



MICHEL VEUILLE

La machinerie des instincts

Avec Christiane, nous allons essayer de faire un séminaire sur le concept de machine – on parle beaucoup de machines ici – moi dans le rôle du biologiste et Christiane dans le rôle de la mathématicienne, puisque cela correspond à nos attributions respectives.

Le problème, c'est qu'en biologie il n'y a pas de concept de machine. Depuis que la biologie est biologie il y a des centaines de constructions machiniques qui ont été imaginées mais il n'y a pas de concept pour en parler.

Dans un premier temps je vais raconter l'histoire de toutes ces machines, parce qu'elles ont une histoire. Je vais parler de la machine comportementale, de la machinerie des instincts et Christiane enchaînera à partir de cela sur les machines mathématiques qui sont beaucoup plus récentes et certainement beaucoup plus formalisées.

Aristote a été le premier zoologiste à décrire l'esclave comme une machine animée (*La Politique*).

Il est d'usage de dire depuis lors que l'idée de machine renverrait à celle de cause mécanique, efficiente.

Selon Georges Canguilhem, l'évocation du fonctionnement machinique d'un être serait une explication d'ordre analogique se référant à ces objets réels consacrés par la technique, assemblages de pivots, de rouages et de cliquets, astreints à quelque tâche utilitaire et infiniment répétable, car inscrite dans l'agencement même de leur structure.

« On peut définir la machine comme une construction artificielle, œuvre de l'homme, dont une fonction essentielle dépend de mécanismes. Un mécanisme, c'est une configuration de solides en mouvement telle que le mouvement n'abolit pas la configuration. Le mécanisme est donc un assemblage de parties déformables avec restauration périodique des mêmes rapports entre les parties. L'assemblage consiste en un système de liaisons comportant des degrés de liberté déterminés (...). La réalisation matérielle de ces degrés de liberté consiste en guides, c'est-à-dire en limitations de mouvements de solides au contact. » (Canguilhem, *La connaissance de la vie*, p. 102)
On a souvent tendance à mépriser ses prédécesseurs ; mais il ne semble pas, en réalité, que quiconque ait jamais songé sérieusement que le corps ou l'esprit puissent fonctionner sur de tels principes.

Et de façon plus générale, que quelqu'un ait jamais cru que l'explication machinique puisse reposer sur les dispositifs utilitaires, ouvriers ou ménagers, utilisés à son époque, sinon dans un but illustratif et métaphorique.

S'il en était ainsi, l'on devrait s'attendre à ce que les structures imaginées par les biologistes du corps et de l'esprit du XIX^e siècle, siècle du machinisme par excellence, aient quelque parenté avec celles manipulées par les ingénieurs de l'époque.

Et sans doute les manuels de dissection et d'anatomie comparée sont-ils saturés de termes évoquant l'architecture fonctionnelle des corps : « animaux articulés », arthropodes, hexapodes, myriapodes ; les psychiatres eux-mêmes discutent sur les « automatismes mentaux ».

S'il fallait s'attacher au sens premier de ces expressions, la science en vogue du XIX^e siècle devrait être quelque forme de chirurgie. En réalité, la discipline de pointe de cette époque, celle dont chacun est convaincu, biologiste ou psychiatre, qu'elle détient le secret du fonctionnement des êtres, c'est la physiologie. Or, la fonction glycogénique du foie n'évoque aucun artetact sorti des ateliers du XIX^e siècle.

La « régulation du milieu intérieur » de Claude Bernard préfigure la notion d'homéostasie, mais n'a rien de commun avec un régulateur à boules, et ne lui doit rien.

Le vivant n'a vraiment été comparé à un automate que dans le siècle qui a suivi Descartes. Les machines étaient alors des jouets : canards qui béquettent et qui défèquent, automate joueur d'échecs, joueuse de tympanon construite pour Marie-Antoinette. C'étaient des objets d'art et de démonstration construits en exemplaire unique. Êtres singuliers et merveilleux, ils fréquentaient le milieu choisi des cours d'Europe et des salons philosophiques. Et quant aux machines utilitaires, il y avait certes, à cette époque, des moulins à eau, mais aucun philosophe n'y a jamais comparé l'esprit.

Les mythes des hommes perdus dans les engrenages de la fabrique, ou détrônés par les robots appartiennent à la culture populaire, mais ils relèvent d'un imaginaire autrement plus riche que celui d'une crainte quelconque inspirée par les objets de l'industrie. Ils ne doivent pas en tous cas masquer le fait que ce sont les machines imaginaires qui ont, en tous temps et en tous lieux, précédé les machines réelles, et non pas l'inverse.

Mais s'il en est ainsi, où est la vertu explicative de la machine ? D'où nous vient cette illusion d'intelligence du fonctionnement d'un être, à la simple évocation de son statut machinique ? Quelle est la plus value de l'argument machinique par rapport à une éventuelle non explication machinique ? À quoi s'oppose la machine dans le langage ?

L'inconscient du machiniste dérouté encore par la prolifération de sa production, car toutes les machineries imaginaires ne sont pas construites sur un mode unique et l'on peut distinguer, au cours de l'histoire des sciences inexactes, biologie et sciences humaines, trois grandes périodes au cours desquelles l'idée de machine s'est renouvelée.

1/ Les automates de l'époque de Descartes sont les seuls qui correspondent à l'image classique du mécanisme. Réalisés pour accomplir quelque tâche nécessairement préétablie par leur constructeur, ils appartiennent à un monde dualiste. La créature renvoie à son créateur : cause finale.

2/ La science du XIX^e siècle se veut matérialiste, moniste, évolutionniste. Les machines biologiques et psychiques sont alors conçues comme des corps organisés, capables d'assurer leur développement par des voies endogènes. L'embryologiste Driesch montre que si l'on dissocie les blastomères d'un œuf d'oursin, chaque blastomère peut régénérer un individu.

