

RODOLPHE GELIN

Dernières nouvelles
de l'intelligence
artificielle



Flammarion

RODOLPHE GELIN

Dernières nouvelles de l'intelligence artificielle

Voitures autonomes, reconnaissance faciale, traduction automatique... Comment fonctionnent les technologies d'intelligence artificielle (IA) qui envahissent notre quotidien ? Par quel miracle est-on passé d'un ordinateur poussif à des algorithmes capables d'apprendre de leurs erreurs et de vaincre les plus grands maîtres d'échecs et de go ?

Dans cette brillante synthèse, l'auteur ouvre la boîte noire de l'IA en révélant la genèse de cette « intelligence », si proche et si éloignée de la nôtre à la fois. Pour en finir avec les idées reçues, il propose un regard neuf et concret sur cet enjeu de société majeur, tout en esquisant le futur de l'IA à travers les pistes de recherche les plus actuelles.

RODOLPHE GELIN est expert en *deep learning* pour le véhicule autonome et le véhicule connecté chez Renault, après une longue carrière d'ingénieur qui l'a conduit à développer Romeo, un robot humanoïde destiné à aider les personnes âgées. Il a notamment publié *L'IA est-elle contre nous ?* aux Éditions de la Documentation française et *L'IA et nous* au Pommier (avec O. Guilhem).

Dernières nouvelles de l'intelligence artificielle

Rodolphe Gelin

Dernières nouvelles de l'intelligence artificielle

Infographies d'Océane Juvin

Flammarion

Schémas : Laurent Blondel/Corédoc
© Flammarion, 2022
ISBN : 978-2-0802-5161-9

INTRODUCTION

Une victoire historique

En mars 2016, une intelligence artificielle, AlphaGo, battait le meilleur joueur du monde de go, Lee Sedol, et empochait la récompense promise, un million de dollars. L'histoire ne dit pas comment les ingénieurs de la société DeepMind ont dépensé cet argent. Probablement ont-ils commencé par payer la note salée d'électricité que la machine avait accumulée. Il est vrai qu'avec ses 1 202 processeurs principaux et ses 176 processeurs graphiques, AlphaGo est un ogre énergétique qu'il avait fallu alimenter les six jours durant, le temps de faire cinq parties, sans parler des mois nécessaires à son entraînement.

À l'époque, l'issue du match a probablement froissé quelques egos dans la communauté des joueurs de go, tandis que les amoureux de

l'intelligence artificielle, eux, se frottaient les mains. Ils tenaient enfin leur revanche sur les oiseaux de mauvais augure qui avaient vu, quelques années plus tôt, la victoire de la machine Deep Blue contre Garry Kasparov comme le chant du cygne de l'intelligence artificielle (IA) : une IA avait peut-être battu le champion du monde d'échecs, mais jamais elle ne détrônerait les humains au go !

Il faut dire que si, en Occident, les échecs passent souvent pour être le roi des jeux, le go est considéré en Asie comme bien plus sophistiqué. Malgré la simplicité de ses règles (toutes les pièces ou « pierres » sont équivalentes), sa combinatoire est beaucoup plus importante (il y a 361 « cases » sur un plateau de go, contre 64 aux échecs). Le jeu requiert à la fois une vision globale et une stratégie à plus long terme. Ces contraintes pesaient sur les logiciels d'alors et empêchaient en 1997, quand Kasparov a perdu, les méthodes de programmation employées pour les jeux d'échecs de briller au go.

Quelle leçon en tirer ? Mieux vaut éviter de déclarer qu'un objectif est définitivement impossible pour l'IA. Au mieux peut-on juger que, dans l'état actuel des connaissances, tel objectif sera difficile à atteindre. J'essaierai de m'en tenir

à cette position dans cet ouvrage. Il est souvent tentant de confondre « nous ne savons pas le faire » avec « c'est impossible ». Mais l'histoire des sciences est là pour nous rappeler qu'il faut résister à cette tentation.

Mais que s'est-il passé en vingt ans ? Comment les informaticiens ont-ils fini par descendre l'humanité de son piédestal au go ? Pour répondre, j'emploierai la comparaison du pêcheur : plutôt que donner du poisson à une personne, ils lui ont appris à pêcher. Voilà la révolution copernicienne qu'ils ont lancée et qui ne bouleverse pas seulement le domaine des jeux, mais bien l'informatique dans son ensemble : ils ont appris aux machines à apprendre.

Au milieu des années 1990, une IA se formant aux échecs apprenait les règles, les meilleurs coups des grands maîtres et, lors d'un match, grâce à la puissance de l'informatique, calculait tous les coups possibles (les siens comme ceux de son adversaire) pour en déduire le meilleur déplacement de ses pièces. L'efficacité de la méthode butait sur un obstacle : l'IA n'apprenait rien de ses défaites. Elle jouait identiquement sa deux millième partie comme sa première. Pour l'améliorer, les programmeurs se penchaient sur les raisons de ses défaites et devaient attendre la mise

à jour suivante pour modifier explicitement le programme.

Apprendre de ses expériences

AlphaGo (et plus encore son successeur AlphaGo Zero – comme zéro connaissance au début de l'apprentissage), lui, ne connaît au départ que les règles du jeu et ne s'emploie qu'à les appliquer. Tel le forgeron qui le devient en forgeant, le logiciel apprend seul, au fil des parties, les coups qui le font gagner ou perdre. Ainsi, AlphaGo monte en gamme au fil de sa « vie ». Si vous l'aviez affronté à ses débuts, sans doute auriez-vous pu n'en faire qu'une bouchée. Mais, après quelques-unes de ses parties d'entraînement, plus ou moins nombreuses selon votre « dan », vous n'auriez bien vite plus fait le poids. Ces techniques d'apprentissage automatique existaient avant 1997, mais n'étaient pas aussi abouties qu'aujourd'hui. L'un de leurs facteurs limitants résidait dans la puissance des ordinateurs de l'époque, largement insuffisante pour les calculs qu'exige ce type de problème.

