre et demeuré pourtant identique. Ainsi, le fait que Jacques sache que l'aspirine calme la douleur mais ne sache pas que l'acide salicylique le fasse aussi ne permet pas de conclure que l'aspirine et l'acide salicylique sont deux choses différentes. Le fait que nous sachions ce qu'est une sensation de chaleur ajouté au fait que nous pouvons ignorer que la chaleur n'est rien d'autre que l'énergie cinétique moyenne des molécules en mouvement n'empêcheront pas que l'on peut assimiler la chaleur et l'énergie cinétique moyenne des molécules compo-

sant un corps. Il ne semble donc pas pertinent d'utiliser le pseudo-attribut "être connu" conjointement avec la loi de Leibniz.

Cela est d'autant moins pertinent que la notion même de connaissance recouvre différents processus cognitifs. En effet, il faut distinguer une connaissance immédiate et sensorielle d'une connaissance élaborée et conceptuelle, souvent médiatisée par tout un appareillage théorique complexe. La sensation de chaleur est immédiatement présente à la conscience, le fait qu'elle résulte de l'agitation des molécules n'a été que tardivement dévoilé par le travail de la physique. Il y a donc des niveaux de connaissance, des strates de perception du monde qui nous entoure, des cadres conceptuels différents à travers lesquels nous percevons les choses.

Ainsi, le fait que Mary n'ait pas d'expérience des sensations de couleurs n'empêche pas qu'elle puisse intellectuellement savoir que les couleurs sont liées à des fréquences d'une onde électromagnétique et que la sensation que nous en avons résulte de la sensibilité des différentes cellules de la rétine à ces différentes fréquences. De même le fait que nous ne puissions pas attribuer à nos désirs, à nos croyances, à nos connaissances mêmes, des caractères spatio-temporels du cerveau, des organes nerveux sensoriels et moteurs et de l'influx nerveux, n'exclut pas que l'on puisse lier les uns aux autres. Simplement la connaissance introspective que nous avons des premiers est de nature différente de la connaissance scientifique que nous pouvons avoir ou espérer avoir des seconds.

On a donc de fortes raisons de rejeter les arguments visant à contester l'idée que la pensée sous ses diverses formes serait inexplicable par l'anatomie et la physiologie du système nerveux. Faut-il pour autant défendre une position fermement réductionniste ? Cette position prétend que tous les énoncés de la psychologie populaire peuvent être ramenés à des énoncés des neurosciences qui leur seraient équivalents comme par exemple les deux énoncés suivant le sont : "la température de ce corps est élevée" et "l'énergie cinétique moyenne des molécules qui composent ce corps est élevée". L'expérience vécue que nous livre l'introspection est très forte et nous avons beaucoup de peine à admettre que l'émotion à peine descriptible que nous pouvons éprouver devant une œuvre d'art puisse se réduire à une suite de phénomènes neuronaux qu'un lourd appareillage pourrait permettre de mesurer. Paul Churchland avance un argument qui ne paraît pas complètement convaincant. Selon cet argument il fut difficile à nos ancêtres d'admettre que le ciel était immobile alors que la terre tournait et d'une manière générale il est toujours difficile de changer le cadre conceptuel à travers lequel nous percevons le monde. Cette difficulté s'exprime très largement à travers de nombreuses opposi-

tions. Par exemple, au niveau psychiatrique on opposera une approche chimiothérapique à une approche psychothérapique, le discours neurologique au discours psychanalytique. À la position d'un matérialisme strictement réductionniste ou éliminationniste tel que le développent Paul et Patricia Churchland on peut substituer une autre conception possible sans pour autant revenir au dualisme.