Ces machines organicistes n'ont pas à être construites. L'évidence de leur fonctionnement ne réside pas dans le jeu prédéterminé des pièces et des rouages, comme chez celles qui sortent de la fabrique. Elles sont constituées d'organes auxquels sont attribués des fonctions spécifiques. Il suffira pour les décrire d'un simple diagramme mettant en évidence la différenciation des parties, et leur agencement respectif. Leur secret n'est pas dans la transmission du mouvement selon des procédures automatiques, mais dans la complexité organisationnelle de l'ensemble.

L'âme de cette époque est le positiviste Herbert Spencer, philosophe de l'organicisme et de l'évolutionnisme, qui base sa philosophie sur deux principes :

- l'instabilité de l'homogène
- la conservation de la force

À chaque niveau de complexité structurale, le positivisme se donne simplement pour tâche de répertorier les propriétés émergentes de la matière, à partir de points d'observation dispersés le long de l'échelle des sciences : cause matérielle.

3/ Le xx^e siècle voit réalisée la continuité substantielle du monde dans toute l'étendue de l'échelle des sciences, et s'intéresse maintenant aux propriétés structurelles. Les « fonctions » de l'organisme n'ont plus leur contenu immanent. Elles sont prises dans des explications formelles qui transcendent leur nature particulière. Les mêmes équations valent pour une variété de systèmes sans apparemment évolutif : cause formelle.

Le comble des machines imaginaires est de n'avoir jamais fonctionné sur le mode de la prétendue « cause efficiente ».

Cet exposé envisagera l'histoire de l'art de fabriquer des machines imaginaires en biologie du comportement animal. Il ne s'attardera pas à la filiation des machines elles-mêmes, car à une même époque, les jeux d'opposition entre écoles multiplient les modèles ; les emprunts et les références multiples, les plagiats, les transferts d'une discipline à une autre, la différence d'habileté de chaque artisan, ou la facture variable d'un même composant selon l'atelier qui l'a produit, rend les généalogies contuses. On peut par contre envisager les lignes directrices dans l'évolution des technologies. Plutôt que de s'intéresser à la ressemblance entre les produits finis, on s'attachera aux changements de traditions dans le travail des machinistes.

Les machines dont je parlerai sont celles de l'instinct situées à la jointure des machines animales et des machines, et ce pour deux raisons.

La première est qu'à une époque donnée, les mêmes schèmes, situés à la frontière du concevable, diffusent dans tous les domaines qui visent à l'intelligibilité du complexe.

La seconde est que l'animalité a toujours été considérée comme l'hypothèse nulle pour l'étude de l'homme, le point zéro de l'humanité, que l'on ait pour but de les opposer de manière dualiste, ou de trouver ce qui fait la spécificité de l'homme à partir de son origine.

I/ Les machines divines

Le corps-machine devient une hypothèse scientifique avec les travaux de H. sur la circulation du sang (1628) et le *Discours de la Méthode* de Descartes (1637) qui développe l'idée de l'animal machine. Il correspond bien alors à son étymologie grecque qui signifie ruse, machination. Le Moyen-Âge connaît les machines de théâtre. Elles supposent un machiniste (Deus ex machina). L'animal-machine n'est que le faire-valoir de son créateur. Il permettra au XVIII^e siècle d'avancer l'argument du « Grand Horloger ». Et jusqu'au début du XIX^e siècle, alors qu'apparaissent les premières théories évolutionnistes (Lamarck, Darwin...), Payley développe l'argument de la montre : l'adaptation des êtres vivants à leur milieu, équilibre merveilleux obtenu par les multiples inventions de la nature, prouve le dessein du créateur (c'était déjà l'argument de Plotin ou de Saint-Augustin : la beauté du monde qui nous entoure est un signe de l'intervention divine).

II/ Au siècle des lumières

Les machines sensualistes de Condillac permettent aux animaux de fonder leur sensibilité sur les besoins qui naissent de leur organisation, et non plus sur un don de la grâce divine : « L'instinct n'est rien, ou c'est un commencement de la connaissance. Car les actions des animaux ne peuvent dépendre que de trois principes : ou d'un pur mécanisme, ou d'un sentiment aveugle, qui ne compare point, qui ne juge point, ou d'un sentiment qui compare, qui juge et qui connaît. » (Condillac, *Traité des animaux*, p. 489).

Si le *Traité des sensations* est dirigé contre Descartes, le *Traité des animaux* est écrit contre Buffon : « Je conçois, écrit-il à son adresse, que si le chien était poussé comme une boule, par deux forces égales et directement contraires, il resterait immobile, et qu'il commencerait à se mouvoir lorsque l'une des deux forces deviendrait supérieure. Mais avant que ces ébranlements donnent des déterminations contraires, il faudrait pouvoir prouver qu'ils donnent des déterminations certaines : précaution que Monsieur de Buffon n'a pas prise ». (Condillac, *ibid.*, p. 446). C'est l'argument de l'âne de Buridan.

La thèse de Condillac annonce déjà la jointure entre le siècle des lumières et le XIX^e siècle qu'opérera Lamarck. Condillac écrit : « Les habitudes naissent du besoin d'exercer ses facultés : par conséquent, le nombre des habitudes est proportionné au nombre des besoins. » (Condillac, *ibid.*, p. 478) Et Lamarck écrit presque en écho : « De grands changements dans les circonstances amènent pour les animaux de grands changements dans leurs besoins, et de pareils changements dans leurs besoins en amènent nécessairement dans leurs actions. Or, si les nouveaux besoins deviennent constants ou très durables, les animaux prennent alors de nouvelles *habitudes*, qui sont aussi durables que les besoins qui les ont fait naître. » (*Philosophie zoologique*, p. 2).

