

Une approche accessible pour tous de l'intelligence artificielle

A la recherche d'une intelligence artificielle à la fois maîtrisable, soutenant le travail humain par une puissance et une rapidité bien supérieures. Est-ce possible ? Retour sur des faits marquants et des repères à saisir pour approcher un domaine hautement complexe mais répandu dans nos vies quotidiennes.

L'intelligence artificielle (IA) est une histoire ancienne. Au début des années 1950, il existait différents noms pour le domaine des « machines à penser », comme la cybernétique, la théorie des automates et le traitement complexe de l'information. C'est dans ces années qu'Alan Turing a imaginé le test fondé sur la faculté d'une machine à imiter la conversation humaine. Il consiste à mettre un humain en confrontation verbale à l'aveugle avec un ordinateur ou un autre humain. Si la personne qui engage les conversations n'est pas capable de dire si son interlocuteur est un ordinateur ou un humain, on peut considérer que le logiciel de l'ordinateur a passé avec succès le test.

En 1956, c'est au collège américain de Dartmouth qu'a eu lieu l'événement qui a donné son nom à l'intelligence artificielle, lors d'un colloque resté

Bertrand Braunschweig est directeur du centre de recherche Inria Saclay - Île-de-France, docteur de l'université Paris-Dauphine et HdR de l'Université Paris VI. Inria (Institut national de recherche en informatique et en automatique) est un établissement public à caractère scientifique et technologique dédié aux sciences du numérique.

Cet article est la retranscription de l'intervention lors du Comité national CFDT Cadres de novembre 2018.

célèbre où la première définition de l'IA a été donnée : le projet est d'avancer sur la conjecture que tous les aspects de l'apprentissage ou d'autres caractéristiques de l'intelligence peuvent être décrits si précisément que l'on puisse fabriquer une machine capable de les simuler.

De nombreuses autres définitions ont été proposées ; Marvin Lee Minsky (1927-2016), fondateur avec McCarthy du groupe d'intelligence artificielle du Massachusetts Institute of Technology (MIT), est l'un des précurseurs de la discipline. Il définit l'IA comme « *la construction de programmes informatiques qui s'adonnent à des tâches qui sont, pour l'instant, accomplies de façon plus satisfaisante par des êtres humains car elles demandent des processus mentaux de haut niveau tels que : l'apprentissage perceptuel, l'organisation de la mémoire et le raisonnement critique* ». En d'autres termes, une intelligence artificielle est avant tout un programme informatique visant à effectuer, au moins aussi bien que des humains, des tâches nécessitant un certain niveau d'intelligence.

L'IA s'est développée depuis soixante ans avec des phases de croissance et d'espairs importants, et des phases de déception et de réduction des investissements connus sous le nom d'« hivers de l'IA ». Le premier hiver, à la fin des années 1960, a suivi la publication de l'ouvrage *Perceptrons* de Marvin Minsky et Seymour Papert, qui démontrait les limites techniques des réseaux de neurones du même nom, incapables de traiter des phénomènes non linéaires. Dans les années 1980, le deuxième hiver a porté sur les systèmes experts, sur lesquels trop d'espairs avaient été fondés. Il s'agissait de modéliser et d'exploiter, souvent par un formalisme de raisonnement logique, la connaissance des experts afin de répondre à des problèmes de diagnostic, de conception, de supervision dans de nombreux domaines d'application (médecine, procédés de production, configuration d'ordinateurs, prospection minière, etc.).

Dans les années 1990 le superordinateur Deep Blue d'IBM et ses médiatiques parties d'échecs masquent une période creuse pour l'IA. C'est la diffusion vers le grand public du web, à partir des années 2000, qui relance la recherche et le développement de l'IA. D'ailleurs on constate aujourd'hui que les plus grands acteurs de l'IA sont ceux de l'Internet : Google, Amazon, Facebook, Apple, Microsoft et les géants chinois Baidu, Alibaba, Tencent et Xiaomi.

Citons certains faits. En 2011, le programme Watson conçu par IBM dans le but de répondre à des questions formulées en langage naturel, gagne au célèbre jeu télévisé Jeopardy !. Il

comprend les questions, répond en langage naturel et puise dans une énorme base de données et de connaissances. En 2012, l'apprentissage profond, une manière de concevoir les réseaux

Une IA est un programme visant à effectuer, au moins aussi bien que des humains, des tâches nécessitant un certain niveau d'intelligence.

neuronaux avec un grand nombre de couches internes, qui repose sur la disponibilité de grandes bases de données annotées, d'algorithmes sophistiqués d'optimisation et de très grandes plates-formes de calcul haute performance, bat les algorithmes classiques de reconnaissance d'images. C'est une date très importante. Citons également quelques exemples, comme :

- DeepMind, rachetée par Google en 2014, qui a démontré ses capacités en maîtrisant un ensemble de jeux sur console Atari avant de battre, deux années de suite, les meilleurs joueurs mondiaux de Go avec le système AlphaGo,

- Les logiciels de reconnaissance vocale intégrés dans les assistants vocaux (Siri, Alexa, Cortana, OK Google) et de traduction en temps réel (Microsoft Skype Translator par exemple) utilisent de l'IA puissante.

Ces exemples ne doivent cependant pas faire oublier les réponses parfois absurdes que chacun de nous a pu entendre avec des applications de commande vocale ou des assistants virtuels.

L'IA est intégrée dans des activités et des entreprises très différentes, du groupe Thalès qui la développe pour quasiment toutes ses branches de sécurité, défense, aéronautique, spatial, etc., à des start-ups dans le secteur médical développant notamment des applications d'analyse et d'aide au diagnostic au service des professionnels de santé et des patients ; dans la banque, la relation client qui est transformée ; dans le transport, la voiture autonome même si le secteur reste émergent et les défis scientifiques nombreux. On peut aussi parler de lunettes intelligentes, de gestion de drones, d'aide à la logistique, etc. Les applications sont partout.