5. Les systèmes hiérarchiques et la question du réductionnisme

Cette position peut nous être une fois de plus suggérée par l'informatique. Il s'agit d'admettre des niveaux relativement indépendants de description et d'explication des systèmes. Dans un système informatique on peut aisément mettre en évidence deux hiérarchies parallèles, l'une logicielle, l'autre matérielle. En effet, à un premier niveau la programmation d'un ordinateur consiste en une suite d'instructions élémentaires directement exécutables par l'unité arithmétique et logique (addition d'octets, décalages de bits vers la gauche ou la droite à l'intérieur d'un octet, opérations booléennes...). Ces instructions sont codées en binaire et rangées de manière successive dans la mémoire pour une exécution séquentielle. Pour une tâche donnée il faudra donc que le programmeur la décompose en une suite de telles instructions, puis les place dans la mémoire de l'ordinateur. Evidemment ces instructions ont la forme de codes spécifiques à chaque type d'unité arithmétique et logique, c'est-à-dire de microprocesseur. Les programmes sont donc difficilement transposables d'une machine à une autre. On imagine la difficulté d'une telle opération dès lors que la tâche en question atteint le moindre degré de complexité ou de généralité. Très tôt on mit au point un programme qui permettait de fournir les instructions élémentaires sous forme de termes mnémotechniques. Ainsi, au-dessus du langage machine proprement dit (suite de uns et de zéros) vient se placer un langage plus accessible, le passage de l'un à l'autre étant assuré par un programme traducteur dénommé *assembleur* et propre lui aussi au microprocesseur. L'assembleur conserve deux défauts majeurs : premièrement le programme reste étroitement lié au type de machine, deuxièmement il reste d'une manipulation difficile puisque ses commandes demeurent proches des opérations élémentaires du microprocesseur. D'où l'idée d'utiliser des langages plus proches du langage naturel et des traducteurs (*interpréteurs* ou *compilateurs*). Par-là on élimine les deux défauts que l'on pouvait reprocher à l'assembleur : facilité d'écriture des programmes et "portabilité" d'une machine à une autre, le compilateur se chargeant d'adapter les programmes aux microprocesseurs particuliers. On superpose ainsi différents

niveaux de langages et, travaillant à l'un de ces niveaux, le programmeur n'a pas besoin de se soucier de la traduction des commandes qu'il écrit dans le langage des niveaux inférieurs. La programmation à l'aide d'un langage de programmation de haut niveau ouvre le champ au développement de l'algorithmique indépendamment de toute préoccupation relative au matériel. Il existe ainsi différents niveaux de description d'un programme informatique. L'étude de l'algorithmique veut même que l'on use d'un pseudo-langage que l'on traduira par la suite dans un langage informatique quelconque. Les problèmes que l'informaticien rencontre sont alors de nature différente suivant le niveau auquel il se place ; à titre d'exemples : problème d'adressage de la mémoire au plus bas, problème de preuve de programmes au plus haut.

Si nous considérons le matériel, nous trouvons aussi une hiérarchie des descriptions possibles. L'unité arithmétique et logique d'un ordinateur est constituée par des portes logiques électroniquement constituées. On peut parfaitement en rester à ce niveau de description pour considérer comment différentes fonctions (additions d'octets, opérations booléennes) sont effectuées sans tenir compte des composants de ces portes logiques. On peut aussi pénétrer en quelque sorte à l'intérieur d'une porte logique et en examiner les éléments, transistors et diodes. À ce niveau-là on peut rendre compte de la fonction assurée par une porte donnée en raison des caractéristiques électriques des composants. La diode est alors vue comme un élément ne laissant passer le courant que dans un seul sens, le transistor comme un interrupteur commandé. On peut encore descendre au niveau de la constitution de ses éléments en examinant comment des courants électriques circulent à l'intérieur des semi-conducteurs dont ils sont faits. Il existe donc un niveau d'explication physique, relevant de l'électronique, et plusieurs niveaux d'explication logique suivant que l'on se place à un degré plus ou moins important d'intégration des portes logiques. D'où la modularité des systèmes physiques en informatique, modularité que l'on retrouve du point de vue logiciel avec la constitution de sous-programmes et même de véritables bibliothèques de routines que le programmeur peut assembler à sa guise en fonction des besoins du programme qu'il élabore.

Toutefois, si l'ordinateur classique avec ses composants électroniques autorise plusieurs niveaux de description, on peut toujours rendre compte de ce qui se passe à un niveau par les éléments inférieurs. Il en va de même dans le domaine logiciel puisqu'on peut toujours traduire un programme écrit dans un langage de haut niveau en assembleur ou en langage machine. Dans les deux cas il n'y a rien de plus dans le tout par rapport aux parties et on peut rester sur une position réductionniste. Les différents niveaux d'explication correspondent à des facilités d'exposi-

tion plus qu'à une composition organique des systèmes. Il résulte aussi de cette hiérarchisation des niveaux de description que la signification elle-même est susceptible d'être comprise de manière plus ou moins modulaire. Le sens d'un objet quelconque peut être saisi de manière plus ou moins globale.

