

Quand l'intelligence artificielle s'inspire de l'intelligence naturelle

"Je propose de réfléchir à la question : Les machines peuvent-elles penser ?"

C'est le mathématicien Alan Turing qui pose cette question et tente de la résoudre dans un article intitulé "les Machines à calculer et l'intelligence" publié dans la revue Mind (cf. D. Hofstadter, Gödel Escher Bach, p.665 et suivantes).

Dans cet article, il décrit une sorte de jeu, "le jeu de l'imitation" connu plus tard sous le nom de "test de Turing", qui se joue à trois : deux hommes et un ordinateur. Le premier homme est dans une pièce séparée et il peut poser des questions aux deux autres joueurs, par le biais d'un clavier ^{écran} ~~de terminal~~ ou d'une autre interface. Le jeu consiste à deviner lequel des deux autres joueurs, le second homme ou l'ordinateur, répond à la question. Si le joueur est incapable de discerner ses deux interlocuteurs, le "test de Turing" est dit positif.

Cette volonté de simuler l'intelligence ne date pas d'hier. Depuis les origines de l'informatique, les ordinateurs ont évolué en devenant de plus en plus "intelligents".

Et bien plus tôt, dès l'antiquité, les hommes ont essayé de construire des machines à leur image, avec les apparences de la vie et de l'intelligence... L'histoire, la mythologie, les traditions,

les contes populaires fourmillent de tels objets : automates de Vaucanson, Pinocchio, golem et autres statues animées.

Or la principale objection que l'on puisse faire à tous ces systèmes (du moins ceux qui ont effectivement été réalisés : les automates) est la même que pour les programmes informatiques : ils ne possèdent pas les caractéristiques essentielles de l'intelligence (cf. D. Hofstadter, p.29-30) :

- " - réagir avec souplesse aux situations qui se présentent ;
- tirer profit de circonstances fortuites ;
- discerner le sens de messages ambigus ou contradictoires ;
- juger de l'importance relative de différents éléments d'une situation ;
- trouver des similitudes entre des situations malgré les différences qui peuvent les séparer ;
- établir des distinctions entre des situations malgré les similitudes qui les rapprochent ;
- synthétiser de nouveaux concepts à partir d'anciens concepts assemblés différemment ;
- trouver des idées nouvelles."

Nous allons voir comment les ordinateurs ont évolué, afin de se rapprocher de plus en plus de cette définition de l'intelligence. Dans cette évolution, nous distinguerons trois étapes, à chacune desquelles correspond une forme différente des matériels ou des logiciels. Nous sommes aujourd'hui à l'aube de la troisième étape.

FIGURE 1.

(1) Première étape : l'ordinateur-calculateur (en anglais "computer").

L'informatique a d'abord servi à calculer. Tous les logiciels classiques, qu'il s'agisse de tableurs, de simulation, de tri, classification, traitement de texte, conception, dessin... ne font rien d'autre que manipuler des nombres.

(2) Deuxième étape : Apparition de "l'intelligence artificielle".

Ce que nous appelons intelligence artificielle (terme inventé par John MacCarthy dans les années cinquante), au contraire, porte sur des symboles qui peuvent être des faits, des objets, des règles... plus généralement des "connaissances".

Le programmeur ne cherche pas à imiter le fonctionnement de l'intelligence, mais seulement à simuler une réponse intelligente, aussi proche de ce que l'on serait en droit d'attendre de la part d'un être doué d'intelligence, dans un domaine d'expertise donné.

~~Ainsi, le système est conçu comme une "boîte noire", du point de vue de l'utilisateur : le logiciel ne contient que des connaissances, telles qu'elles ont été entrées par un homme (expert), et un programme pour enchaîner ces connaissances. Quant au matériel, c'est un ordinateur d'architecture classique.~~

(3) Troisième étape : machines neuromimétiques.

Cette nouvelle génération d'ordinateurs a un fonctionnement directement inspiré de celui du cerveau. Même la terminologie de ces machines est empruntée ^(peut-être abusivement) à la biologie : on parle de neurones et de synapses, comme pour le système nerveux.

On les appelle aussi aussi "réseaux neuronaux", "mémoires associatives" ou "systèmes connexionnistes". Nous verrons, par la suite, la raison de ces dénominations.

Les informaticiens qui les étudient espèrent, avec ces systèmes, résoudre des problèmes sur lesquels l'intelligence artificielle bute encore.

Pourquoi l'intelligence artificielle ?

