



## LES MACHINES SONT-ELLES INTELLIGENTES ?

[Jacques Bouveresse](#)

Éditions du Croquant | « Zilsel »

2022/2 N° 11 | pages 399 à 403

ISSN 2551-8313

ISBN 9782365123723

DOI 10.3917/zil.011.0399

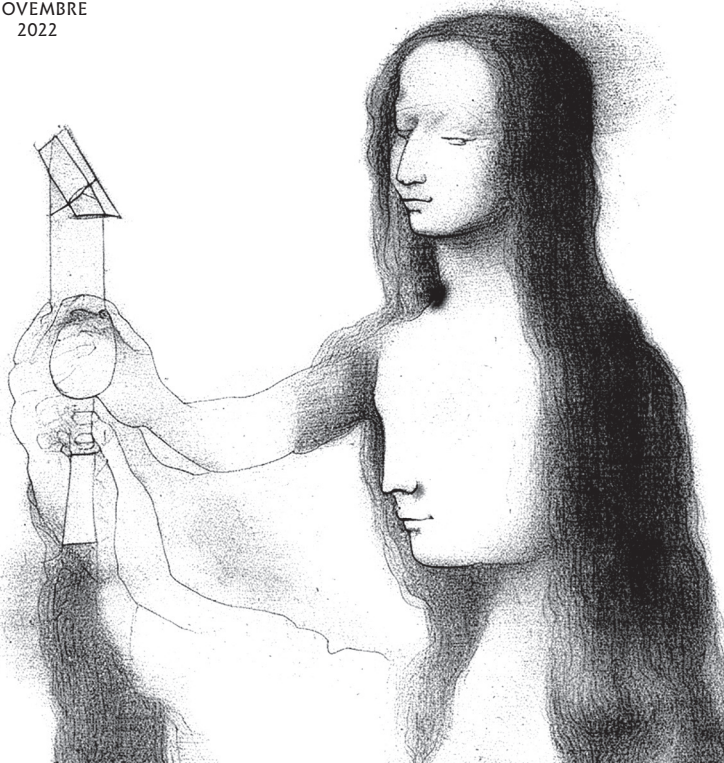
Article disponible en ligne à l'adresse :

<https://www.cairn.info/revue-zilsel-2022-2-page-399.htm>

Distribution électronique Cairn.info pour Éditions du Croquant.

© Éditions du Croquant. Tous droits réservés pour tous pays.

La reproduction ou représentation de cet article, notamment par photocopie, n'est autorisée que dans les limites des conditions générales d'utilisation du site ou, le cas échéant, des conditions générales de la licence souscrite par votre établissement. Toute autre reproduction ou représentation, en tout ou partie, sous quelque forme et de quelque manière que ce soit, est interdite sauf accord préalable et écrit de l'éditeur, en dehors des cas prévus par la législation en vigueur en France. Il est précisé que son stockage dans une base de données est également interdit.



# Les machines sont-elles intelligentes?<sup>1</sup>

Jacques Bouveresse

Un certain nombre de gens (en particulier des philosophes) croient encore aujourd'hui à la possibilité de démontrer que des capacités et des activités humaines de nature diverse ne peuvent, pour des raisons a priori, être reproduites par des machines « intelligentes ». Cependant, les affirmations que l'on croyait pouvoir proférer sur les prétendues limitations intrinsèques des machines ont été si souvent réfutées par des progrès ultérieurs que la plus grande prudence devrait, de toute évidence, s'imposer désormais dans ce domaine. Même si les ordinateurs sont encore loin aujourd'hui d'avoir confirmé les pronostics et réalisé les espérances enthousiastes de la première heure, il n'en reste pas moins qu'ils sont en mesure de faire et font couramment des choses que la philosophie défaitiste des sceptiques n'a jamais été en mesure de rêver. Contrairement à ce que l'on entend malheureusement encore affirmer chez les profanes, ils sont capables de performances exceptionnelles dans des jeux comme les dames ou le backgammon, ils sont en train de devenir des joueurs d'échecs plus qu'honorables, ils peuvent produire des démonstrations inédites et surprenantes de théorèmes mathématiques intéressants et tenir une conversation assez sophistiquée, dans une langue relativement naturelle, avec un interlocuteur humain, sur un sujet convenablement délimité. Consultés sur un cas médical, ils sont en mesure, le cas échéant, de raisonner aussi bien ou mieux que des experts humains et, qui plus est, de corriger et d'améliorer, en communiquant avec eux, leurs diagnostics et leurs méthodes. De façon générale, des aptitudes essentiellement pratiques, dans lesquelles des choses comme l'« intuition » ou le jugement jouent un rôle décisif, ne semblent plus aussi complètement en dehors de leur portée qu'on ne le supposait fréquemment jusqu'à une date récente.