6. Réseaux de neurones formels et propriétés émergentes

La question est différente avec les réseaux de neurones formels. Peu importe la façon dont sont réalisées les différentes cellules de ces réseaux, les capacités dont ils font preuve ne sont pas celles des cellules prises isolément. Chaque neurone artificiel a pour fonction de faire la somme de ses entrées puis, en fonction de la valeur de cette somme, d'activer ou non sa sortie. La capacité d'apprentissage, celle de reconnaître des formes et toute autre capacité qui peut apparaître à l'utilisation de ces réseaux, apparaissent comme émergentes. On veut dire par là que la description des éléments ne rend pas compte du fonctionnement du tout. D'où l'expression : "il y a plus dans le tout que dans la somme de ses parties". On parle aussi à propos de cette propriété de holisme. L'étude des systèmes, la systémique, a beaucoup insisté sur cet aspect de certains dispositifs reprenant des idées de différents domaines, de la psychologie de la forme aux systèmes dynamiques. On a alors souvent mis en avant le fait que notre esprit, nos états mentaux, notre conscience seraient des propriétés émergentes de l'organisation de notre cerveau ou de notre système nerveux. Dès lors, la neurologie et la neurobiologie qui procèdent par analyse et décomposition, par description anatomique et physiologique des éléments qui composent le système cérébral, seraient condamnées irrémédiablement à l'échec du fait de leur incapacité à prendre en charge le tout. Le réductionnisme est donc lui aussi condamné par principe.

La notion de propriété émergente ne va pas sans problème et on peut être tenté de l'utiliser de la même manière que la physique pré-galiléenne utilisait les propriétés occultes, le mot valant explication. On peut aussi tenter d'en donner une description plus rigoureuse. Ou plutôt il faudra probablement trouver dans le jeu des relations qui constitue un système une description rigoureuse qui ne sera pas nécessairement chaque fois la même. Pour chaque système, la notion de propriété émergente recouvre probablement des relations différentes, différemment modélisables même si l'on peut parfois découvrir des analogies intéressantes et révélatrices. Je voudrais pour finir relever l'une d'entre elles qui se trouve au cœur du problème que je viens d'aborder.

Un réseau de neurones formels possédant n cellules en entrée et m cellules en sortie peut être considéré comme un dispositif capable de transformer un vecteur de dimension n en un vecteur de dimension m . Supposons, par exemple un tel réseau capable de reconnaître les caractères alphabétiques manuscrits. Ces derniers sont projetés sur une matrice de cellules photoélectriques de 10 par 10. Ainsi chaque occurrence présentée au réseau correspond à un vecteur de dimension 100. En sortie, nous disposerons d'une série de 80 diodes électroluminescentes, chacune correspondant à un caractère alphanumérique. Le fonctionnement du réseau est tel que seule s'allume la diode correspondant au motif présenté sur la matrice d'entrée. De fait, nous obtenons en sortie un nouveau vecteur de dimension 80 dont chaque position ne peut prendre que les valeurs 0 ou 1 et tel qu'une seule position est à 1 et toutes les autres à 0. On pourrait évidemment remplacer ces diodes par des dispositifs de mesure de tension et considérer que le vecteur de sortie est une suite de 80 valeurs différentes dont une et une seule est beaucoup plus grande que les autres. Peu importe. Le réseau fonctionne bien comme un transformateur de vecteurs.

Nous retrouvons dans le système nerveux quelque chose d'assez semblable. Considérons, par exemple, le cas de la vision des couleurs⁸. Nous savons que celles-ci correspondent à différentes longueurs d'onde de la lumière. Une faible partie du spectre des ondes électromagnétiques est perceptible à nos yeux, c'est-à-dire est capable de provoquer une excitation des cellules qui tapissent notre rétine, sensibles aux couleurs et que l'on nomme les *cônes*. Ce spectre visible s'étend de 380 à 760 nanomètres. On a découvert qu'il existe trois sortes de cônes : certains sont sensibles aux faibles longueurs d'onde, d'autres aux longueurs d'onde moyennes et les troisièmes aux grandes longueurs d'onde de cet intervalle 380-760 nanomètres. Ainsi lorsqu'une lumière colorée va frapper notre rétine, chaque type de cellules va réagir plus ou moins en fonction du type et de la couleur. Chaque cône réagit en émettant un signal à une fréquence donnée. Les cellules sensibles aux faibles longueurs d'onde émettront à une fréquence F_f , celles qui sont sensibles aux longueurs d'onde moyennes à une fréquence F_m et, enfin, celles qui sont sensibles aux grandes longueurs d'onde à une fréquence F_g . Les trois valeurs $\langle F_f, F_m, F_g \rangle$ constituent un vecteur représentatif de la couleur donnée. Il représente une information sensorielle qui sera transmise au cortex