L'informatique classique semble être dépassée. Parce que nous nous rendons compte que la puissance de calcul, c'est très bien, mais ce n'est pas suffisant. Il faut toujours qu'il y ait un homme derrière la machine, pour interpréter les résultats, prendre les décisions, donner une opinion, intervenir dans les cas critiques, arrêter la machine ou au contraire en mettre une autre en route.

L'intelligence artificielle a été annoncée comme la panacée : selon ses partisans, elle apporterait la possibilité de gérer une entreprise, une usine, d'animer des robots, de faire face à la plupart des problèmes complexes, sans nécessiter d'intervention humaine.

La notion de "cerveau électronique", expression un peu inquiétante que l'on employait au début de l'ère informatique, redevient d'actualité et le mythe de l'ordinateur tout-puissant, qui tel Hal (2001, l'Odyssée de l'espace, Arthur Clarke) arrive à dominer l'homme, refait surface.

Pour le moment, heureusement, les fonctions de ces machines dites intelligentes sont encore limitées aux systèmes experts, programmes d'aide à la conception, à la décision et autres "X intelligemment assisté par ordinateur".

Ces systèmes sont, en effet, assez efficaces pour nous aider dans diverses tâches, comme si nous avions à notre disposition, quand nous le souhaitons, un expert dans le domaine qui nous intéresse. Ils semblent doués d'intelligence, nous donnant l'illusion de nous trouver en face d'un interlocuteur aussi humain que nous, mais plus savant, plus rapide, plus efficace pour résoudre les problèmes qui nous sont posés.

Comment un système informatique résout un problème

FIGURE 2.

(1) En informatique classique, le système (un programme, en l'occurrence) explicite entièrement la solution d'un problème, par une séquence d'instructions qui seront exécutées ^{par la machine} dans l'ordre voulu par le programmeur (algorithme).

(2) Un système d'intelligence artificielle (~~moteur d'inférence + base de connaissances~~) contient les éléments pour résoudre une catégorie de problèmes. L'ordre d'exécution des différentes étapes du programme n'est pas précisé à l'avance. Un tel programme est heuristique (il trouve par lui-même la séquence de connaissances à enchaîner).

(3) Un système neuromimétique est capable d'apprendre à résoudre différents problèmes. Une fois qu'il a appris un problème, il peut résoudre tous les autres problèmes de cette catégorie, ou suffisamment voisins, en généralisant à partir de cet exemple.

Nous allons voir comment il procède.

Pour simuler le raisonnement humain, la première chose à faire est d'observer son fonctionnement. C'est cette réflexion, rassemblant des spécialistes de diverses disciplines (biologistes, neurophysiologistes, physiciens, mathématiciens...) qui est à l'origine d'une nouvelle espèce d'ordinateurs : les systèmes *dits* "neuromimétiques", c'est-à-dire qui imitent les neurones, les cellules nerveuses constituant le cerveau.

On a d'abord essayé d'assimiler le fonctionnement du cerveau à celui d'une machine. C'est la démarche de la cybernétique, science initiée par Norbert Wiener en 1948, dans un ouvrage intitulé Cybernetics or Control and Communication in the Animal and the Machine.

Ces travaux ont donné naissance à toute une variété d'"animaux cybernétiques", sortes d'automates doués de certaines facultés qui, jusqu'alors, semblaient caractériser le règne vivant : la capacité d'adaptation à l'environnement, l'apprentissage, l'autorégulation, dont une application est l'"homéostasie" (faculté de maintenir constantes les conditions de la vie, par exemple la température interne du corps chez les animaux homéothermes).

Bref retour à l'I.A.

Ces recherches, qui paraissaient prometteuses, ont pourtant été interrompues au profit de l'intelligence artificielle (la deuxième étape ~~des figures 1 et 2~~). Celle-ci ne remettait nullement en cause la structure des ordinateurs. Il s'agissait simplement de concevoir d'autres types de logiciels, dont le mode de fonctionnement différait des programmes classiques (algorithmiques). Et les progrès de la technologie mettaient à la disposition des informaticiens des machines suffisamment puissantes et des mémoires assez spacieuses pour les exécuter.

Les plus répandus et les seuls programmes qui soient aujourd'hui réellement opérationnels, en intelligence artificielle, sont les systèmes experts.

