

chercheurs en IA, pensent plutôt qu'il faudra attendre 2040, voire 2100.

Il se pourrait que l'on se trouve dans une situation analogue à celle du jeu d'échecs : on y arrivera, mais pas dans les délais annoncés, car le problème est réellement difficile, ce que l'on découvre année après année. Le plus raisonnable serait sans doute de ne pas tenter aujourd'hui de pronostics !

Liée à cette éventuelle superintelligence artificielle, l'idée d'une « singularité » à venir est évoquée sans qu'on puisse savoir vraiment de quoi il s'agit. Une date où se situerait une sorte d'asymptote verticale dans la courbe mesurant le niveau des intelligences artificielles créées sur Terre ? Un point où les machines écrasant les humains prendraient le pouvoir ? Un instant au-delà duquel il est impossible de comprendre ce qui va se passer à cause des bouleversements entraînés par l'existence d'intelligences artificielles nous dépassant ?

Tout cela n'est pas très sérieux, même en prenant en compte la célèbre « loi de Moore » (vérifiée en gros depuis cinquante ans) qui indique un doublement de capacité des dispositifs informatiques tous les dix-huit mois ou deux ans. D'une part, une fonction exponentielle comme celle envisagée par la loi de Moore n'a pas d'asymptote verticale. D'autre part, dans notre monde physique, toute croissance exponentielle rencontre rapidement des obstacles qui l'arrêtent et personne ne défend sérieusement que la loi de Moore (qui s'essouffle déjà) continuera à être valide indéfiniment. Dernier point : l'augmentation des capacités de calcul et de mémoire ne constitue pas la garantie qu'on saura les utiliser pour créer cette fameuse intelligence générale qui nous résiste.

Les annonces des Gafam (Google, Apple, Facebook, Amazon, Microsoft) sur ces sujets doivent être examinées avec méfiance. Certains spécialistes comme Jean-Gabriel Ganascia, président du comité d'éthique du CNRS et lui-même chercheur en intelligence artificielle, sont plus rigoureux et lucides que les agents publicitaires des firmes de technologie. Jean-Gabriel Ganascia considère que ceux-ci cherchent à promouvoir un projet du dépassement de la politique et des États-nations par des entreprises à l'avidité sans borne. L'idée de la singularité, utilisée pour amuser et effrayer, aurait en réalité surtout pour but d'intimider et de masquer leurs ambitions dominatrices.

Les difficultés à anticiper avec précision n'interdisent pas de s'adonner au jeu de l'anticipation en s'interrogeant sur l'apparition d'une superintelligence artificielle, quitte pour cela à imaginer plusieurs scénarios, sans donner d'échéances et en envisageant qu'éventuellement aucun ne se réalise. Jouons.

Nick Bostrom définit dans son livre une superintelligence comme tout intellect qui

excède radicalement les capacités cognitives humaines et ce dans tous les domaines. Deux hypothèses assez différentes sont envisagées.

(a) Les machines à base de silicium, y compris les ordinateurs quantiques s'ils se révèlent utiles, progressent au point que les dispositifs construits nous égalent et nous dépassent dans toutes les tâches intellectuelles possibles, cela, comme pour AlphaZero, sans que nous soyons associés directement à leur fonctionnement et sans que des éléments biologiques aient été fusionnés ou connectés à ces créatures intelligentes artificielles.

(b) Seule une association entre des éléments (humains ou non) provenant de la biologie et des dispositifs informatiques, donc un mixte de technologies inventées par l'homme et de parties utilisant ce que la vie sur Terre a patiemment élaboré, réussit ce dépassement.

ARTILECTS OU CYBORGS ?

Les êtres intelligents provenant de l'hypothèse (a) sont dénommés *artilects* (pour *artificial intellect*) par Hugo de Garis, un chercheur qui a développé un scénario pour le futur sur lequel nous allons revenir.

La science-fiction a envisagé l'hypothèse (b) depuis longtemps et dénommé cyborgs (pour *cybernetic organisms*) ces êtres hybrides superintelligents grâce au mélange et à l'imbrication de biologie et d'informatique. Imaginons par exemple que, pour suppléer à notre mémoire insuffisante, on nous greffe des puces contenant des téraoctets de données qui nous permettraient directement par la pensée de savoir tout ce qu'il y a dans l'encyclopédie Wikipédia, et que grâce à un module de calcul mathématique lui aussi accessible instantanément par la pensée, nous sachions répondre à n'importe quelle question du type « Combien fait 7^{20} ? » ou « Quelle est la dérivée cinquième de $\sin(\cos(x))$? » Notre intelligence en serait accrue !

Une troisième hypothèse est parfois évoquée, celle d'un cerveau planétaire naissant des réseaux de machines qui s'échangent des informations et constituent une sorte d'être décentralisé et pensant ; nous y reviendrons.

LE PESSIMISME DE HUGO DE GARIS

Pour Hugo de Garis, « la société se divisera en trois groupes philosophiques majeurs, agressivement opposés. Le premier groupe sera constitué par les *cosmistes* (terme dérivé du mot *cosmos*), favorables à la construction d'artilects. Le deuxième groupe sera constitué des *terrans* (dérivé du mot *Terra*, la Terre) opposés à la construction d'artilects. Le troisième groupe sera celui des *cyborgistes*, qui souhaitent devenir eux-mêmes des artilects en ajoutant des composants à leur propre cerveau humain. »

>