

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

Dans l'évolution de l'informatique, nous pouvons distinguer trois grandes étapes, correspondant à des ordinateurs de plus en plus "intelligents".

- La première étape correspond à l'ordinateur-calculateur (en anglais "computer"), que l'on affectait pourtant déjà de propriétés anthropomorphiques en le désignant par "cerveau électronique".
- Le terme d'"intelligence" appliqué à l'ordinateur apparaît dans les années cinquante avec l'"intelligence artificielle", expression due à John MacCarthy. Il serait plus juste de désigner ces programmes par "systèmes de traitement symboliques" ou "systèmes à base de connaissances", car ils ne fonctionnent pas du tout comme l'intelligence. Ces systèmes dits "experts" se rapprochent cependant de l'intelligence humaine par les réponses qu'ils fournissent, aussi proches que possibles de celles que l'on pourrait attendre d'un être doué d'intelligence, dans un domaine d'expertise donné. Au reste, l'ordinateur fonctionne comme un calculateur, mais avec des logiciels plus complexes.
- La troisième étape correspond à des principes de fonctionnement tout différents, car directement inspirés du cerveau. Ces machines dites "neuromimétiques" sont constituées d'un très grand nombre d'éléments simples, les "neurones", qui sont reliés par des connexions ou "synapses". Etant donné l'importance de ces connexions, ces systèmes sont aussi appelés "réseaux" ou "systèmes connexionnistes".

Cette étape se distingue principalement des deux premières par le fait que, comme dans le cerveau humain, mémoire, unité centrale de

traitement et programme sont indissociables. Dès lors, il n'y a plus de distinction entre logiciel et matériel. Ces systèmes fonctionnent, en fait, comme une mémoire active qui, telle la mémoire humaine, est capable d'associer des concepts, des symboles, des idées, d'où leur nom de "mémoires associatives".

Les systèmes neuromimétiques ont été formalisés par John Hopfield (California Institute of Technology) depuis 1982. Selon cette théorie des "neurones formels", un neurone (noeud du réseau) peut se trouver dans deux états : activé ou non activé. Chaque neurone est relié à tous les autres par des connexions ou synapses (mailles du réseau) qui peuvent être excitatrices, inhibitrices ou inexistantes. L'ensemble des états de ces synapses constitue la matrice synaptique.

La mémorisation dans un tel réseau, ou apprentissage, se fait en modifiant les coefficients synaptiques. Un contenu de mémoire correspond alors à une certaine matrice synaptique, et ne peut être localisé en un point du réseau. Il s'ensuit que la notion d'adresse, essentielle en informatique classique, n'existe plus. Comment, dans ces conditions, accéder aux données mémorisées ? Les seules entrées possibles sont également des contenus de mémoire, auxquels on peut accéder en activant ou désactivant des neurones. Aussi les mémoires associatives sont-elles encore appelées "mémoires adressables par le contenu" ou CAM ("Content adressable memory").

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.

On distingue les mémoires auto-associatives et les mémoires 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 d'une phase d'apprentissage. Ils s'appliquent bien à la reconnaissance de formes. Les systèmes hétéro-associatifs mémorisent une relation entre deux ou plusieurs configurations ; s'il s'agit de l'association entre une hypothèse et une conclusion, de tels systèmes peuvent résoudre des problèmes typiques d'intelligence artificielle, à la place des systèmes experts.

Les systèmes neuromimétiques, comme les humains, sont capables d'apprentissage. Loin d'être limités à une catégorie de problèmes, comme les systèmes experts qui sont programmés pour résoudre un type de problème (diagnostic, par exemple) dans un domaine bien limité (chimie, médecine, électronique...), ils sont capables d'apprendre à résoudre différents problèmes, cet apprentissage se faisant généralement par l'exemple. Une fois que le système a appris à résoudre un problème, il peut aussi faire face à tous les autres problèmes de cette catégorie, ou suffisamment voisins.

L'avantage des systèmes neuromimétiques, par rapport aux programmes informatiques classiques, est la prise en compte de la notion de voisinage, de ressemblance, et plus généralement d'ensembles flous, notion fort utile dans les problèmes de reconnaissance de forme. Mais les avantages quasi humains de ces systèmes (plus grande souplesse, adaptabilité, capacité d'apprentissage...) ont leur contrepartie : une moindre fiabilité, moindre répétabilité, relative lenteur de calcul.

Une autre différence essentielle avec l'informatique classique réside dans le fait que l'on ne sait pas comment s'effectue le raisonnement,

dans un système neuromimétique. Bien qu'il puisse fonctionner comme un système expert, aucun cognitivien n'a fourni au réseau neuromimétique les connaissances ou les règles sur lesquelles se fonde son "raisonnement", aucun informaticien n'a écrit de programme pour mettre en oeuvre ce raisonnement (le moteur d'inférence).

Il est certain que l'étude des systèmes neuromimétiques nous aidera à comprendre l'intelligence humaine. Pourtant, si perfectionnée et complexe soit-elle, la machine ne parviendra probablement jamais à égaler certaines caractéristiques de la pensée humaine. En particulier, 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. La machine traite de nombres, des images, des symboles, ou plus exactement des représentations convenues de ces symboles ; elle ne les comprend pas.

Pour conclure, nous tenterons une analogie avec la "loi de complémentarité" que les physiciens atomiques ont appliquée au comportement dualiste onde/corpuscule des objets de la microphysique :

- Soit l'intelligence est adaptée aux calculs, classifications et raisonnements parfaitement rigoureux, dans un domaine bien limité : c'est l'informatique classique, incluant l'intelligence artificielle.
- Soit elle est capable d'apprentissage, d'adaptation, d'autonomie, voire de créativité (au détriment de la précision, de la rigueur et de la vitesse de calcul), dans un domaine illimité : c'est l'intelligence naturelle.

Et en dépit des rapides progrès de l'informatique, gageons que cette dernière forme d'intelligence restera encore longtemps le propre de l'homme.