



Jean-Pierre Hauet
Rédacteur en chef de la REE

L'intelligence artificielle : prothèse ou orthèse ?

L'intelligence artificielle (IA) fait partie des technologies pseudo-cycliques, de celles sur lesquelles on fonde de grands espoirs jusqu'à ce que le soufflé retombe et avant qu'une nouvelle génération d'ingénieurs et de responsables politiques reprenne le flambeau, quelque 25 ou 30 ans plus tard. L'IA est née dans les années 50 des travaux d'Alan Turing. Elle a beaucoup mobilisé les scientifiques à l'époque puis a connu un nouveau pic d'engouement dans les années 1985 avec les systèmes experts, la logique floue et les premiers réseaux neuronaux, mais les résultats, une nouvelle fois, n'étaient pas au rendez-vous.

Cependant les technologies pseudo-cycliques parfois décollent et ce décollage peut être très rapide : on l'a vu avec les écrans plats et les cellules photovoltaïques. Aujourd'hui, l'IA connaît un redémarrage très puissant avec des résultats spectaculaires, en particulier pour tout ce qui a trait à la reconnaissance de formes ou d'images, avec des applications médicales assez impressionnantes, notamment dans le domaine du diagnostic avec, par exemple, la possibilité de détecter par analyse d'image des mélanomes malins, avec des taux d'erreur plus faibles que ceux des meilleurs spécialistes humains.

L'IA est à présent couramment utilisée dans des applications grand public telles que la reconnaissance de la parole et les assistants vocaux ou IPA (Intelligent Personal Assistant) tels que Google Assistant, Siri, Alexa, Bixby, Nina, etc. On trouve aussi l'IA dans le domaine du traitement statistique, avec l'analyse des données massives et la détection des signaux faibles, dans les jeux, dans la robotique et la cobotique, etc.

Ces applications de plus en plus sophistiquées laissent à penser que l'on est à présent capable de reproduire les mécanismes du cerveau humain et donc

de l'intelligence humaine et un grand nombre y voit à la fois un progrès et une menace, car si une machine est capable de se substituer à l'esprit humain, qui peut dire où elle pourra s'arrêter et qui pourra l'arrêter ? Les déclarations de personnalités éminentes telles que Stephen Hawking et Elon Musk ont donné corps à ces craintes, en prédisant l'apocalypse, suite à une prise de contrôle du monde entier par des dictateurs numériques avec lesquels les humains ne pourraient pas rivaliser.

La question est donc posée de savoir ce qu'est l'intelligence artificielle d'aujourd'hui, s'il y a une IA « forte » se démarquant de l'IA « faible » des années 1980 ou, pour être dans l'air du temps, une IA 2.0 dont il faudrait se réjouir mais dont il serait sage de se méfier.

Le problème de la définition de l'IA n'est pas nouveau et il y a toujours eu quelque flou autour de ce concept. L'IA est en effet usuellement définie comme l'ensemble des théories et techniques mises en œuvre pour réaliser des machines capables de simuler l'intelligence. « Simuler l'intelligence », cela pose plusieurs questions :

- s'agit-il de l'intelligence humaine ou d'une autre forme d'intelligence, l'intelligence des animaux, des insectes ou une intelligence simplement rationnelle ?
- que veut-dire « simuler » ? S'agit-il simplement de développer des comportements, des formes de raisonnement ayant l'apparence de l'intelligence humaine ? Ou bien, avec l'IA que l'on dit nouvelle, aurait-on la prétention d'avoir reproduit les mécanismes de fonctionnement interne du cerveau, c'est-à-dire de pouvoir « émuler » et non plus « simuler » en portant ces mécanismes sur une nouvelle machine ? Ces machines seraient-elles alors dotées d'un esprit semblable à celui de l'être humain ? Mais n'est-ce qu'une apparence ou bien la réalité intrinsèque ? Et toute l'activité du cer-

veau humain et pas seulement certaines fonctions spécifiques, pourrait-elle ainsi être transplantée sur une autre machine ?

En deux mots, l'IA est-elle une Intelligence Artificielle ou une Intelligence Augmentée ? Un substitut à l'intelligence naturelle ou simplement un outil qui en améliore l'efficacité ? Une prothèse ou une orthèse ?