Evoquons, pour mémoire, certaines de leurs caractéristiques essentielles :

- Leur fonction est de simuler le comportement d'un expert dans son domaine.
- Ils sont constitués de deux parties indépendantes : le moteur d'inférences (raisonnement) et la base de connaissances (de même qu'en informatique classique on peut distinguer, d'une part, le programme ou séquence d'instructions, et, d'autre part, les données).
- Grâce à cette séparation entre raisonnement et connaissances, il est possible, à n'importe quel stade du traitement, de rajouter des connaissances, sans pour autant avoir à reprendre la séance depuis le

début. Pour cela, le système dispose généralement d'un module d'acquisition de connaissances. Les systèmes les plus élaborés sont capables de restructurer leurs connaissances en fonction des nouvelles acquisitions (métaconnaissances).

- Contrairement aux programmes classiques, les systèmes experts sont souvent capables de fournir une conclusion sensée, même si leurs données sont incomplètes.

Bien qu'il soit beaucoup plus souple qu'un programme classique, un système expert n'est capable que d'un type de "raisonnement" donné (celui qui lui est fourni par son moteur d'inférence), et ne peut se fonder que sur les connaissances qu'un ou plusieurs experts ont bien voulu lui inculquer. Sa base de connaissances, si grande soit-elle, ne peut jamais être exhaustive. Contrairement à un expert humain, l'efficacité d'un système expert cesse brutalement dès que l'on sort de son domaine de compétence ou qu'un cas nouveau apparaît.

Cette limitation ne peut être palliée que par la présence d'un homme susceptible d'intervenir dès que le système butera sur un problème qu'il ne peut résoudre faute de connaissances.

Cet inconvénient n'est généralement pas trop gênant, dans la mesure où les systèmes experts, comme leur nom l'indique, sont limités à un domaine d'expertise, donc relativement facile à cerner : dans les applications financières, la base de connaissances comprend des règlements, des critères... ; dans les applications industrielles, il suffit généralement de connaître la structure des machines, le mode de fonctionnement et le scénario des pannes les plus fréquentes ; en médecine, on connaît les symptômes des maladies les plus fréquentes ; etc.

C'est pourquoi un système expert portant sur un domaine restreint et bien délimité, comme par exemple le diagnostic de pannes d'une machine ou de maladie, l'identification d'une plante, d'un objet, la recherche des démarches à effectuer pour trouver un emploi ou créer une entreprise... peut être très efficace, voire surpasser les possibilités d'un expert humain, dans 80 ou 90 % des cas (les 10 ou 20 % restants étant des cas d'exception ou critiques, qui doivent être résolus individuellement par l'expert humain).

Mais d'autres applications de l'intelligence artificielle sont beaucoup plus délicates à mettre en oeuvre. Parmi les plus importantes, citons la traduction automatique, certains jeux de réflexion (échecs, tours de Hanoï, backgammon...), la reconnaissance de formes, la compréhension du langage naturel.

Pour les jeux, le problème est en général assez facile à résoudre théoriquement, du moment que l'on connaît les règles du jeu et que le nombre de pièces et de positions est limité. Mais l'inconvénient est que le nombre de combinaisons devient vite exorbitant, aussi faut-il se limiter à un nombre réduit de coups, et l'ordinateur est alors généralement battu par les bons joueurs.

La situation est encore bien plus grave dans les autres domaines cités, notamment la reconnaissance de formes et le traitement du langage naturel. Dans ces cas, les connaissances requises sont non seulement très nombreuses, mais elles ne sont pas dénombrables, elles sont diffuses, parfois subjectives, incertaines, incomplètes...

Prenons, par exemple, la reconnaissance de formes : un visage, des éléments d'un paysage, une écriture manuscrite. La reconnaissance est facile s'il s'agit de reconnaître quelque chose ou quelqu'un parmi un

nombre limité de candidats, d'objets. Dans ce cas, il s'agit de comparer la forme à identifier aux différents modèles représentant ces candidats, et d'opter pour le plus proche voisin. Mais ce cas est exceptionnel.

D'une façon générale, ce type de problème présuppose un grand nombre de connaissances acquises, consciemment ou inconsciemment, souvent depuis la naissance, voire auparavant, qui englobent un patrimoine génétique, un contexte culturel, un environnement, l'expérience vécue, pensée, rêvée, la ou les traditions, etc., de sorte qu'il ne nous est même pas possible de savoir si l'ensemble des connaissances que nous devons posséder pour le résoudre est très grand ou infini.

Alors que toute machine (ordinateur, système expert ou autre programme d'intelligence artificielle) traite un ensemble bien déterminé et délimité d'objets (des nombres, des données, des connaissances, des formes...) - système fermé - , l'intelligence humaine ne sait pas toujours, a priori, à quel ensemble de connaissances elle devra faire appel - système ouvert.