⁸ Pour une étude développée de ces hypothèses et des procédures expérimentales mises en œuvre pour les étayer, on se reportera à P. S. Churchland et T. S. Sejnowski, *The Computational Brain*, MIT Press, Cambridge, Massachusetts, 1992, et en particulier au chapitre 4, "Representing the World", pp. 141-238.

visuel primaire le long du nerf optique, à travers le corps genouillé latéral. On peut d'ailleurs représenter graphiquement toutes les couleurs dans un espace à trois dimensions. Chaque axe correspond à un type de cellules et est étalonné en pourcentage de l'excitation maximum que peuvent délivrer les cônes, de 0 à 100 %. On s'aperçoit que des couleurs ressenties comme proches (le rouge et l'orangé, par exemple) seront représentées par des points proches dans cet espace. De même, une couleur ressentie comme intermédiaire entre deux autres (comme l'orangé entre le jaune et le rouge) aura un point représentatif dans l'espace situé entre les points représentatifs des deux autres.

Cette représentation vectorielle des couleurs peut évidemment être transposée à d'autres informations sensorielles. Sur la langue, par exemple, on peut distinguer quatre types de cellules sensibles à différentes substances. Un goût donné sera codé par un vecteur de quatre valeurs et la représentation des goûts pourra se faire dans un espace à quatre dimensions. S'il nous est difficile de nous représenter un espace possédant plus de trois dimensions, le mathématicien peut parfaitement travailler sur un tel espace. D'autres catégories des sens (comme les odeurs, les formes, etc.) exigeront probablement des vecteurs plus grands encore. Par exemple, on estime que les odeurs sont saisies à travers trente types de cellules sensibles différentes, ce qui donne des vecteurs correspondant à trente valeurs et un espace de représentation à trente dimensions. La position de notre corps dans l'espace (proprioception) est fournie à notre cerveau par une multitude de cellules sensibles réparties dans tout notre organisme et reliées à l'encéphale par les faisceaux de Goll et de Burdach, et au cervelet par les faisceaux spino-cérébelleux. Mais là encore la théorie vectorielle de la représentation peut s'appliquer, même si les vecteurs doivent être encore plus grands. Inversement, on peut penser que notre système nerveux élabore, à partir des vecteurs sensoriels, un vecteur qui va commander par l'intermédiaire des nerfs moteurs, les mouvements de nos muscles et provoquer nos mouvements volontaires ou non.

Je ne développerai pas plus cette théorie vectorielle que les neurosciences confirment peu à peu mais j'en retiendrai deux leçons : premièrement on peut supposer que le traitement de l'information, au niveau cérébral, est un traitement de vecteurs plus ou moins grands ; deuxièmement, on retrouve des représentations en usage dans le domaine des réseaux de neurones formels même s'il demeure de grandes différences entre les systèmes nerveux et les artefacts.

Un traitement de vecteurs suppose un dispositif avec autant d'entrées que le vecteur d'entrée possède de valeurs et fournissant en sortie un nouveau vecteur. Un tel dispositif est réalisé par l'organisation réti-

culaire des neurones formels. Trouve-t-on la même chose en neuroanatomie ? Les techniques de visualisation de l'anatomie et de l'activité cérébrales ne nous fournissent pas encore des images suffisamment fines de la matière grise de l'encéphale pour que l'on puisse y repérer de tels réseaux. Cependant le système d'interconnexions des neurones laisse supposer que de tels dispositifs s'y trouvent, probablement à un niveau de complexité extrême compte tenu du nombre immense de neurones et de synapses. Par ailleurs, dans le cervelet qui joue un rôle important dans la proprioception, dans la faculté d'équilibre et dans la coordination de nos mouvements, la taille de certains neurones (les cellules de Purkinje) nous permettent de retrouver cette organisation réticulaire susceptible d'effectuer un traitement vectoriel.⁹ Les axones des cellules de Purkinje qui sortent du cervelet portent le vecteur de sortie (on a alors un vecteur de plusieurs millions de valeurs). Or les excitations sur ces axones dépendent des excitations reçues sur les dendrites buissonnantes de ces cellules et fournies par les fibres parallèles ainsi que des poids synaptiques des connexions entre ces fibres et les dendrites des cellules de Purkinje. Les fibres parallèles portent les vecteurs en entrée et les cellules de Purkinje effectuent le traitement. Cette façon de voir est, certes, caricaturale car le système est complexifié par la présence des fibres moussues et des fibres grimpantes qui jouent probablement un rôle dans l'évolution des poids synaptiques. On remarquera toutefois que dans la comparaison entre l'organisation réticulaire du cervelet et les réseaux artificiels, nous sommes déjà au-delà d'une simple et vague analogie.