1. Article initialement publié dans *La Recherche*, N°170, 1985, p.1126-1127. Nous remercions la rédaction de *La Recherche* qui a autorisé la rédaction de Zilsel à reproduire ce texte.

**Classiques**

Les machines sont-elles intelligentes ?

## **Reproduire nos connaissances, mais aussi notre ignorance**

Les théorèmes de limitation, en particulier le théorème de Gödel, ont été abondamment utilisés à une certaine époque pour tenter de démontrer qu'un être humain peut faire des choses qu'aucune machine ne peut ni ne pourrait jamais faire. Mais, en tout cas, il est clair que les incapacités et les impossibilités significatives, lorsqu'on essaie d'apprécier les performances des machines, ne sont pas celles des algorithmes eux-mêmes. Car, comme l'observe Daniel C. Dennett, les limitations (gödeliennes ou autres) imposées à ce qui peut être fait à l'aide d'un algorithme ne sont pas nécessairement des limitations imposées aux capacités d'invention ou d'innovation dont peut faire preuve un mécanisme particulier qui utilise cet algorithme.

Dans le cas du jeu d'échecs, il existe un algorithme qui permettrait en principe à une machine de gagner à coup sûr de façon complètement inintelligente sans recourir à autre chose qu'à la « force brute ». Mais l'utilisation de cet algorithme exigerait une puissance de calcul si colossale qu'il est à peine imaginable qu'une machine quelconque la possède un jour. De ce fait, les programmes d'échecs que l'on expérimente ne reposent pas sur des algorithmes conçus pour mettre systématiquement l'adversaire échec et mat, mais sur des algorithmes permettant de jouer plus ou moins bien (c'est-à-dire, avec des chances plus ou moins grandes de succès) aux échecs. Si on les considère non plus en tant que simples algorithmes, mais du point de vue heuristique, ce qui fait d'eux de bons joueurs est, comme dans le cas d'un être humain, leur aptitude à prendre des risques raisonnables dans une situation d'incertitude relative, à ignorer des possibilités reconnues immédiatement comme dépourvues d'intérêt, à tirer des conclusions qui vont au-delà de ce qui est strictement autorisé par les prémisses, etc. Pour pouvoir simuler de façon acceptable l'intelligence humaine, une machine doit reproduire non seulement notre savoir, mais également en un certain sens, notre ignorance et notre façon de l'utiliser.

Dans l'état actuel des choses, ni les partisans ni les adversaires de l'intelligence artificielle ne semblent disposer d'arguments susceptibles de justifier a priori l'assurance des premiers ou le scepticisme des seconds. Les critiques de philosophes comme Hubert Dreyfus donnent l'impression de relever tantôt de l'objection de principe, appuyée sur des difficultés ou des impossibi-

lités constitutivement insurmontables, tantôt de la simple mise en garde qui, tout en concédant que la tâche n'est pas forcément irréalisable, soutient qu'elle est pour l'instant conçue de façon radicalement inadéquate et que les moyens utilisés actuellement ne peuvent conduire à autre chose qu'un échec retentissant.

En conclusion d'un entretien récent<sup>2</sup>, Dreyfus admettait que l'ordinateur réellement intelligent pourrait exister un jour ; mais selon lui, lorsque les chercheurs se seront attaqués réellement au problème de la simulation de la connaissance usuelle, telle qu'elle est exemplifiée par le comportement le plus ordinaire de l'être humain ou même simplement de l'animal dans son environnement familier, ils s'apercevront que le délai requis n'est pas de l'ordre de trois cents ans, mais plus probablement de trois mille ans. On ne peut évidemment que lui donner raison, lorsqu'il proteste contre les illusions entretenues par des gens qui ont cru être sur le point de comprendre la nature réelle de choses comme la perception, la mémoire la compréhension ou la connaissance, simplement parce qu'ils disposent de programmes relativement simples capables, à leurs yeux, de simuler adéquatement des activités aussi extraordinairement complexes que celles-là. Mais, du point de vue philosophique, on est tenté de dire que le temps ne fait rien à l'affaire et qu'il importe peu que le problème soit résolu dans trente ans ou dans trois mille ans, à partir du moment où aucune impossibilité logique ou conceptuelle n'interdit qu'il le soit un jour.