Une machine, si elle est confrontée à un objet sortant de son "univers", reste muette ou fournit une réponse aléatoire, absurde. L'homme est capable "d'inventer" une solution, d'improviser.

Apprentissage et association

L'intelligence humaine, donc, est capable de sortir de tout "univers" prédéfini, en apprenant, en s'adaptant, en imaginant, dès qu'elle rencontre l'inconnu. Ou, du moins, dès qu'elle peut relier cet

"inconnu" à quelque chose de déjà connu. De nouvelles liaisons s'établissent dans notre cerveau. Car c'est précisément ainsi qu'il fonctionne - pour autant que nous puissions connaître ce fonctionnement, la connaissance sur le cerveau étant encore incomplète et controversée, et les résultats des recherches pouvant être interprétés différemment.

Rappel sur le cerveau :

FIGURE 3.

Structure d'un neurone

Le cerveau humain, comme celui des animaux, est constitué d'un très grand nombre de cellules, les neurones (environ 100 milliards). Ceux-ci sont constitués d'un corps cellulaire à peu près sphérique ou polygonal, duquel partent des centaines de filaments, les dendrites (de forme arborescente, d'où leur nom, dérivé du grec "dendron", arbre), par l'intermédiaire desquelles des connexions peuvent s'établir avec les neurones voisins. Une autre excroissance, l'axone, beaucoup plus longue que les dendrites puisqu'elle peut atteindre quelques dizaines de centimètres, représenterait le tronc ; plus épais aussi, l'axone est constitué d'une fibre nerveuse entourée d'une couche isolante. À son extrémité, elle se ramifie en un réseau de fibres qui transmettent les signaux aux dendrites des cellules voisines. L'espace que l'influx nerveux doit franchir pour passer d'un neurone à l'autre est appelé synapse (du grec "sunapto", j'attache). Il peut y avoir jusqu'à 10 000 entrées synaptiques pour un neurone.

L'influx nerveux a une double nature : signal électrique à l'intérieur de la cellule, il est converti en un processus chimique lorsqu'il franchit une synapse.

Le processus de conduction du signal est le suivant : le neurone collecte sous forme de potentiels et de courants électriques, au niveau des dendrites, les signaux qui lui parviennent des neurones voisins. Ces potentiels se propagent par conduction le long de l'axone jusqu'à ses terminaisons. Arrivé au niveau des synapses, le signal électrique est transmis par réaction chimique aux neurones voisins, par l'intermédiaire de leurs dendrites, où de nouveaux potentiels sont créés. C'est donc un processus de "reconnaissance" chimique qui permet aux cellules du cerveau de se connecter l'une à l'autre en créant des liaisons synaptiques.

Mémoire associative

L'apprentissage consiste à établir une série de connexions, traçant ainsi un "chemin" à travers le réseau de neurones. Chaque fois que le même objet ou événement est rencontré, ce chemin sera consolidé, ou des voies plus ou moins proches sont créées. Dès lors, le premier élément de connaissance sera toujours associé, dans notre mémoire, au second. Un tel mécanisme porte le nom de "mémoire associative".

La faculté d'association est, en effet, une caractéristique essentielle de la pensée humaine. Notre cerveau fonctionne par "association d'idées". La formation d'un concept ou d'une

représentation dans le cerveau résulte d'un couplage entre différents neurones ou groupes de neurones. Les connaissances sont ainsi encodées sous forme de "traces" dans les connexions, lesquelles sont établies ou modifiées au cours de l'apprentissage.

Lorsque nous effectuons une addition, par exemple " $2 + 2 = 4$ ", notre cerveau ne va pas chercher deux chiffres dans différentes cases pour effectuer l'opération "addition" et ranger le résultat dans une autre case. En fait, le cerveau ne fait qu'associer deux "objets" : "2 et 2" d'une part, et "4" d'autre part.

Un souvenir est rappelé à la mémoire lorsque notre pensée suit certains cheminements, dont le point de départ peut être une sensation qui avait déjà été éprouvée lors de l'événement dont nous nous souvenons. Le philosophe et biologiste William James avait déjà remarqué, à la fin du siècle dernier, que "quand deux processus cérébraux élémentaires ont été actifs simultanément ou l'un après l'autre, l'un d'eux, lorsqu'il se produit, tend à propager son excitation à l'autre."