Cette nouvelle approche a réservé des surprises lors de l'affrontement entre AlphaGo et

Lee Sedol. L'IA a joué des coups qui ont désarçonné les plus grands spécialistes du jeu ! Ces coups ont été qualifiés *a posteriori* d'« inhumains », tandis que d'autres, au contraire, auraient valu des remontrances de son professeur à tout élève qui les aurait tentés. Pourtant, ils ont fonctionné pendant le match. Certains sont même enseignés depuis dans les écoles. AlphaGo a-t-il fait preuve de créativité, de la même façon que Dick Fosbury a révolutionné le saut en hauteur en 1968, lorsqu'il a franchi l'obstacle en tournant le dos à la barre plutôt qu'en la faisant passer sous son ventre ? Ou a-t-il plutôt recueilli les fruits d'une exploration méthodique dont la peur de perdre, qui lui est inconnue, nous empêche parfois de jouir ?

Lors de son apprentissage, AlphaGo Zero a joué des millions de parties contre lui-même et a nécessairement beaucoup, beaucoup perdu. Mais, en perdant, il apprenait beaucoup, car il mémorisait des erreurs à ne plus commettre. Dans un autre jeu pratiqué par l'IA, le poker, un comportement du programme a, de la même façon, étonné certains observateurs : l'ordinateur semblait capable de bluffer. Il était susceptible de miser gros, alors que son jeu le destinait à une défaite quasi certaine. Une lecture naïve de ce comportement laisserait penser que les IA sont

aussi espiègles que les humains et connaissent les ressorts de la psychologie humaine. Or, là encore, cette attitude est issue de l'apprentissage de la machine. Lors de ses multiples parties, l'IA a dû s'apercevoir que le couple formé de la qualité des cartes et de la mise confère toute la valeur d'une main. Et qu'elle doit s'appuyer sur cette valeur pour prendre ses décisions.

L'exemple du poker illustre cette tentation bien naturelle qui consiste à projeter notre façon de réfléchir sur celle de l'IA. C'est la voie la plus directe vers l'anthropomorphisation de l'IA, à qui on prête toutes sortes de qualités et de défauts humains dont celui, bien connu, de vouloir conquérir le monde et d'écraser ceux qui lui barreront le chemin ! Dans les prochaines pages, nous nous emploierons ensemble à mieux connaître cette IA, qui n'est pas une personne mais une technologie dont les domaines d'application sont innombrables, depuis les moteurs de recherche sur Internet jusqu'à la voiture autonome, en passant par le ciblage des publicités ou la synthèse de nouvelles protéines. Nous pourrions donc nous permettre, sans IRM fonctionnelle, de regarder comment elle résout ces problèmes complexes. Nous comprendrions ainsi en quoi cette intelligence est à la fois proche (les chercheurs

ne cachent pas l'inspiration biologique de leurs travaux) et pourtant si éloignée de la nôtre. Sur ces bases solides, il nous sera possible d'imaginer l'éventuel futur de l'IA et de ses usages. Et d'anticiper, en bannissant évidemment le mot « impossible », les domaines où il semble difficile qu'elle surpasse les humains.

I

LA PRÉHISTOIRE DE L'IA

Comme toutes les technologies, l'IA a une histoire. Si des esprits brillants, assez bien identifiés (Alan Turing, John McCarthy, Marvin Lee Minsky, etc.) l'ont formalisée et ont commencé à lui donner vie au milieu du XX^e siècle, elle ne sort pas de nulle part. C'est l'aboutissement logique des efforts que nous faisons, depuis toujours, pour nous simplifier la tâche. Entre le levier et les réacteurs nucléaires, l'homme a franchi un nombre considérable d'étapes lui permettant de démultiplier son interaction physique sur son environnement. L'histoire de la démultiplication de ses capacités intellectuelles est beaucoup plus récente et, jusqu'à il y a peu, beaucoup moins technologique. Un des premiers jalons de cette quête a probablement été l'écriture, qui a permis de fixer les idées, de les partager, de s'y appuyer pour en construire de nouvelles. L'écriture matérialise la mémoire

humaine, qui a été l'un des rouages majeurs du développement de l'intelligence.

Corollaire de l'écriture, les chiffres ont également grandement compté pour améliorer l'efficacité de l'esprit humain et notre compréhension du monde. En effet, tout autant que les mots qui nous servent à décrire les choses, les nombres sont eux aussi des supports de cet imaginaire, à la nuance près que les règles de la grammaire et de la syntaxe sont remplacées par des équations issues des mathématiques et de la physique. De la même façon que les astronomes ont cherché les équations expliquant le mouvement des planètes, d'autres scientifiques ont cherché celles qui expliquaient ceux de nos raisonnements. Sans le savoir, ces pionniers ont jeté les bases de la future intelligence artificielle.

Un monde kafkaïen

Ces explorateurs de la pensée, qui sont-ils ? Les philosophes de l'Antiquité ont été les premiers à formaliser la nature d'un raisonnement logique. Au IV^e siècle avant notre ère, Aristote en a identifié l'un des piliers : le syllogisme. L'exemple classique est la démonstration du fait