Avec Condillac tout se passe comme si les animaux pouvaient faire l'économie du créateur. Il devait nécessairement s'ensuivre une renégociation des rôles entre Dieu, les animaux et les hommes. L'animal se rapproche de l'homme, en même temps que le dualisme entre le corps et l'esprit de l'homme se transforme en une dualité fragile.

« Retrançons d'un homme fait le *moi de la réflexion*, on conçoit qu'avec le seul *moi d'habitude*, il ne saura plus se conduire lorsqu'il éprouvera quelqu'un de ces besoins qui demandent de nouvelles vues et de nouvelles combinaisons. Mais il se conduira encore parfaitement bien, toutes les fois qu'il n'aura qu'à répéter ce qui est dans l'usage de faire. Le moi d'habitude suffit donc aux besoins qui sont absolument nécessaires à la conservation de l'animal. Or, l'instinct n'est que cette habitude privée de réflexion. » (*Ibid.* Condillac, p. 489)

« Il y a dans la bête ce degré d'intelligence que nous appelons *instinct* ; et dans l'homme, ce degré supérieur, que nous appelons *raison*. » (*Ibid.*, p. 529).

Cette conception anthropomorphique de l'instinct animal. et parallèlement, cette dualité psychique de l'homme entre *moi animal* et *moi raisonnable* seront des constantes de la science du XIX^e siècle jusqu'aux alentours de 1900.

Il n'est pas anodin, pour terminer sur Condillac de noter qu'il utilisait l'expression « d'êtres organisés » pour désigner les animaux. Vers la fin du XIX^e siècle, le terme « d'évolution des êtres organisés » sera synonyme de transformisme.

III/ Darwin

Darwin était fort préoccupé du problème de l'origine du psychisme humain, et lui a consacré des chapitres entiers de plusieurs de ses ouvrages.

Ses conceptions sur la question n'avaient rien de trivial. Il a nettement avancé l'idée, inédite à son époque, que la sexualité humaine commençait chez le nourrisson, thèse que connaissait Freud, et dont on sait la place dans sa théorie psychanalytique.

Au XIX^e siècle, c'est la conscience qui était conçue comme apanage de l'humanité. On ne méconnaissait pas l'existence de phénomènes inconscients, mais ceux-ci exprimaient plutôt notre nature animale. La biologie de Darwin était loin, cependant, de les opposer. Organiciste, comme toute celle du XIX^e siècle, elle était caractérisée, non seulement par l'idée d'hérédité, mais aussi par celle de développement.

Pour Darwin, les capacités intellectuelles s'étaient développées sous l'influence conjuguée de l'hérédité et de la sélection naturelle, et avaient fait naître le langage, « moitié-art, moitié-instinct ».

Quant à la conscience, elle naissait d'un conflit intérieur entre instincts sociaux et pulsions irréflechies, arbitré par la raison.

« J'ai cherché à prouver que le sens moral (de l'homme) résulte premièrement de la nature des instincts sociaux toujours présents et persistants ; secondement de l'influence qu'ont sur lui l'approbation et le blâme de ses semblables ; troisièmement de l'immense développement de ses facultés mentales et de la vivacité avec laquelle les événements passés viennent à se retracer à lui ; et par ces derniers points, il diffère complètement des autres animaux. »

Cette disposition d'esprit entraîne l'homme à regarder malgré lui en arrière et en avant, et à comparer les impressions des événements et des actes passés.

« Aussi, lorsqu'un désir, lorsqu'une passion temporaire l'emporte sur ses instincts sociaux, il réfléchit, il compare les impressions maintenant affaiblies de ces pulsions primitives (?) avec l'instinct social toujours présent, et il éprouve alors ce sentiment de mécontentement que laissent après eux tous les instincts auxquels on n'a pas obéi.

Il prend en conséquence la résolution d'agir différemment à l'avenir, – c'est là ce qui constitue la conscience. Tout instinct qui est constamment le plus fort ou le plus persistant éveille un sentiment que nous exprimons en disant qu'il faut lui obéir. » (Darwin, *La Descendance de l'Homme*, pp. 668-9).

Sous une grande naïveté, on discerne déjà les grandes lignes d'un autre système, qui aura été grandement influencé par l'organicisme du XIX^e siècle, celui des topiques freudiennes, et de l'affrontement entre le moi, le surmoi et les pulsions du ça.

La grande différence est que la partie inconsciente du psychisme humain était alors conçue comme exprimant sa partie animale. Il n'en reste pas moins que la psychanalyse n'est pas tombée du ciel.

Darwin était tant préoccupé du problème qu'il lui consacre encore un ouvrage : *L'expression des émotions chez l'homme et les animaux*. Le thème traité était celui de «...l'acquisition par le développement de certaines actions réflexes dans lesquelles les muscles qui ne peuvent être influencés par la volonté sont mis en action (...) ».

« Car la conscience, dont dépend le sens de l'utilité, ne peut être intervenue dans le cas d'actions effectuées par des muscles involontaires. Le mouvement magnifiquement adapté de l'iris lorsque la rétine est stimulée par trop ou trop peu de lumière, est un exemple du problème. » (Darwin, « Origin of certain instincts », *Nature*, 1873).

Le livre de Darwin se signale notamment par l'importance accordée au psychiatre français Duchenne de Boulogne. C'était l'un des maîtres de Charcot. Ses travaux de neurologie mettaient à contribution deux des inventions scientifiques du XIX^e siècle : l'électricité et la photographie. Il soignait l'hystérie par des séances d'électricité. Par ailleurs il avait collaboré avec Nadar. Aujourd'hui la fixation des expressions sur la pellicule est considérée comme un art. À l'époque c'était une découverte d'ampleur scientifique. De leur collaboration Nadar avait acquis son habileté à dresser un portrait expressif de ses contemporains. Duchenne qui classifiait les maladies mentales comme on classifie les espèces, en avait tiré des photographies des principales expressions de base chez l'homme, que Darwin regardait comme un « progrès considérable » et qui illustraient tout son ouvrage.