Tout cela paraît, somme toute, assez simple. Et pourtant nous avons bien du mal à comprendre comment le cerveau peut accomplir toutes ses fonctions, aussi complexes que variées. C'est ici qu'intervient le grand nombre de cellules mises en jeu : nous avons dit qu'elles sont plusieurs milliards. On imagine que les interactions entre plusieurs milliards de cellules, chacune pouvant être connectée à plusieurs centaines, voire milliers d'autres, donnent un système extrêmement compliqué : il y a environ 10^{14} relations potentielles entre les neurones d'un homme (soit cent mille milliards).

Remarque :

Bien qu'il comprenne un grand nombre d'éléments actifs, un système neuromimétique diffère fondamentalement dans son fonctionnement de ce que l'on appelle "architectures parallèles" : dans ces machines appelées supercalculateurs, plusieurs processeurs effectuent des traitements simultanément, mais chaque processeur fonctionne de manière classique, selon la conception de Johannes von Neumann. Ce mathématicien d'origine hongroise est, en effet, à l'origine de la conception actuelle des ordinateurs, qui fonctionnent grâce à un programme (logiciel) distinct de la machine (matériel).

Les mémoires associatives, comme le cerveau, ne fonctionnent pas du tout de la même façon. L'étude de ces mémoires intéresse les informaticiens depuis quelques années. L'ingénieur allemand Konrad Zuse, qui d'ailleurs a travaillé dès l'origine sur la conception des calculateurs programmables (en même temps que J. von Neumann, aux Etats-Unis), en donne la définition suivante (1980) : "La mémoire associative est une mémoire dans laquelle on a accès au contenu de la mémoire, non à une adresse. Nous savons que le cerveau humain fonctionne sur ce principe..."

Cette idée que la mémoire est faite d'associations entre concepts est bien plus ancienne, puisqu'on en trouve déjà des traces à l'époque d'Aristote. Elle a fait son apparition en informatique grâce aux recherches effectuées sur l'utilisation d'associations simples pour représenter la signification des mots dans les bases de données. Plus récemment, on l'a utilisé pour représenter les connaissances dans les systèmes experts, sous la forme de "réseaux sémantiques", ou pour structurer les informations dans certains types de programmes.

appelés "hypertexte" ; dans tous ces cas, les informations (données, connaissances, objets, faits, événements...) sont associés dans une structure de forme généralement arborescente.

Mais de telles structures de mémoire étaient jusqu'alors traduites uniquement dans le logiciel. Suivant cette représentation en réseau, les informations constituent les noeuds, et il est possible de cheminer dans le réseau en suivant les mailles, correspondant aux associations entre ces informations. Ces mémoires sont essentiellement passives. Il faut un programme ~~(moteur d'inférence- d'un système expert, logiciel d'hypertexte, etc.)~~ pour les "activer", c'est-à-dire créer et suivre ces séries d'associations qui sont à la base du raisonnement.

Au contraire, dans le système cérébral humain, la mémoire elle-même est l'élément actif. Mémoire, unité centrale de traitement et programme ne font qu'un. Dès lors, il n'y a plus de distinction entre matériel et logiciel. La mémoire résulte d'une organisation de la structure matérielle, et il suffit d'activer un de ses éléments pour faire évoluer tout le système, pour effectuer le traitement.

C'est ainsi, semble-t-il, que fonctionne notre cerveau : une simulation d'un neurone ou d'un groupe de neurones (induite par une sensation - vue, ouïe, toucher, etc.), et tout un mécanisme se met en marche.

Rappel sur les mémoires d'ordinateurs :

Dans les mémoires d'ordinateurs classiques, les informations sont organisées de manière séquentielle, sous la forme de séries de bits ("binary units" prenant la valeur 0 ou 1). Pour faciliter leur accès, ces bits sont regroupés en "mots" de 8 ou 16 bits, parfois plus, rangés en ordre séquentiel dans des cases numérotées par une "adresse". Ainsi, chaque information est accessible, de manière univoque, par cette adresse.

Dans une mémoire associative, les informations ne peuvent pas être localisées à des emplacements déterminés, mais chaque donnée mémorisée est distribuée sur l'ensemble du réseau. Les neurophysiologistes ont montré qu'il en est de même pour le cerveau : nos souvenirs n'ont pas leur siège dans certaines zones cervicales ; ils sont distribués dans toute la structure, si bien que la destruction d'une partie des neurones (accidentellement ou au cours du vieillissement) n'a généralement pas pour effet d'effacer certains souvenirs particuliers, mais plutôt d'altérer globalement toute la fonction de mémorisation.

Nous reconnaissons là une des caractéristiques des hologrammes, sortes de photographies en relief, sur chaque point desquelles est enregistrée l'information provenant de tous les points de l'objet photographié. On a remarqué qu'un fragment d'hologramme fournit la même image que l'hologramme entier, mais avec une perte de qualité d'autant plus grande que le fragment est petit.