On a cru pendant longtemps que les activités humaines les plus difficiles à reproduire artificiellement devraient nécessairement être les plus intellectuelles et les plus abstraites. On sait aujourd'hui qu'il n'en est rien. Une connaissance hautement théorique explicable sous la forme d'un ensemble déterminé de règles formelles est beaucoup plus facile à simuler qu'une aptitude éminemment pratique (comme, par exemple, la connaissance du langage) qui, en dépit de ce que croient les linguistes comme Noam Chomsky, est peut-être trop complexe pour pouvoir être représentée sous la forme de la connaissance (implicite) d'une théorie que le locuteur utiliserait sans le savoir. Et les capacités perceptives de l'homme se sont révélées, dans l'ensemble, beaucoup plus réfractaires à la programmation que certaines de ses aptitudes intellectuelles les plus sophistiquées.

Keith Gunderson a proposé de distinguer entre les caractéristiques de l'esprit humain qui se prêtent et celles qui résistent à la

2. Hubert Dreyfus, in *Le Monde de l'Éducation*, N°116, mai 1985.

programmation. Les succès les plus remarquables de l'intelligence artificielle ont été remportés, pour l'essentiel, dans le domaine de la simulation du comportement de l'homo sapiens, les erreurs et les échecs les plus spectaculaires proviennent généralement de la résistance de l'homo sentiens à tout traitement du même genre. Les adversaires de l'intelligence artificielle en ont tiré parfois la conclusion que la conscience et l'affectivité constituaient le dernier bastion que les machines « pensantes » n'investiront à coup sûr jamais : une machine est peut-être capable à la rigueur de pensée, mais certainement pas de sentir, d'éprouver de la douleur ou des émotions, etc.

## **Faire penser les machines et leur donner la faculté d'éprouver des émotions sont deux problèmes de nature différente**

La douleur, par exemple, n'est pas un état définissable de façon purement fonctionnelle et susceptible, comme tous les états de genre d'un grand nombre de réalisations différentes, dont certaines pourraient éventuellement avoir pour support un simple mécanisme programmé de façon adéquate. Elle se définit avant tout par une qualité phénoménologique irréductible qui peut être ou ne pas être présente, mais ne peut justement pas être « simulée » comme peut l'être le comportement de l'homme qui souffre. La conclusion à tirer de cette observation n'est cependant peut-être pas que les particularités de l'homo sentiens constituent un domaine définitivement réservé et protégé, mais plutôt que, comme le suggère Dennett, « la tâche qui consiste à faire en sorte qu'une machine sente est très différente de celle qui consiste à faire en sorte qu'elle pense »<sup>3</sup> et constitue plutôt un problème de « matériel » – hardware – qu'un problème de programmation.

Dennett voit dans l'intelligence artificielle « une recherche de l'espèce la plus abstraite sur la possibilité de l'intelligence ou de la connaissance »<sup>4</sup>. Au lieu d'essayer de répondre directement à des questions de psychologie empirique du type : « Comment notre système nerveux accomplit-il, dans les faits, des tâches de l'espèce X, Y ou Z ? », il est souvent intéressant de poser des questions plus

3. Daniel C. Dennett, *Brainstorms: Philosophical Essays on Mind and Psychological*, Montgomery, Bradford Books Publishers, 1978, p. 190.

4. Hubert Dreyfus, art. cit.

générales du type : « À quelles conditions un système quelconque (possédant les caractéristiques A, B, C...) pourrait-il effectuer X, Y ou Z ? » C'est d'une certaine manière exactement ce que Chomsky a cherché à faire en formulant la question « Quel genre de caractéristiques (abstraites) un système quelconque devrait-il posséder pour être pourvu de ce que nous appelons la connaissance ou la compréhension d'une langue naturelle ? »

Il est clair que ce que l'intelligence artificielle considère comme l'un de ses avantages essentiels, sa tendance à aborder les questions à un niveau d'abstraction délibérément plus élevé que celui des recherches psychologiques usuelles sur l'intelligence, constitue en même temps, aux yeux de ses adversaires, l'un de ses principaux points faibles. Dreyfus souligne de façon très caractéristique que, pour lui : « Notre intelligence ne réside pas dans ce que nous savons, mais dans ce que nous sommes : nous avons un corps et des aptitudes physiques, des émotions (ce qui nous émeut, nous irrite, nous intéresse, etc.). Nous n'avons pas à connaître de choses à leur sujet, parce que nous sommes ce corps, ces émotions. »<sup>5</sup> Mais ce genre de critique risque de déboucher finalement sur une constatation aussi évidente que dépourvue d'intérêt, à savoir qu'il est effectivement vain d'espérer comprendre mieux ce que nous sommes en construisant des êtres artificiels capables de faire certaines choses que nous faisons, si l'on commence par exiger d'eux qu'ils soient exactement ce que nous sommes.

5. Daniel C. Dennett, *Brainstorms*, op. cit.