Un exemple de problème traité par Darwin et qui donne une bonne idée de la problématique de l'ouvrage est celui du rougissement.

Certains psychiatres et criminologues influencés par les idées évolutionnistes disaient que les hommes ayant une mentalité primitive dénuée de sens moral étaient incapables de rougir. Darwin envoya une circulaire à tous les médecins de l'empire britannique, leur demandant, entre autres questions, s'ils avaient déjà vu rougir des sauvages. Et les réponses parviennent :

- Les juifs rougissent comme les aryens.
- Les Chinois possèdent une expression qui signifie « rougir de honte ».
- Il est fréquent de voir la rougeur sur les joues des plus belles femmes de Tahiti.
- Mr. Washington Matthews a vu souvent rougir les jeunes *squaws* (femmes) appartenant aux diverses tribus indiennes de l'Amérique du Nord.
- Plusieurs observateurs dignes de foi assurent avoir vu sur le visage des nègres quelque chose qui ressemblait à la rougeur, sous l'influence de circonstances qui l'auraient excitée chez les blancs ; leur peau était pourtant d'un noir d'ébène.

Quelques-uns décrivent ce phénomène en disant que chez eux, la rougeur est brune (...).

- Chez une négresse albinos parfaitement caractérisée, décrite par Buffon, on voyait une teinte légèrement cramoisie s'étendre sur ses joues lorsqu'elle se montrait nue. Les cicatrices de la peau demeurent blanches très longtemps chez les nègres, et le docteur Burgess, qui eut de fréquentes occasions d'observer des balafres de ce genre sur le visage d'une négresse, la vit distinctement « devenir rouge, chaque fois qu'on lui parlait sans qu'elle s'y attendît, ou qu'on l'injurait d'une façon grossière ».

- Les races sémitiques rougissent aisément, comme on devait s'y attendre d'après leur ressemblance générale avec les races aryennes. (*L'expression*, p. 338).

Le mystère des automatismes inconscients de l'homme reste entier.

En conclusion, Darwin donnera des mécanismes comportementaux non conscients une explication neurologique basée sur trois principes :

- « Les mouvements utiles à l'accomplissement d'un désir ou au soulagement d'une sensation pénible finissent, s'ils se répètent fréquemment, par devenir si habituels qu'ils se reproduisent toutes les fois qu'apparaissent ce désir ou cette sensation. »

- Principe de l'antithèse : « Un usage constant (...) affermit en nous l'habitude d'exécuter volontairement les actes opposés (...) (parfois même), dans un état d'esprit déterminé, une tendance involontaire, irrésistible, à l'accomplissement d'actes absolument contraires doit se produire ». Rappelons que c'est par ce principe que les psychiatres expliquaient ce fait d'observation bien connu que, par exemple, une personne réputée pour sa grande chasteté est parfois assaillie d'idées obscènes.

- Action directe de l'énergie nerveuse sur le système nerveux : « Cela revient à dire que la force nerveuse prend volontiers des voies qu'elle a déjà parcourues. » Thèse reprise en 1950 par Hebb, et qui est encore la théorie de l'apprentissage la plus prisée des neurologues.

L'inconscient travesti

À ce point de l'histoire des machines imaginaires, nous pouvons résumer ainsi le chemin parcouru depuis Descartes.

Le dualisme entre le corps-machine et l'âme n'a été remis en cause que sur la base d'une confusion croissante entre l'animal et l'humain.

La description des mœurs animales au XIX^e siècle est devenue très anthropomorphe. Il a fallu accepter une dualité du psychisme humain, en surajoutant à sa raison ce qui apparaît pour le moment comme un soubassement animal, s'exprimant par des « automatismes » échappant à sa conscience.

L'inconscient s'avance masqué, sous la figure de l'animal.

Nous approchons du point crucial où l'homme et l'animal vont se dévêtir et échanger leurs défroques. Au début du *xxe* siècle, l'animal va redevenir un pur automate, et l'inconscient va devenir l'attribut essentiel de l'humanité.

Au delà de ce point de rupture, Lacan pourra dire : « Il n'y a d'inconscient que chez l'être parlant. Chez les autres qui n'ont d'être qu'à ce qu'ils soient nommés, bien qu'ils s'imposent du réel, il y a de l'instinct, soit le savoir qu'implique leur survie.

Encore n'est-ce que pour notre pensée, peut-être là inadéquate. » (Lacan, *Télévision*, p. 15)

Pour l'instant, cette confusion des rôles, on la trouve pour l'animal, dans le livre sur *L'évolution mentale chez les animaux*, écrit en 1884 par Romanes, élève chéri de Darwin, et dont Freud devait faire une lecture abondamment annotée :

« Je me suis efforcé de démontrer que l'origine des instincts peut être ce que j'ai appelé primaire ou secondaire, c'est-à-dire que je crois qu'il y a un nombre de faits établissant que les instincts peuvent naître soit par la fixation, au moyen de la sélection naturelle, d'habitudes dépourvues de but précis, qui se trouvent être avantageuses ; dans ce cas, ces habitudes deviennent des instincts sans que l'intelligence s'en soit jamais mêlé ; soit par la transformation d'habitudes, originellement intelligentes, en actes automatiques, grâce à la répétition. » (Romanes, *op. cit.*)

Côté homme, l'inconscient se trouve sur le fil du rasoir avec Janet, qui reprend l'idée de Charcot selon laquelle l'hystérie est une maladie des idées, de l'imagination, qui crée le concept de subconscient et déclare en 1895 : « Je ne serais pas surpris si, – pour comprendre les idées fixes en général, on était obligé de commencer par l'étude des idées fixes subconscientes. » (*Névroses et idées fixes*, p. 231).