Cette analogie holographique du cerveau (mise en évidence et étudiée par le biologiste Karl Pribram) explique la pertinence des systèmes

neuromimétiques à base d'hologrammes (cf. ci-après).

Comme, dans une mémoire associative, les informations ne sont pas localisées, la notion d'adresse n'existe pas. Comment, dans ces conditions, accéder à ces informations ? Les seules entrées possibles sont également des contenus de mémoire. Aussi les mémoires associatives sont-elles aussi appelées "mémoires adressables par contenu" (CAM = "Content Adressable Memory"). Il suffit que la donnée entrée soit associée à celle que l'on veut récupérer en sortie.

Chaque information mémorisée correspond à un état d'équilibre du système neuromimétique. Un modèle géométrique nous aidera à comprendre le fonctionnement de ces mémoires.

FIGURE 4.

Représentation des états d'un système neuromimétique

Représentons l'ensemble des états du système par une surface dans l'espace. On peut la matérialiser par une nappe en caoutchouc tendue. Initialement, avant tout apprentissage (cas idéal, mais ne correspondant pas à la réalité dans le cas du cerveau), la nappe est parfaitement lisse, plane, horizontale ; tous les états sont également probables ("tabula rasa").

L'apprentissage ou la mémorisation d'une information consiste à creuser une "vallée" dans cette surface (en posant, par exemple, un poids sur la nappe). A l'issue de la phase d'apprentissage, le système, suivant notre analogie, se présente comme une nappe de caoutchouc parsemée de creux.

En phase d'utilisation, entrer une information (à reconnaître ou identifier) consiste à lâcher une bille à partir d'un point de cette surface. Si la bille se trouve déjà au fond d'un creux, elle y reste évidemment (ce cas correspond à la présentation d'un modèle identique à l'un de ceux qui ont été mémorisés). Si la bille est lâchée à proximité d'un creux, elle est attirée vers celui-ci et s'arrête lorsqu'elle arrive au fond (modèle proche d'un des modèles mémorisés). Enfin, si elle se situe trop loin d'un creux, ou à égale distance entre deux creux, son évolution est indéterminée.

Ce modèle met en évidence différentes propriétés des mémoires associatives, dont voici quelques-unes :

- Un état mémorisé par un apprentissage préalable devient un "attracteur" (bassin d'attraction) pour tous les états suffisamment voisins.
- La capacité de la mémoire n'est pas illimitée : si deux attracteurs sont trop proches, la classification ne sera plus déterministe.
- La vitesse de classification (de reconnaissance) est indépendante du nombre d'attracteurs (de modèles mémorisés) (contrairement aux programmes classiques de reconnaissance de formes qui sont d'autant plus longs à exécuter qu'ils ont beaucoup de formes en mémoire).

La surface en caoutchouc n'est qu'une image car, comme nous l'avons vu, les informations ne sont pas localisées, mais distribuées sur l'ensemble de la structure. Cette image correspond en réalité aux valeurs de l'énergie du système. Les informations mémorisées correspondent aux points d'énergie minimale, donc des états d'équilibre stable, vers lesquels sont attirés les états d'énergie supérieure.

On distingue les mémoires auto-associatives, qui associent à un modèle présenté en entrée le modèle mémorisé (l'état du réseau) le plus proche, et les mémoires hétéro-associatives, qui stockent une relation entre deux ou plusieurs modèles (états). Dans ce dernier cas, il faut un peu modifier notre image de la nappe en caoutchouc : au lieu des creux attracteurs, ce sont des "vallées" qui entraînent la bille d'un point (ou d'une région) à un autre point de la surface.

Les réalisations de systèmes neuromimétiques

Les systèmes neuromimétiques sont donc une tentative pour reproduire ce fonctionnement. Ces structures sont constituées d'un grand nombre d'éléments simples (les neurones), reliés par des connexions (synapses). Etant donné l'importance de ces connexions, on appelle aussi ces systèmes "réseaux" ou "systèmes connexionnistes".

Ces systèmes ont des propriétés d'autant plus intéressantes que le nombre d'éléments est plus important. Il est, bien sûr, illusoire actuellement d'espérer avoisiner les propriétés du cerveau avec ses dizaines de milliards de neurones, mais des systèmes plus simples se sont avérés déjà tout à fait intéressants.