Pour l'informatique traditionnelle, la réponse était claire. Lorsqu'un humain calcule $15+7$, il ne décompose pas chacun des nombres en base 2 avant d'appliquer les règles du calcul binaire. Mais l'IA nouvelle est différente : elle procède par apprentissage et la modélisation à laquelle celle-ci fait appel, repose sur des réseaux de neurones structurés en plusieurs couches dont les « poids synaptiques » c'est-à-dire les coefficients d'influence d'un neurone sur les autres, peuvent être modulés par l'apprentissage. On parle alors de *deep learning* ou d'apprentissage profond, car il se produit alors un phénomène de boîte noire : au fur et à mesure que des tests sont faits sur des échantillons toujours plus nombreux, le modèle s'affine au prix de calculs très complexes, le taux d'erreur diminue sans que l'on sache exactement ce que représente le modèle interne ainsi construit et quel raisonnement est suivi pour parvenir à une conclusion donnée. On atteint des niveaux de perfectionnement assez bluffants comme en atteste la démonstration de prise de rendez-vous chez le coiffeur faite par le système Google Duplex en mai 2018, dans laquelle rien ne permettait de distinguer la voix de la machine de celle d'un humain, pas même les « hmm, hmm » renvoyés en cas d'hésitation.

Mais cette démonstration, comme bien d'autres, vient alimenter le débat : les machines ont-elles réellement compris quelque chose au dialogue auquel elles participent ou bien ne sont-elles que des robots qui exploitent mécaniquement une immense base de connaissances, sans comprendre le sens des mots et des phrases échangées, en utilisant des échappatoires lorsque la question sort des limites de leur base de données ?

Beaucoup estiment que l'IA bute sur les capacités de calcul des machines et font valoir que le cerveau est une belle machine qui contient des milliards de cellules et peut-être un million de milliards de synapses, c'est-à-dire de connexions entre ces cellules. C'est un ordinateur biologique et son rendement thermodynamique est bien meilleur que celui des meilleurs calculateurs actuels. Avec une puissance de 28 W, correspondant approximativement à la puissance dissipée en moyenne

par un cerveau humain, on peut faire fonctionner aujourd'hui un calculateur qui aura une puissance à peu près équivalente au 1/10 000 d'un cerveau humain. Ceci veut dire que l'ordinateur le plus puissant au Monde, le Summit, construit aux Etats-Unis et qui absorbe 15 MW, a une puissance de calcul équivalente à celle d'un hypothétique cerveau humain qui aurait des dimensions quatre fois supérieures. La différence n'est pas considérable, on mesure en conséquence la difficulté à faire supporter par un calculateur de taille modeste et a fortiori par un simple composant, des fonctions aussi complexes que celles attachées à la conduite des véhicules.

Cependant cet argument sur la puissance du cerveau est insuffisant pour disqualifier l'IA et écarter l'hypothèse d'un rattrapage du cerveau par les machines. L'électronique d'aujourd'hui a un rendement thermodynamique, calculé par référence à la limite de Landauer¹, extrêmement mauvais et il est à parier que nos descendants lorsqu'ils examineront les machines que nous utilisons aujourd'hui, seront remplis d'étonnement et de condescendance : nos calculateurs actuels sont beaucoup moins performants que la lampe à huile comparée à l'éclairage par LED. Mais des voies de progrès existent qui se traduiront par des ruptures technologiques majeures : recours à la nanoélectronique, à l'électronique de spin et aux nanooscillateurs à transfert de spin, peut-être aux calculateurs quantiques... Ces technologies arriveront à maturité et donc les machines continueront à se miniaturiser en offrant toujours plus de puissance de calcul et de capacité de mémoire. Des programmes extraordinairement complexes et d'immenses bases de données pourront venir se loger dans des machines d'une taille cette fois plus petite que celle de notre cerveau, donnant l'impression que la compétition est perdue.

Mais la différence entre un humain et une machine n'est pas une question de puissance de calcul. Il y a d'autres éléments plus fondamentaux et, à vrai dire, plus mystérieux.

Prenons le cas d'un robot tueur qui va surgir au milieu d'un forum organisé, par exemple !, par la SEE. Il y reconnaîtra sans aucune hésitation les visages qu'il recherche, s'il a été programmé pour cela. Le GIGN sera appelé en urgence, mais croit-on que le chef de peloton pourra entamer avec lui un dialogue en rappelant l'existence

¹ La limite de Landauer est l'énergie minimale nécessaire pour faire basculer un bit d'information. Elle s'écrit $E = kT \ln 2$, formule dans laquelle k est la constante de Boltzmann ($1,38 \times 10^{-23} \text{ J.K}^{-1}$), T la température en °K et $\ln 2$ le logarithme népérien de 2 (0,69315).

de sa vieille mère qu'il a laissée au pays ou en tentant de le faire parler de son enfance malheureuse... Pense-t-on qu'une machine, qu'un robot puisse connaître des émotions, qu'il ait une sensibilité, qu'il puisse connaître des humeurs et tenir compte d'un contexte évolutif ?