Il y a deux grandes sortes d'IA. Une première méthode très coûteuse mais bien maîtrisée : les modèles « faits à la main » (règles, modèles logiques, graphes, ontologies qui représentent les connaissances d'un domaine) qui sont explicables puisque l'ensemble du raisonnement conduit par le système peut être analysé et décortiqué, ayant été conçu par des humains. Et il y a les systèmes d'apprentissage automatique à partir des données. Ce sont les réseaux neuronaux, les algorithmes d'évolution artificielle, les systèmes auto-immunes...

Tout l'enjeu est la confiance entre l'homme et ces systèmes.

Les réseaux neuronaux apprennent à partir de très grandes quantités de données, souvent annotées par des opérateurs humains (« *il y a un chat dans l'image* », « *c'est une cellule cancéreuse* », « *dans cette situation, je freine* »). Combinant grands volumes de données, algorithmes d'apprentissage et formidable puissance de calcul, ils peuvent atteindre, sur des cas très précis, des performances égales ou supérieures aux meilleurs opérateurs humains.

On parle beaucoup de l'impact sociétal de l'IA. Je cite en référence le « guide de survie » édité par Microsoft : placer l'humain au cœur de la machine (l'IA au service de l'humain), contrôler l'IA (contrôle vigilant des programmes, véracité des données), démocratiser

Hiérarchie de mots-clés pour l'IA établie par l'Inria

Toutes les typologies se valant, il a été décidé d'opter pour la hiérarchie de mots-clés proposée par la communauté des responsables d'équipes-projets Inria, afin de mieux identifier leurs contributions aux sciences du numérique en général. Celle-ci ne couvre que les domaines de l'IA sur lesquelles l'Inria effectue des recherches ; il y a encore d'autres technologies, voir par exemple le rapport publié par #FranceIA en 2017, ou le rapport publié en avril 2018 par l'Académie des Technologies (source : *IA Les défis actuels et l'action d'Inria*, sept. 2016).

Connaissances Bases de connaissances. Extraction & nettoyage de connaissances. Inférence. Web sémantique. Ontologies.

Apprentissage automatique Apprentissage supervisé. Apprentissage (partiellement) non-supervisé. Apprentissage séquentiel et par renforcement. Optimisation pour l'apprentissage. Méthodes bayésiennes. Réseaux de neurones ou neuronaux. Méthodes à noyau. Apprentissage profond. Fouille de données. Analyse de données massives.

Traitement du langage naturel

Traitement des signaux Parole. Vision (reconnaissance d'objets, reconnaissance d'activités, recherche dans des banques d'images et de vidéos, reconstruction 3D et spatio-temporelle, suivi d'objets et analyse des mouvements, localisation d'objets, asservissement visuel.

Traitement (y compris les véhicules autonomes) Conception. Perception. Décision. Action. Interactions avec les robots (environnement/humains/robots). Flottes de robots. Apprentissage des robots. Cognition pour la robotique et les systèmes.

Neurosciences, sciences cognitives Compréhension et stimulation du cerveau et du système nerveux. Sciences cognitives.

Algorithmique de l'IA Programmation logique et ASP. Déduction, preuve. Théories SAT. Raisonnement causal, temporel, incertain. Programmation par contraintes. Recherche heuristique. Planification et ordonnancement.

Aide à la décision

l'IA (accessible à tous). Tout l'enjeu est la confiance entre l'homme et ces systèmes. De nombreuses initiatives nationales et internationales s'intéressent de près aux conséquences possibles d'une large diffusion de l'IA dans la société.

Les débats sur l'IA ont commencé au vingtième siècle avec les lois de la robotique d'Isaac Asimov

mais s'intensifient aujourd'hui en raison des récentes avancées, décrites précédemment, dans le domaine. Selon la théorie de la singularité technologique, une ère de domination des machines sur l'homme verra le jour lorsque les systèmes d'intelligence artificielle deviendront super-intelligents. Mais les spécialistes de l'IA affirment que cet horizon est bien lointain, personne ne peut dire s'il sera atteint un jour.

Plus nous développons l'intelligence artificielle, plus nous devons collectivement réfléchir à son impact. Nous en attendons d'abord de nombreux bénéfices : amélioration du diagnostic médical et développement de la médecine personnalisée ; optimisation de l'usage de l'énergie, et meilleure maîtrise de l'environnement ; exécution de tâches peu valorisantes pour l'humain

Plus nous développons l'intelligence artificielle, plus nous devons collectivement réfléchir à son impact.

de manière plus économique et plus sûre (comme conduire une voiture sur autoroute), etc. Mais, nous savons aussi qu'utilisée à grande échelle, l'intelligence artificielle peut comporter

des risques et constituer quantité de défis pour les humains, en particulier si les intelligences artificielles ne sont pas conçues et encadrées de façon à respecter et protéger les humains. Si, par exemple, l'optimisation et les performances sont le seul objectif, cela peut conduire à des situations où les utilisateurs sont instrumentalisés, abusés, manipulés, etc. par des agents. C'est aussi le message de notre pays dans son programme national pour l'IA « *AI for Humanity* », suite au rapport coordonné par Cédric Villani fin mars 2018¹ : la recherche en IA doit être globale, pluridisciplinaire et inclure une réflexion de fond sur la société que nous construisons avec elle.

¹ Marc Schoenauer, Yann Bonnet, Charly Berthet, Anne-Charlotte Cornut, François Levin, Bertrand Rondepierre, « *Donner un sens à l'intelligence artificielle : pour une stratégie nationale et européenne* ».