Malgré tout, ce subconscient reste pour lui de nature essentiellement pathologique, le résultat d'une défaillance de la volonté : « Cette conservation des groupements anciens (d'idées) une fois constitués, une fois organisés, nous explique les idées fixes et les obsessions (...). »

« La force de ces idées fixes vient précisément de la faiblesse des idées nouvelles acquises à chaque moment ; c'est par nos pensées actuelles que nous résistons à la marée montante de nos souvenirs. Supprimez ce "réducteur antagoniste", et nos souvenirs anciens vont se reproduire, se combiner de mille manières, d'une façon facile, automatique et irréversible. Les souvenirs, les répétitions de mots, les rêveries sont tout près de notre conscience. Il suffit d'un moment de distraction pour qu'ils affleurent, il suffit d'un moment de sommeil pour qu'ils se répandent sur tout l'esprit pendant les rêves et les cauchemars (...). Il suffit de se réveiller un moment, de reprendre la claire perception des choses nouvelles et de ses propres changements pour voir s'évanouir tous ces fantômes. » (Janet, *ibid.*, p. 53).

Le Carnaval de Genève

Avant d'abandonner l'homme pour considérer les nouveaux paradigmes de l'instinct, situons précisément ce point du début du siècle, ce carrefour où homme et animal changent mutuellement de peau.

Freud n'a pas encore sa notoriété. Nous sommes à Genève en 1909. L'association de psychologie physiologique fondée par Charcot tient le *vie* congrès international de psychologie organisé par F. et Claparède. Parmi les nombreux points à l'ordre du jour, deux tables rondes retiennent l'attention. L'une sur le subconscient avec Janet et Morton-Prince, l'autre sur les tropismes, c'est-à-dire la description mécaniste du comportement animal, avec les zoologistes Jacques Lab et Jennings. Signe des temps, on a même invité un botaniste, Francis Darwin.

En ce qui concerne le *subconscient*, déjà en concurrence avec la notion d'*inconscient*, Morton-Prince expose les six acceptions alors en usage : l'inconscient.

Subconscient : 1/ Cette partie du champ de notre conscience qui à tout moment, est en dehors des formes de l'attention.

Conscient : 2/ Psychologie anormale ; idées dissociées ou séparées ; séparées de la conscience personnelle majeure, des formes de l'attention, de telle façon que le sujet est entièrement inconscient, bien qu'elles ne soient pas inertes mais actives.

Subconscient : 3/ Écriture ou parole automatique.

Subconscient : 4/ Coconscient + expériences conscientes passées qui sont soit oubliées, soit hors de l'esprit.

Inconscient : 5/ Moi caché : tout esprit serait double, (...) le moi éveillé et le moi submergé.

Inconscient : 6/ Idées subliminales (...). Le moi personne devient même une conscience inférieure émergeant d'une conscience supérieure, parfois conçue comme une partie d'un monde transcendant.

Avec la deuxième table ronde, nous venons à un domaine purement animal, et nous quittons définitivement le XIX^e siècle. Elle oppose Jacques Lob, le promoteur du « mécanisme » à Jennings, l'auteur de la théorie de l'apprentissage par « essais et erreurs ».

Pour le premier, la base du comportement est dans les tropismes, c'est-à-dire dans des réactions passives d'orientation de l'animal par rapport à son milieu.

Par exemple, dans une cage expérimentale, on met un insecte, organisme à symétrie bilatérale. On éclaire une source lumineuse frappant l'animal de côté. L. explique que sous l'effet de cette dissymétrie de la stimulation, les réactions d'oxydo-réduction se produisent à vitesse différente des deux côtés du corps. Cette motricité inégale fera tourner l'animal lors de son déplacement vers la source lumineuse. Elle s'égalisera lorsque les deux côtés du corps seront également illuminés (cf. papillons). À cette théorie physico-chimique de l'orientation, il ne manque que d'expliquer le S.N.C. Elle se forme en réaction contre l'anthropomorphisme du XIX^e siècle, et aura un succès éphémère. C'est la première fois, et la dernière, depuis Descartes, que l'animal est décrit comme une machine pure. Ceci n'est rendu possible que par le refus des auteurs concernés de considérer l'utilité adaptative des comportements.

C'est ainsi que s'achève le processus d'échange des qualités de l'homme et de l'animal. L'homme ne peut être connu que par son inconscient. L'animal ne peut être connu que par ses actes extérieurs, son comportement. Mais ce n'est pas parce qu'il a hérité de la raison. C'est parce qu'il est connu par l'intermédiaire d'un processus rationnel : le dispositif expérimental où le scientifique l'enferme. Si nouveau dualisme il y a, il est méthodologique.

L'éthologie

L'étude des tropismes ne permet pas cependant d'étudier le comportement naturel : d'une part, parce qu'il évacue tout sens adaptatif des conduites ; d'autre part, parce qu'il nie l'existence de toute motivation endogène, l'animal n'étant connu que par ses réactions au dispositif expérimental.

À quoi sert de savoir qu'un papillon est attiré par un lampadaire ?

Tinbagen, en décrivant le comportement sexuel de l'épinoche transpose métaphoriquement le dispositif expérimental dans la nature, permettant d'objectiver l'existence de stimuli déclencheurs, et donc de garder la même description formelle.

Il faut cependant introduire la notion de *motivation*, en admettant l'existence de niveaux *hiérarchiques* supérieurs du comportement, celui des grands *instincts*.

Cette étude de *l'instinct*, dans le cadre de la description rigide héritée du *mécanicisme*, est baptisée par Konrad Lorenz *éthologie objective* :

« *Comportement appétitif* : phase variable d'introduction d'un schème ou d'une séquence comportementale instinctive. *Acte consommatoire* : acte constituant la fin d'un schème ou, d'une séquence comportementale instinctive. » (Thorpe)

Les pulsions

La notion d'instinct amène à poser un problème dont nous n'avons pas parlé, celui du moteur. À toute machine, il faut un moteur. Chez Aristote, le moteur était présent : c'était l'âme, située dans le cœur, plus chaud que l'organisme ; le cerveau ne servait qu'à refroidir. Animal : animé : âme sensible. Chez Descartes, c'était un autre dualisme : l'âme spirituelle siège dans l'organe pinéal ; le cœur réchauffe encore l'organisme par les « esprits animaux ». Au XIX^e siècle, ce sont l'hérédité et les besoins (d'où naissaient les habitudes) qui étaient le moteur du comportement. Dans la première moitié du XX^e siècle, tant pour l'éthologie que pour la psychanalyse, ces besoins impérieux de l'hérédité prennent la forme de pulsions.