Les premiers modèles s'inspirant du cerveau apparaissent dans les années cinquante. L'un d'eux, le "perceptron", conçu par Frank Rosenblatt (Cornell Aeronautical Laboratory, 1949), est capable d'apprendre à distinguer plusieurs figures.

En 1960, le biologiste et cybernéticien W.R. Ashby crée le premier système artificiel auto-organisé : l'homéostat. Capable de trouver

par elle-même son équilibre, cette machine "autonome" n'avait aucune destination pratique, mais elle a servi de base à toute une série d'études sur les réseaux de neurones, qui commencent à donner leurs fruits depuis quelques années.

Théorie des neurones formels

Le point de départ de ce regain d'intérêt, ce sont les travaux de l'Américain John Hopfield (California Institute of Technology, 1982). Ce chercheur a développé une théorie des "neurones formels" (l'idée de neurones formels date, en fait, de 1943, avec les travaux de McCulloch et Pitts) : chacun de ces neurones est un élément simple pouvant se trouver dans deux états : activé (1) ou non activé (-1). Chacun d'eux est relié à tous les autres par des connexions (synapses) qui peuvent être excitatrices, inhibitrices ou inexistantes. Si l'on désigne par C_{ij} le coefficient affecté à la synapse entre les neurones i et j , C_{ij} est positif si la synapse est excitatrice, négatif si elle est inhibitrice, et nul si elle est inexistante.

FIGURE 5 :

Réseau de cinq neurones entièrement connectés : chaque neurone est relié à tous les autres par un coefficient synaptique C_{ij} .

Chaque neurone est une sorte d'intégrateur, qui fait la somme de tous les signaux qui lui parviennent, c'est-à-dire de l'état de tous les autres neurones du réseau, pondéré par le coefficient synaptique correspondant. Cette somme est ensuite comparée à un seuil : si elle est inférieure à ce seuil, le neurone est non activé (état -1), si elle est supérieure il est activé (+1).

(Sur la figure, les cercles représentent les neurones avec le seuil correspondant θ_i , et les points représentent les synapses.)

Le modèle de Hopfield comprend en outre la rétroaction : le signal de sortie est réintroduit en entrée, après seuillage, et le processus recommence jusqu'à ce que le système ait atteint un état d'équilibre.

Au départ, toutes les synapses ont un poids égal ("tabula rasa"). Le système est initialisé en activant certains neurones (mis à l'état +1). L'ensemble évolue alors spontanément vers un état stable que nous pouvons considérer comme le résultat du traitement. Dans ce cas, il est insignifiant, puisque le système n'a encore rien appris.

L'apprentissage se fait en modifiant les poids synaptiques. Un contenu de mémoire correspond alors à une certaine matrice synaptique.

En pratique, cet apprentissage peut se faire de manière itérative : on présente à l'entrée l'une des données à apprendre, et l'on calcule les paramètres du réseau, de façon à obtenir, à la sortie, la réponse recherchée ; le processus est réitéré jusqu'à ce que toutes les données aient été mémorisées (ou que la capacité maximale ait été atteinte).

(Léon Personnaz, ESPCI)

Le traitement informatique, dans ces systèmes neuromimétiques, consiste à faire évoluer le réseau, à partir d'une configuration quelconque des neurones, vers une configuration stable, correspondant à une information mémorisée.

Nous avons déjà fait, précédemment, la distinction entre mémoires auto-associatives et hétéro-associatives. Les premières associent à une configuration (donnée, image ou autre information) du système la configuration la plus proche parmi celles qui ont été mémorisées préalablement, au cours de la phase d'apprentissage.

Les mémoires auto-associatives s'appliquent bien à la reconnaissance de formes. Présenter une forme au système équivaut à mettre chacun des neurones dans l'état correspondant à cette configuration. Si cet état a déjà été préalablement appris, le réseau reste inchangé. Sinon, il évolue jusqu'à un état stable correspondant à la configuration la plus proche (attracteur). C'est ainsi qu'il est possible de rétablir une forme incomplète ou erronée, de reconnaître des caractères manuscrits, même déformés.

Les systèmes hétéro-associatifs mémorisent une relation entre deux ou plusieurs configurations ; ils peuvent alors simuler le raisonnement, et notamment résoudre des problèmes typiques d'intelligence artificielle : diagnostic, classification, traduction automatique, par exemple.

Un réseau hétéro-associatif peut même remplacer un système expert, s'il a mémorisé l'association entre une hypothèse et une conclusion. Dans ce cas, l'état présenté au système correspond à une hypothèse ; le système fournit la conclusion correspondante (qui a été associée à l'hypothèse au cours de la phase d'apprentissage).

