

Penser l'intelligence artificielle

Éric Delassus

Parce que l'intelligence artificielle, ou ce que l'on désigne par ce terme, est en train d'envahir nos vies selon un processus de développement qui apparaît comme irréversible, il devient urgent de s'interroger sur ce que désigne réellement cette expression, autrement dit de penser l'intelligence artificielle. Cette nécessité est principalement due au fait que celle-ci fait l'objet de nombreux fantasmes, c'est-à-dire de représentations imaginaires fortement chargées d'affects. Si l'on se réfère à la définition que donne du fantasme Élisabeth Roudinesco dans son *Dictionnaire amoureux de la psychanalyse* : « Production de l'imaginaire, scénario, fiction ou rêve diurne, le fantasme représente de façon déformée une réalité », nous avons précisément affaire à ce type de représentation dans l'imaginaire collectif, lorsque nous parlons d'intelligence artificielle. D'autant que ces représentations présentent toute l'ambivalence du fantasme, elles peuvent être à l'origine d'impulsions créatrices féconde, comme elles peuvent nourrir des peurs paralysantes. C'est sur ce point d'ailleurs que se situe le problème posé par cette manière d'appréhender l'intelligence artificielle, car le fantasme relève généralement d'une réaction spontanée plutôt que d'une véritable pensée.

Penser ne signifie pas simplement avoir des idées. Penser signifie d'abord réfléchir, c'est-à-dire procéder à un retour sur soi de la pensée qui permet, pour parler comme Spinoza, de produire des idées de ses idées. Aussi, penser l'intelligence artificielle consiste tout d'abord à s'efforcer d'en produire une idée adéquate, c'est-à-dire une idée qui exprime pleinement ce qu'elle est, à partir d'une interrogation sur les représentations que nous en avons et qui peuvent tout aussi bien être source de fascination et de séduction que de répulsion ou de rejet.

Ce qu'il est d'ailleurs intéressant de souligner tient dans le fait qu'à ce sujet, fascination et répulsion reposent toutes deux sur une même représentation de l'intelligence artificielle liée à la notion de puissance. Si nous sommes fascinés par l'intelligence artificielle, c'est qu'elle nourrit notre désir d'effectuer grâce à la puissance dont elle dispose des opérations que l'intelligence humaine est incapable d'accomplir. Cependant, cette puissance nous inquiète et peut aller jusqu'à susciter en nous un sentiment de répulsion, voire d'effroi. Nous craignons, en effet, qu'un jour, cette puissance de l'intelligence artificielle se retourne contre ses inventeurs et qu'elle rende possible la domination de la machine sur l'homme. De nombreux œuvres de science-fiction ont d'ailleurs exploité, souvent avec talent, cette hypothèse. C'est, par exemple, le cas des films *Matrix* qui nous décrivent un monde dans lequel les machines ayant pris le pouvoir cultivent les humains pour produire leur énergie et les font vivre dans une réalité totalement virtuelle.

L'opinion a donc tendance à osciller entre technophilie et technophobie lorsqu'il s'agit d'intelligen-

ce artificielle et toute la difficulté consiste à trouver la voie qui permettra de sortir de cette alternative.

Cette voie n'est autre que celle de la pensée et du questionnement et ce qu'il faut principalement interroger, c'est le rapport entre l'intelligence artificielle et l'intelligence humaine. Peut-on considérer l'intelligence humaine comme purement naturelle en l'opposant à celle des machines, en d'autres termes la distinction entre une intelligence qui serait naturelle et une autre qui serait artificielle est-elle pertinente ?

Principalement, nous devons interroger les paradigmes en fonction desquels nous concevons ces deux formes d'intelligence. Est-il pertinent de considérer que le cerveau humain fonctionne comme un ordinateur ou de concevoir le fonctionnement des ordinateurs sur le modèle du cerveau humain ? L'intelligence d'une machine est-elle, en effet, comparable à celle d'un être vivant ? Par définition, ce que l'on a coutume de nommer « intelligence » désigne en ensemble varié de compétences et de connaissances issues d'une genèse souvent longue et complexe. Pendant longtemps, les intelligences artificielles étaient particulières, c'est-à-dire dévolue à une tâche bien précise, et incapables d'évoluer. Aujourd'hui, avec le *deep-learning*, les machines dites intelligentes sont en capacité d'apprendre de leurs erreurs et d'évoluer et peut-être serons-nous sous peu en mesure de mettre en œuvre une intelligence artificielle générale, c'est-à-dire de s'adapter à de nouvelles tâches pour lesquelles elles n'ont pas été initialement conçues. Cependant, même si l'intelligence artificielle, comme l'intelligence humaine, peut s'inscrire dans une genèse, dans un processus évolutif, il n'est pas certain que la genèse de l'intelligence artificielle puisse être comparable à celle d'une intelligence vivante. Peut-être parce qu'il y a dans l'intelligence vivante une dimension affective et émotionnelle qui pour le moment, n'est pas présente dans l'intelligence artificielle, mais qui pourrait éventuellement s'y développer, mais certainement sous une autre forme. En effet, pour le moment, les machines ne s'engendrent pas et ne sont pas le fruit d'une gestation, comme c'est le cas pour les êtres humains. La machine, parce qu'elle est artificielle, reste un objet fabriqué. Et même si les machines grâce à la robotique s'avéraient capables de se reproduire, elle ne pourrait le faire que sur le mode de la fabrication et non selon celui de la génération.