On voit comment peu à peu les questions relatives à la pensée trouvent des formulations en termes d'organisation cérébrale et de structures dynamiques. La notion d'émergence d'une propriété, dans ce cas peut se traduire en termes de traitement de vecteurs. On voit aussi toute la pertinence que les dispositifs connexionnistes peuvent avoir pour modéliser certaines fonctions mentales. Toutefois il demeure de grandes différences entre les structures cérébrales et les réseaux de neurones formels. J'en retiendrai deux. Tout d'abord il y a un immense écart quantitatif entre les quelques centaines de cellules des artefacts et les milliards de neurones du cerveau dotés chacun de milliers de synapses. Cette différence quantitative implique nécessairement une différence de complexité tout aussi vertigineuse. Ensuite, il existe un caractère particulier des neurones : l'apprentissage chez les êtres vivants, dont l'homme, ne tient pas seulement à la possibilité de modifier des poids synaptiques mais aussi à la possibilité que des synapses disparaissent et que d'autres

⁹ Cf. P. Buser et M. Imbert, *Mécanismes fondamentaux et centres nerveux*, Hermann, Paris, 1993, chapitre 18 "Le cervelet et son cortex", pp. 483-510.

surgissent, à la possibilité qu'un axone croisse, s'allonge comme s'il recherchait à nouer de nouvelles connexions. On sait maintenant comment une telle croissance se fait et comment elle est guidée par l'expérience et par des mécanismes chimiques¹⁰. Mais cette connaissance ne nous permet pas encore de reproduire de tels phénomènes encore qu'elle soit une première étape sur ce chemin. Les tentatives de connexions de neurones biologiques et artificiels constituent une étape importante dans ce sens.

7. Conclusion

En conclusion je dirai que si le cerveau a manifestement un pouvoir organisateur — substitut matériel des cadres *a priori* de la sensibilité tels que Kant les postule dans l'*Esthétique transcendantale* —, il faut encore que ce qui est saisi possède une organisation qui entre en correspondance avec ces cadres. L'ancienne discussion entre l'importance respective de l'inné et de l'acquis ou autrement entre empirisme et innéisme se trouve renouvelée et déplacée du terrain métaphysique à celui des sciences expérimentales. L'esprit, ou le cerveau, n'est pas une *tabula rasa* sur laquelle l'expérience viendrait inscrire le sens. Le cerveau est perpétuellement actif, intégrant constamment les informations fournies par les sens à travers une véritable activité computationnelle. Conformé initialement par le codage génétique, à chaque instant sa structure dépend de manière complexe autant de cette forme initiale que de l'expérience sensorielle accumulée. La structure initiale, riche en neurones mais encore pauvre en synapses dépend aussi de toute l'évolution au terme de laquelle se trouve notre espèce. La mise en place des synapses au cours de l'enfance dépend, elle aussi, autant de facteurs génétiques que du flux de l'information sensorielle. Ajoutons encore que le cerveau n'est pas un simple intermédiaire entre les entrées sensorielles et les sorties motrices, mais acquiert et possède une certaine capacité à s'auto-contrôler, à s'auto-modifier. De ce fait, les causes de nos actions et comportements tiennent autant à la faculté résultante de nous projeter dans l'avenir et à agir en fonction de buts que nous nous représentons qu'à des réponses immédiates à des stimuli sensoriels.

Ainsi, ces découvertes des neurosciences, même si elles laissent encore de nombreuses questions non résolues, jettent une lumière toute particulière sur beaucoup d'aspects de notre vie mentale. On peut désor-

¹⁰ Cf. C. Shatz, "La maturation du cerveau" in *Pour la Science*, n° 181, novembre 1992, pp. 50-59.

mais trouver une signification toute matérialiste et espérer une formulation quasi scientifique de l'intuition que formule Roger Caillois dans le domaine de l'art :