Car au fond, les impressionnantes technologies d'apprentissage dont on s'étonne ne restent que des algorithmes, basés sur de gigantesques matrices de convolution. Où sont l'histoire, l'éducation, la sensibilité, sans aller jusqu'à parler de la conscience... ?

Mais même face à ces algorithmes, dont on pourrait soutenir qu'ils représentent une partie de la capacité de raisonnement humain à défaut d'en représenter l'intégralité, on peut s'interroger. La reconnaissance de formes ou d'objets, demande l'analyse des résultats de dizaines de milliers de tests avant de pouvoir ramener le taux d'erreur à un niveau acceptable. C'est d'ailleurs l'un des problèmes qui se trouvera posé par la certification des véhicules autonomes : quand pourra-t-on arrêter les tests et considérer que l'on a atteint le niveau de sécurité requis, avec l'assurance de n'avoir rien oublié ? Un enfant est plus expéditif : une fois qu'on lui a présenté une pomme, peut-être deux, il saura reconnaître toutes les pommes et il n'est point besoin pour cela de lui en présenter toutes les variétés ! Est-ce simplement une question de puissance de calcul ? A coup sûr non. Certains disent aujourd'hui que les calculs de coefficients synaptiques ne représentent pas les mécanismes du cerveau et qu'il faut leur substituer une approche impulsionnelle (*spike neural networks*) dans laquelle un neurone ne s'active pas systématiquement mais seulement lorsque son potentiel excède un certain seuil, auquel cas il envoie une impulsion vers les neurones avec lesquels il correspond qui se trouvent ainsi à leur tour activés. Comment être sûr que l'on s'est véritablement rapproché des mécanismes du cerveau ?

Mais il y a plus grave : quand je décide de lever mon bras, je le lève alors que personne ne m'y oblige, sans qu'il n'y ait aucun stimulus externe. Je lève mon bras parce que j'ai décidé de le lever. Je lève mon bras mais je peux aussi bien prendre mon porte-plume et faire acte de création littéraire ou artistique, un beau matin

sans que rien ne m'y oblige. Ce pouvoir d'initiative et de création a quelque chose de mystérieux car on n'a jamais, que je sache, vu une machine décider de faire autre chose que ce pour quoi elle a été programmée. Elle peut le faire bien ou le faire mal et dans ce cas être à l'origine de dégâts considérables. Un assistant « intelligent » pourra répondre de façon plus ou moins pertinente aux questions d'une personne âgée. Mais peut-on imaginer de le voir tout d'un coup rendre son tablier pour aller se consacrer à d'autres tâches et développer par exemple une activité créatrice ?

Le cerveau humain et sans doute aussi l'intelligence animale conservent leur mystère qui ne relève peut-être pas seulement de la technique. Attendons-nous à voir des machines de plus en plus puissantes et sophistiquées qui allégeront les humains de plus en plus de tâches, accroîtront leurs capacités, pourront le cas échéant avoir un pouvoir destructeur mais ne remplaceront pas l'homme dans sa capacité créatrice.

Il y a malheureusement un autre péril dont il faut tenir compte : celui de la perte des connaissances et du savoir-faire, et je devrais dire de « l'abêtissement » d'une fraction croissante des populations qui d'ores et déjà ne savent plus faire une division à la main, ne savent plus lire une carte et se précipitent vers leur smartphone au moindre problème. En fait cette dégénérescence du savoir n'est pas universelle : il subsiste fort heureusement des personnes qui maîtrisent la conception et la fabrication des machines. Mais ces spécialistes sont une minorité et on peut donc soutenir que les technologies numériques malgré tous les services qu'elles rendent, sont un facteur d'inégalité, créant un fossé énorme entre ceux qui les maîtrisent et ceux qui ne savent rien faire de

mieux que de les utiliser sans avoir la moindre idée sur la façon dont elles fonctionnent². Notre revue REE, à son niveau, s'efforce de combler ce vide et d'offrir à ses lecteurs la possibilité de rester dans le peloton de tête sans avoir recours à l'intelligence artificielle 2.0 ! ■

JEAN-PIERRE HAUET est rédacteur en chef de la REE. Ancien élève de l'École polytechnique et ingénieur au corps des mines, il a notamment dirigé le laboratoire central du Groupe Alcatel-Alsthom avant de devenir Chief Technology Officer du groupe Alstom. Il est l'auteur de nombreux ouvrages et publications. Ses travaux actuels sont axés sur le domaine de l'énergie et sur celui de la cybersécurité.

² Sur cette question cruciale de l'IA, potentiel générateur d'inégalités, le lecteur est invité à se reporter à l'ouvrage du neurobiologiste Laurent Alexandre intitulé « La guerre des intelligences », paru en octobre 2017 aux Editions JC Lattès.