Il semble donc nécessaire, pour penser l'intelligence artificielle, de nous affranchir des paradigmes en fonction desquels nous les pensons l'une et l'autre. Autrement dit, ne plus penser que l'intelligence artificielle et l'intelligence humaine pourraient être identiques, la première n'ayant de différence que de degré par rapport à la seconde. Ne pas les penser non plus comme deux forces en compétition, mais plutôt tenter de penser leur articulation dans le cadre d'une intelligence artificielle adossée à l'intelligence humaine. Aussi, est-il nécessaire ce penser l'intelligence humaine dans ce qu'elle a de spécifique, afin qu'elle puisse continuer à effectuer les opérations que l'intelligence artificielle ne pourra jamais faire et, réciproquement, de penser la spécificité de l'intelligence artifi-

cielle afin de lui faire effectuer ce que l'intelligence humaine n'est pas en mesure de faire.

Ainsi, pourrions-nous conclure avec l'intelligence artificielle, ce que nous serions tentés d'appeler, non sans une certaine ironie, « une nouvelle alliance » avec l'intelligence artificielle. Certes pas pour faire de celle-ci un nouveau Dieu, loin de là, mais pour tenter de penser la manière dont ces deux intelligences distinctes peuvent s'articuler l'une à l'autre, mais toujours en vue de l'intérêt humain. Ce qu'il faut tenter de penser ici, c'est les modalités selon lesquelles l'intelligence parviendra à poursuivre son processus de développement en y intégrant l'existence de l'intelligence artificielle à laquelle elle va devoir s'adapter. Toute innovation technologique est productrice d'un monde, c'est-à-dire d'un horizon de sens dans lequel doit s'efforcer de s'orienter. Il nous faut donc réfléchir au type de monde que va produire le développement de l'intelligence artificielle et apprendre à nous adapter à ces transformations pour mieux les apprivoiser.

L'intelligence humaine est une faculté d'adaptation, elle est le produit d'une histoire et du développement d'une multiplicité de cultures qui durant des millénaires se sont rencontrées plus ou moins pacifiquement. Cette intelligence s'est construite, elle n'a pas été construite et elle va devoir continuer de se construire en s'adaptant à la présence de l'intelligence artificielle, en s'alliant avec elle.

Peut-être un processus inverse est-il en train de s'initier pour l'intelligence artificielle. Pour le moment, celle-ci ne semble pas avoir atteint un degré suffisant pour parvenir à la conscience d'elle-même, elle consiste principalement à mimer les manifestations extérieures de l'intelligence humaine dans l'effectuation d'opérations bien spécifiques. Il est néanmoins permis de se demander si un jour, peut-être proche, un saut qualitatif ne sera pas accompli. Peut-être un jour serons-nous confrontés à des machines conscientes, voire à des machines susceptibles de ressentir des émotions ?

Nous sommes là confrontés à un problème véritablement métaphysique. Si l'on considère l'esprit comme une réalité en soi, substantiellement distincte du corps auquel il est uni, on est en droit de considérer que jamais une machine ne pourra atteindre un tel niveau de développement. En revanche, si l'on pense que l'esprit émerge d'une certaine organisation complexe de la matière, rien n'empêche d'envisager son apparition dans des produits de l'intelligence humaine dotés eux-mêmes d'une intelligence artificielle. Cela dit, à supposer que puisse voir le jour des machines dotées d'une intelligence consciente et peut-être également capable d'affects, cette conscience et ces affects seront ceux d'une machine et ne seront donc pas en tous points comparables à ceux d'un être humain. D'où la nécessité de penser cette intelligence pour ne pas se laisser surprendre par les effets qu'elle pourrait produire.

Dans la *Lettre au marquis de Newcastle du 23 novembre 1646*, Descartes affirme que ce qui distingue un homme d'un automate, c'est sa capacité à dialoguer. Ce qui fait que nous reconnaissons que nous avons affaire à un être humain et non à une machine mimant parfaitement son comportement

consiste dans la capacité de l'homme à parler et surtout à s'exprimer « à propos », c'est-à-dire en produisant un discours en adéquation avec la situation dans laquelle il se trouve et correspondant au discours de son interlocuteur :

Enfin il n'y a aucune de nos actions extérieures, qui puisse assurer ceux qui les examinent, que notre corps n'est pas seulement une machine qui se remue de soi-même, mais qu'il y a aussi en lui une âme qui a des pensées, excepté les paroles, ou autres signes faits à propos des sujets qui se présentent, sans se rapporter à aucune passion.

Cet argument n'est pas sans rappeler le fameux test de Turing, ce mathématicien anglais qui, dans un article de 1950, qui considérait qu'on peut affirmer qu'une machine pense lorsqu'elle est en mesure de dialoguer avec un être humain sans que personne soit en mesure de déceler qu'il s'agit d'un programme informatique et non d'un réel sujet humain. Aujourd'hui, précisément, la thèse cartésienne est remise en question, car de nombreuses machines imitent à s'y méprendre la capacité d'un être humain à entrer en conversation avec l'un de ses semblables. Cela ne concerne, il est vrai que des sujets bien précis et bien délimités, mais rien ne dit que le champ d'application de l'intelligence artificielle ne dépassera pas ses limites actuelles.

On peut désormais se demander si à force de permettre à des machines de mimer toutes les manifestations extérieures d'une intelligence consciente, nous ne sommes pas également en train de transformer leur mode de fonctionnement interne et d'en faire de réelles consciences artificielles.

Probablement, nous faudra-t-il nous adapter à ces types d'intelligence artificielle dans les décennies qui viennent. C'est pourquoi il nous faut penser l'intelligence artificielle, si nous ne voulons pas être surpris par les conséquences de son développement.