Métaphores énergétiques.

« *Instinct* : système inné ou adapté de coordination interne du système nerveux dans son ensemble qui, lorsqu'il est activé, trouve expression dans un comportement culminant en un schème d'action héréditaire. Il est organisé sur une base hiérarchique, tant du côté afférent que du côté efférent.

Lorsqu'il est "chargé", il révèle l'existence d'un potentiel d'action spécifique et d'une prédisposition au déclenchement par un déclencheur comportemental.

Activité de déplacement : activité résultant de l'activation par surcharge (potentiel d'action spécifique) d'un ou plusieurs instincts. Il semble apparaître lorsque un (ou des) instinct (s) ne reçoit pas l'occasion de la décharge adéquate par son ou ses actes consommatoires. » (Thorpe)

Freud écrit en écho :

« La puissance du ça exprime la finalité propre de la vie de l'individu (...).

Nous donnons aux forces qui agissent à l'arrière-plan des besoins impérieux du ça et qui représentent dans le psychisme les exigences d'ordre somatique, le nom de *pulsions* (...).

Nous avons appris que les pulsions peuvent changer de but (par déplacement) et aussi qu'elles sont capables de se substituer les unes aux autres, l'énergie de l'une pouvant se transférer à une autre. » (*Abrégé*, p. 7).

En réalité, si les bases conceptuelles de l'éthologie et de la psychanalyse sont différentes, voire incompatibles, les règles de l'assemblage machinique, les composantes de base (non compris leur *sens* conceptuel) sont identiques.

- 1/ un organicisme déterritorialisé, abstrait par rapport à l'organicisme du XIXe siècle.
 - Chez Freud, les topiques.
 - Chez les éthologistes, la description formelle de l'éthogramme, reflet du dispositif formel de prise de connaissance, reflété dans la réification de l'objet décrit.
 - 2/ L'origine pulsionnelle des conduites.
- Il n'y a plus, comme au XIXe siècle, un sujet, réglant ses habitudes en fonction de ses besoins.

Mais il reste encore un diable machiniste caché dans les coulisses.

La Cybernétique

L'apparition de la cybernétique et des ordinateurs va faire disparaître ce qu'il restait encore d'animal (d'animé) dans les machines comportementales, pour ne garder que l'aspect formel.

Bateson :

- 1/ L'esprit est un agrégat de parties ou de composants en interaction.
- 2/ L'interaction entre les parties de l'esprit est mise en action par la différence, et la différence est un phénomène insubstantiel délocalisé dans l'espace et le temps ; la différence s'apparente à la négentropie et à l'entropie plutôt qu'à l'énergie.
- 3/ Les processus mentaux requièrent une énergie collatérale.
- 4/ Les processus mentaux requièrent des chaînes de détermination circulaires (ou plus complexes).
- 5/ Dans les processus mentaux, les effets de différence doivent être vus comme des transformations (c'est-à-dire des versions codées) d'événements les ayant précédé. Les règles d'une telle transformation doivent être relativement stables (c'est-à-dire plus stables que le contenu) mais sont elles-mêmes sujettes à transformation.
- 6/ La description et la classification de ces processus de transformation révèle une hiérarchie de types logiques immanents au phénomène.

L'intérêt de cette définition de l'esprit est de ne pas être une définition de l'esprit mais de la cybernétique. Elle est fondamentalement valable pour n'importe qui dès lors que l'on présuppose qu'un formalisme cybernétique s'y applique. Elle n'est immanente que du contexte culturel dont elle émane.

Le problème de l'analogie entre l'homme et la machine est étroitement lié à des symbolismes culturels et linguistiques associés à cet objet technique. Toute assimilation – admise ou refusée – de l'homme à la machine a rapport à des modalités de signification centrées principalement sur la notion de simulation.

Je vais vous donner quelques jalons, qui vont de l'animal-machine de Descartes à la question de la pensée des ordinateurs.

Ce n'est pas l'objet technique qu'est la machine qui est en question ici, mais son fonctionnement comme modèle, comme métaphore à l'intérieur d'un certain type de discours. Ce qui est en jeu, c'est la liaison de la machine à des fonctions radicalement humaines ; la question de l'être humain a toujours été considérée dans une perspective mécaniste de savoir s'il était une machine, et la machine a toujours été considérée comme simulacre du vivant.

À la fin du XVII^e siècle, les dictionnaires s'accordent sur les trois sens suivants du mot « machine » :

1. Instrument servant à transformer une force naturelle (engin).
2. agencement de parties fonctionnant par lui-même (machinerie).
3. Invention, ruse, moyens mis en œuvre en vue d'une fin (machination).

Le second sens de ce mot est une innovation par rapport à la conception antique de la machine, dans laquelle la source d'énergie est extérieure à la machine.

Les dictionnaires de la Renaissance indiquent le sens « ruse » comme prépondérant. Mais à l'époque classique, Furetière écrit de la machine :

« Engin, assemblage de pièces fait par l'art des mécaniques, qui sert à augmenter la vertu des forces mouvantes. On donne le nom de machine en général à tout ce qui n'a de mouvement que par l'artifice des hommes. »

À l'époque de Descartes, l'accent est mis sur l'artifice, sur l'ingéniosité qui soumet le réel. La machine apparaît comme une ruse du faible, comme un appareil à produire des illusions (en particulier au théâtre). Elle ruse avec la nature pour produire l'illusion du réel ; elle est, par excellence, ce qui sert à la simulation.