L'harmonie, d'abord, apparaît comme une médiation entre une structure mathématique générale et un organisme sensitif. Il semble à première vue paradoxal que, reposant sur le nombre, son effet soit directement physiologique, et non intellectuel, comme on attendrait inconsidérément. La question s'éclaire si l'on pense qu'il existe une structure mathématique de l'univers régissant aussi bien l'homme que le milieu (ces deux termes n'ayant d'ailleurs de valeur absolue que dans le déplorable langage anthropocentrique), dépendant sans doute en dernière analyse du principe d'économie (comme tend à le montrer par exemple le calcul de l'angle optimum de la disposition en hélice des feuilles sur la tige pour la réception maxima des rayons solaires) et dont dépendent à leur tour aussi bien la morphologie des cristaux que le rythme de croissance des coquillages et des plantes.¹¹

Le sens, comme ici l'émotion esthétique associée à la perception de certaines harmonies, tient donc à la rencontre dans une conception nouvelle de l'automate entre une forme cérébrale dynamique et une forme perçue. Nous avons peut-être perdu l'exorbitant privilège d'un esprit que l'on n'aurait partagé qu'avec Dieu ; nous y gagnerons peut-être aussi les fondements plus assurés d'une nouvelle éthique¹² dégagée des dogmatismes et des sectarismes, vertiges potentiels jusqu'ici de toute philosophie.

Université de Bourgogne, gerard.chazal@u-bourgogne.fr

¹¹ R. Caillois, *Approches de l'imaginaire*, Gallimard, Paris, 1974, pp. 44-45.

¹² Cf. J.P. Changeux, "Point de vue d'un neurophysiologiste sur les fondements de l'éthique" in M. Meulders, M. Crommelinck et B. Feltz (eds.), *Pourquoi la science ? Impacts et limites de la recherche*, Champ Vallon, Seyssel, 1997, pp. 185-202.

BIBLIOGRAPHIE

- [1] BEAUNE Jean-Claude, *Le vagabond et la machine*, Champ-Vallon, Seyssel, 1983.
- [2] BEAUNE Jean-Claude, *Les spectres mécaniques*, Champ-Vallon, Seyssel, 1988.
- [3] BUSER Pierre & IMBERT Michel, *Mécanismes fondamentaux et centres nerveux*, Hermann, Paris, 1993.
- [4] CAILLOIS Roger, *Approches de l'imaginaire*, Gallimard, Paris, 1974.
- [5] CHANGEUX Jean-Pierre, "Point de vue d'un neurophysiologiste sur les fondements de l'éthique", *Pourquoi la science ? Impacts et limites de la recherche*, Champ Vallon, Seyssel, 1997, pp. 185-202.
- [6] CHURCHLAND Patricia S. & SEJNOWSKI Terrence J., *The Computational Brain*, MIT Press, Cambridge, Massachusetts, 1992.
- [7] CHURCHLAND Paul, *A neurocomputational Perspective*, MIT Press, Cambridge, Massachusetts, 1985.
- [8] DESCARTES René, *De l'homme, Œuvres philosophiques*, Garnier Flammarion, (éd. F. Alquié), tome I.
- [9] JACKSON Frank, "Epiphenomenal Qualia", *Philosophical Quaterly*, 32, 127, 1982, pp. 127-136.
- [10] JACKSON Frank, "What Mary didn't know", *Journal of Philosophy*, 83, 5, 1986, pp. 291-295.
- [11] McCULLOCH Warren. & PITTS Walter. "A logical calculus of the ideas immanent in nervous system", *Bulletin of Mathematical Biophysics*, 5, 1943, pp. 115-133.
- [12] NAGEL Thomas, "What is it like to be a bat?", *Philosophical Review*, 83, 4, 1974, pp. 435-450.
- [13] NEUMANN John von, *Théorie générale et logique des automates*, Champ-Vallon, Seyssel, 1996.
- [14] ROSENBLATT Frank, *Principles of Neurodynamics*, Spartan Books, New York, 1959.

Cahiers François Viète

La revue du *Centre François Viète*
équipe de recherche en
Épistémologie, Histoire des Sciences et des Techniques
de l'Université de Nantes
(EA 1161)
www.cfv.univ-nantes.fr

Les *Cahiers François Viète* publient chaque année les principales conférences du séminaire d'Histoire des sciences et des techniques du Centre. Des numéros spéciaux peuvent être consacrés à des thèmes particuliers, correspondant par exemple à des actes de colloque.

Responsable de publication - Stéphane Tirard

Comité de rédaction

Guy Boistel (Nantes)	Jacques Gapaillard (Nantes)
Céline Briée (Nantes)	Jean-Louis Kerouanton (Nantes)
Olivier Bruneau (Nancy)	Pierre Teissier (Nantes)

Secrétaire de rédaction - Sylvie Guionnet