Implémentations des systèmes neuromimétiques

Les systèmes neuromimétiques peuvent être simulés sur des ordinateurs classiques, mais ils perdent alors un certain nombre de leurs qualités, notamment la rapidité de traitement.

L'optique est particulièrement appropriée pour réaliser des systèmes neuromimétiques. Nous avons déjà souligné l'analogie entre cerveau et hologramme. De plus, l'optique est bien adaptée au traitement parallèle inhérent à ces systèmes.

Demetri Psaltis (California Institute of Technology) et Nabil Farhat (University of Pennsylvania) ont été les premiers à avoir eu l'idée de réaliser des réseaux neuromimétiques optiques, appliquant le modèle de Hopfield.

La première réalisation date de 1984. Elle comprend un alignement de n diodes constituant autant de neurones binaires. Les connexions se font optiquement, par l'intermédiaire d'une matrice à $n \times n$ cellules (la matrice synaptique).

FIGURE 6.

Modèle optique de réseau neuromimétique

Chacune des diodes peut être allumée (neurone actif) ou éteinte (neurone inactif). Chaque élément de la matrice synaptique a une transparence qui peut prendre différentes valeurs, du blanc (transparence totale) au noir (opacité totale), et pouvant être modifiée par l'apprentissage.

La lumière émise par une photodiode de la rangée horizontale, N_i , illumine, via une lentille cylindrique, une colonne de la matrice synaptique. La lumière transmise par cette matrice dépend de l'état des photodiodes émettrices (N_i) et de la transparence de la matrice elle-même. Chaque ligne de la matrice est focalisée, par l'intermédiaire d'une lentille cylindrique, sur une photodiode d'une rangée verticale (N'_j).

Chacune de ces diodes (N'_j) est reliée à la diode émettrice correspondante (N_i) par une boucle de rétroaction non linéaire. Le signal reçu est comparé à un seuil θ_j . Selon le résultat de la comparaison, la diode N_i sera allumée ou éteinte, et le processus est réitéré jusqu'à ce qu'il aboutisse à un état stable de l'ensemble des diodes (neurones).

La matrice synaptique peut être remplacée par un hologramme qui présente l'avantage de comporter un nombre quasi infini de points faisant office de neurones (au lieu des n diodes du dispositif précédent).

Quelques remarques en guise de conclusion

L'intelligence artificielle, telle qu'elle est mise en oeuvre dans les systèmes experts, est une technique (un ensemble de méthodes de programmation) particulièrement adaptée à certaines applications : diagnostic, aide à la décision... dans des domaines où il existe des spécialistes, mais où des algorithmes ne sont pas toujours disponibles : médecine, chimie, géologie, électronique...).

En revanche, on peut assimiler l'autre approche de l'intelligence (celle que nous avons décrite ici) à une science, l'étude de l'esprit humain en vue de reconstituer, à l'aide de moyens artificiels, des raisonnements et des actions intelligentes.

Alors que l'intelligence artificielle classique n'est qu'une approche empirique de l'intelligence, les systèmes neuromimétiques constituent un modèle pour étudier l'intelligence. Et le fait que l'on ait réussi à simuler, avec ces machines, certains aspects de l'activité intellectuelle constitue un encouragement à poursuivre dans la voie d'une investigation biologique des phénomènes cognitifs.

Ressemblances avec l'intelligence naturelle : souplesse, adaptabilité, capacité d'apprentissage...

Un système neuromimétique, à l'instar du cerveau, est capable de généralisation, d'abstraction, et donne l'impression de pouvoir structurer les connaissances sans que cela ait été programmé préalablement.

Nous y sommes presque arrivés, à la fin de ce siècle, et pourtant
cette opinion mérite ~~cependant~~ d'être tempérée.

L'ordinateur n'a accès qu'à la forme, pas au fond ; aux signes qui ne produisent pas de sens ; aux mots et aux relations entre ces mots, pas aux idées qu'ils expriment (cf. Umberto Eco, Le nom de la rose, ou le tableau de Magritte, "Ceci n'est pas une pipe"). L'ordinateur traite des nombres, des images, des symboles, ou plus exactement des représentations convenues de ces symboles ; il ne les comprend pas.

L'autre versant de la culture, celui qui produit le sens, la connaissance (et non une "représentation des connaissances"), l'art, la poésie, la musique... demeure encore inaccessible à toute machine, fût-elle neuromimétique.

Et en dépit de nos louables efforts en recherche informatique, gageons que cette dernière forme d'intelligence restera encore longtemps le propre de l'homme.