Pour Descartes, la machine simule tous les phénomènes en tant qu'ils se déroulent automatiquement, c'est-à-dire sans intervention de l'âme. D'où l'homologie bien connue entre l'animal et la machine, entre le corps (privé d'âme) et la machine.

En se calquant sur les gestes de l'homme, la machine représente en retour un modèle pour comprendre comment fonctionne le corps.

Le mécanisme de Diderot est différent de celui de Descartes. La machine vivante est formée de la réunion du corps anatomique et du corps pensant, souffrant, jouissant. Elle ne se réduit pas à une machine affublée d'une âme. Le réel est entièrement mécanique, ce qui entraîne qu'il est objet de science. Le vivant est une machine d'une extrême complexité, mais c'est sa machinicité même qui le rend intelligible. Diderot associe mécanisme et raisonnement : le syllogisme fonctionne comme une machine et la machine comme un syllogisme.

On voit apparaître ici la possibilité que la machine simule la pensée, le raisonnement. Il est remarquable que Diderot ne décrit jamais la machine, mais seulement son fonctionnement.

Le sens de ruse-machination du mot machine disparaît à la fin du XVIII^e siècle. Pendant l'ère industrielle – triomphe du machinisme – on ne trouve pas trace du sens de machination, comme si la réalité de la machine occultait tout imaginaire.

Au XIX^e siècle, la machine devient un instrument de domination entre les mains des puissants.

Cependant, depuis la machine arithmétique de Pascal (1640), les machines à calculer se développaient mais elles n'avaient pas de mémoire et n'étaient pas programmables. Vers 1850, Charles Babbage invente ce qui est considéré comme le véritable ancêtre des ordinateurs, la Machine

Analytique. Cette machine possède une mémoire (store), une unité de calcul et de décision (mill), et des programmes sur cartes perforées. Elle est bien sûr entièrement mécanique. Lady Ada Lovelace écrit : « The Analytical engine Weaves algebraic patterns just as the Jacquard-loom weaves flowers and leaves. » (Cette machine ne fut jamais réellement construite). Mais elle écrit aussi : « La Machine Analytique n'a pas la prétention de donner naissance à quoi que ce soit. Elle peut faire tout ce que nous savons lui ordonner de faire. »

Les machines actuelles peuvent être considérées comme un système de connexions avec un organe qui joue le rôle d'un cerveau sensori-moteur.

En fait, bien avant la naissance du premier ordinateur vers 1940, la logique mathématique s'était longuement intéressée au problème de mécaniser la pensée (i. e. ce qui nous distingue des autres espèces). Chez Aristote, ce sont les syllogismes, chez Euclide, la géométrie. Dans les deux cas, ce sont des modèles déductifs. Au XIX^e siècle, les travaux de Boole et de De Morgan, puis ceux de Ege, Peano et Hilbert vont dans le même sens.

Les logiciens et les mathématiciens voulaient formaliser la notion intuitive de calculabilité (qui ne tient pas compte des contraintes matérielles). Plusieurs formulations équivalentes furent proposées en 1936, mais nous ne nous intéressons qu'à celle qui fait intervenir la notion de machine. Alan Turing invente un objet mathématique qui s'appellera « machine de Turing ».

Un nombre est dit « calculable » si son développement décimal peut être écrit par une machine. Turing écrit : « We may compare a man in the process of computing a real number to a machine which is only capable of a finite number of conditions... »

Les machines de Turing sont capables de calculer des fonctions : une fonction $f(x_1, x_2, \dots, x_n)$ est effectivement calculable s'il existe une procédure mécanique pour déterminer la valeur $f(k_1, k_2, \dots, k_n)$ quand les k_i sont donnés (par exemple, l'addition).

Cette procédure mécanique s'appelle un algorithme, c'est-à-dire une procédure mécanique déterministe qui, appliquée à une certaine classe d'entrées symboliques (input) fournit éventuellement un symbole de sortie (output).

entrée -----> boîte noire -----> sortie

Exemple d'algorithme : - l'addition de deux nombres entiers

- écrire les deux nombres l'un en dessous de l'autre, alignés à droite

- prendre la colonne la plus à droite

- faire la somme arithmétique de ces deux chiffres

- si le résultat est inférieur à 9, l'écrire en dessous. Donner à R la valeur 0

- si le résultat est supérieur à 9, écrire en dessous le chiffre égal à somme - 10. Donner à R la valeur 1

- passer à la colonne immédiatement à gauche. Ajouter R aux deux chiffres

- recommencer jusqu'au bout.

Un algorithme est donc

1/ un ensemble d'instructions de taille finie.

2/ un agent de calcul.

3/ des mémoires (pour effectuer, stocker).

4/ un temps discret (pas de continu).

5/ déterministe et mécanique (pas de hasard).

Ceci est similaire à un ordinateur : 1/ programme. 2/ circuiterie et logique. 3/ mémoire. 4/ digital. 5/ mécanique.

Autres exemples d'algorithmes : recette de cuisine ; fichetricot. Il faut bien distinguer l'algorithme lui-même des entrées et des sorties (ingrédients-recette-gâteau).

Une machine de Turing comporte :

- une unité centrale, qui peut prendre un certain nombre fini d'états internes.
- une bande infinie dans les deux sens (mais finie à tout moment) sur laquelle sont inscrites au départ les données à traiter, et sur laquelle viennent s'inscrire les calculs intermédiaires (ceci entraîne que l'on a une mémoire potentiellement infinie).
- une tête de lecture-écriture qui n'opère que sur une seule case à la fois et qui peut : a/ remplacer le symbole lu par un nouveau symbole. b/ déplacer éventuellement la bande d'une case à droite ou à gauche. c/ s'arrêter (ne rien faire).

Dans une situation déterminée par un état interne et par un symbole lu sur la bande, la machine change d'état et agit sur la bande.

Il y a un nombre fini de règles de transition, et, dans une situation donnée, une seule évolution possible, c'est-à-dire que la machine est déterministe. La machine produit son propre temps, qui est discret.

On associe un algorithme à une machine : un mot est imprimé sur la bande. La tête démarre sur la lettre du mot la plus à gauche, la machine étant dans un certain état initial. Si la machine stoppe, le mot écrit sur le ruban à ce moment là est la valeur de l'algorithme.

L'intérêt essentiel des machines de Turing est qu'elles permettent de manipuler la notion de décidabilité. On dit qu'une proposition P est décidable si le problème de décider si P est vraie ou fausse est récursivement résoluble, c'est-à-dire s'il existe un procédé déductif mécanique (une machine de Turing) qui donne la réponse vrai ou faux après un nombre fini de calculs.

Une théorie est décidable s'il existe une machine de Turing qui peut, pour chaque formule, dire en un temps fini si cette formule appartient ou non à la théorie.

Le résultat le plus fameux à propos d'indécidabilité est sans doute le théorème de Gödel (1931) qui dit : dans toute formulation axiomatique consistante de l'arithmétique, il y a des propositions indécidables. La preuve de ce théorème repose sur l'écriture d'un énoncé mathématique auto-référent, de la même manière que le paradoxe d'Epiménide le Crétois est un énoncé auto-référent du langage. Le problème est que les énoncés de l'arithmétique portent sur des nombres, et ne sont donc pas auto-référents, i. e. ne portent pas sur des énoncés.

Gödel a eu l'idée de coder par un certain nombre (dit de Gödel) tout énoncé de l'arithmétique. Ainsi, un énoncé de l'arithmétique peut être compris à deux niveaux différents : 1/ comme énoncé de l'arithmétique. 2/ comme énoncé à propos d'énoncés de l'arithmétique.

Ensuite Gödel considère l'énoncé suivant (G) : « Cet énoncé n'a pas de preuve dans ce système. » Cet énoncé (G) n'est pas prouvable dans le système, mais vrai. Ainsi la démontrabilité est plus faible que la vérité.

Revenons aux machines de Turing. Le théorème de Gödel a comme conséquence qu'il n'existe pas de machine de Turing qui puisse, au bout d'un temps fini, dire si une machine de Turing quelconque peut s'arrêter. Autrement dit, aucun programme ne peut, pour un programme quelconque, dire si celui-ci s'arrêtera ou non.

Vers 1940, Turing et d'autres commencèrent à concevoir les ordinateurs (appelés calculateurs), qui sont à la convergence de trois théories : la théorie du raisonnement axiomatique ; l'étude du calcul mécanique ; la psychologie de l'intelligence.

Intelligence ! Vers 1950, ce qui agite beaucoup les logiciens, les mathématiciens et les philosophes, c'est la question suivante « les machines peuvent-elles penser ? »

Alan Turing, dans un célèbre article de 1950, imagine un jeu de simulation, de deux individus, lequel est humain et lequel est une machine ?

Le théorème de Gödel a été souvent utilisé pour tenter de montrer que l'esprit humain ne peut se réduire à la machine. Ainsi Lucas déclare que dans toute machine, il y a une proposition

indémontrable, qui ne peut donc pas être écrite par la machine. Celle-ci ne connaît donc pas cette proposition. Or, l'humain connaît une proposition indémontrable.

Cet argument a été réfuté plusieurs fois. En voici une due à Hofstadter :

« Un homme un jour rencontre un objet inconnu, une femme. Il n'en a jamais vue, et il est d'abord émerveillé par sa ressemblance avec lui. Mais au bout d'un certain temps, il dit aux autres hommes : "Je peux regarder son visage, ce qu'elle ne peut pas faire. Donc elle n'est pas semblable à moi !" D'où la supériorité des hommes sur les femmes (incidemment, le même argument montre que cet homme est supérieur aux autres hommes, mais il ne le mentionne pas. La femme répond : "Mais moi, je peux voir ton visage, ce que tu ne peux pas faire. Nous sommes donc égaux." Et il lui répond : "Mais tu te trompes si tu penses que tu peux voir mon visage. Ce que vous les femmes faites n'est pas ce que nous les hommes faisons. C'est d'une qualité inférieure et ne peut porter le même nom. Tu peux l'appeler voir-femme. Maintenant le fait que tu peux voir-femme mon visage n'a pas d'importance, parce que la situation n'est pas symétrique. Tu vois ?" "Je vois-femme", répliqua-femme la femme, et elle s'en alla-femme". »

Dans son article de 1950, Turing jetait les bases de ce que l'on appelle l'Intelligence Artificielle. Il prévoyait des machines qui pourraient voir, parler, etc.

Hofstadter poursuit cette approche. Pour lui, le cerveau est isomorphe à un système formel. Tout aspect de la pensée peut être considéré comme une description de haut niveau, d'un système qui, à un niveau inférieur, est gouverné par des lois simples et formelles. Ce système est le cerveau. Le niveau logiciel, visible est informel, tandis que le niveau matériel, caché, est un mécanisme complexe qui effectue des transitions d'état en état, selon des lois physiquement implantées, et selon l'impact de signaux.

Le but de l'Intelligence Artificielle est de décrire tout ce qui se passe au niveau supérieur. Thèse : au niveau supérieur, les procédés mentaux sont récursifs. Tout processus mental peut donc, pour l'I.A. être simulé par un programme d'ordinateur dont le langage sous-jacent contient les fonctions partielles récursives. Les machines vont évoluer de plus en plus près des mécanismes du cerveau (cf. les récents développements de la neuroinformatique).

Les machines ont servi de modèle à la pensée ; le cerveau et son fonctionnement sert maintenant de modèle pour la construction de machines.