



MAIS L'IA C'EST QUOI ÇA VIENT D'OÙ ?

Avant la quête de l'intelligence artificielle, il y avait la quête de la vie artificielle. De la colombe d'Archytas qui battait des ailes jusqu'à l'avènement des robots au début du XX^e siècle, en passant par le canard et le joueur de flute de Vaucanson, les inventions de Leonardo Da Vinci et les automates de Pierre Jaquet-Droz, l'humanité s'est ingéniée à essayer de recréer la vie à travers des constructions mécaniques ingénieuses. Mais cette apparence de vie se contentait de reproduire des actions prédéfinies, immuables, et surtout sans prendre de décisions. Il faudra attendre l'essor des bases théoriques du traitement de l'information (informatique, cybernétique, formalisation de l'information et du signal, toutes nées durant la 2^{ème} guerre mondiale) pour voir se développer en parallèle une réflexion sur la possibilité de créer des machines ou des algorithmes capables d'imiter l'intelligence humaine, c'est-à-dire reconnaître, évaluer, prédire et décider.

Le domaine de l'IA a de fait été forgé par la confluence de deux axes de recherche ; d'une part des théoriciens en neuroscience essayant de proposer des architectures expliquant le fonctionnement de notre système nerveux (l'IA connexionniste), et d'autre part des logiciens développant des algorithmes de raisonnement visant à prendre des décisions (l'IA symbolique). Ces deux communautés ont alterné succès et

déconvenues, se partageant les financements et le devant de la scène.

En 1943, Warren McCulloch et Walter Pitts ont proposé le premier modèle de neurone artificiel, qui allait devenir la brique élémentaire des réseaux de neurones, l'architecture qui est derrière la plupart des outils d'IA utilisés de nos jours, comme *ChatGPT*. Les ordinateurs programmables n'existant pas encore, les premières implémentations de ces "neurones" furent des machines dédiées, comme le SNARC construit en 1951 par Marvin Minsky et Dean Edmonds (l'auteur a eu la chance d'être en contact en 2009 avec Marvin Minsky qui – faisant toujours montre d'un esprit curieux – a bien voulu discuter d'un de ses travaux dans un domaine tout à fait différent. Un honneur rare). Le terme "Intelligence Artificielle" a été créé par John McCarthy pour le *Dartmouth Summer Research Project on Artificial Intelligence* qu'il organise en 1956 avec Minsky, événement considéré comme marquant la naissance du domaine en tant que tel. Le terme semble avoir été choisi en partie pour sa généralité et en partie pour des raisons politiques, afin d'éviter d'autres termes en vigueur à l'époque. Durant l'école d'été, Allen Newell et Herbert Simon (futur prix Nobel d'économie) présentèrent le *Logic Theorist*, probablement le premier programme d'IA, capable de prouver des théorèmes mathématiques, parfois de manière plus claire et élégante que leur auteur original.

Les avancées des réseaux de neurones continuèrent, notamment avec les travaux de Frank Rosenblatt. En 1958, il décrit une architecture utilisant une couche de plusieurs neurones McCulloch-Pitts améliorés (des *perceptrons*) qui apprend en modulant les poids synaptiques (les connexions entre les neurones). Le système est programmé sur un ordinateur IBM, mais également implémenté "en dur" dans une machine dédiée à la reconnaissance d'images. Cependant, la difficulté de faire apprendre au système des motifs complexes lassa les financeurs. Minsky et Seymour Papert enfoncèrent le clou avec leur livre *Perceptrons* en 1969, montrant qu'une seule couche de neurones ne pouvait pas conduire de calculs non-linéaires, signant ainsi le début de l'"hiver de l'IA". Des perceptrons multicouches existaient déjà, comme le système Tobermory du même Rosenblatt, constitué de 4 couches et 12 000 poids synaptiques, dédié à la reconnaissance vocale. Cependant, l'algorithme qui allait rendre possible l'apprentissage profond des réseaux à plusieurs couches (*deep learning* en anglais), à savoir la *backpropagation* (rétropropagation du gradient) bien que découvert indépendamment par Seppo Linnainmaa et Paul Werbos au début des années 1970 ne sera popularisé que dans les années 1980, grâce à des gens comme John Hopfield et Geoffrey Hinton (prix Nobel de physique 2024 pour leurs travaux sur les réseaux de



Dartmouth Summer Research Project on Artificial Intelligence, 1956.

De gauche à droite, Oliver Selfridge, Nathaniel Rochester, Ray Solomonoff, Marvin Minsky, Peter Milner, John McCarthy, Claude Shannon. Reproduction avec l'aimable autorisation de la famille Minsky.

neurones). De plus les ordinateurs de l'époque manquaient de puissance, de mémoire, et les jeux de données étaient rares et limités.

Les deux dernières décennies du XX^e siècle ont donc vu le retour en force de l'IA symbolique, en particulier des systèmes experts. Ces systèmes, dont le pionnier fut Edward Feigenbaum dans les années 1960, se fondent sur des règles décidées d'avance par leurs concepteurs, historiquement implémentées par des conditions imbriquées. Les systèmes experts ont été – et sont toujours – utilisés dans de nombreux domaines dont la santé, la chimie, l'astronomie, et la défense. Ils sont apparus au grand public quand le système IBM *DeepBlue* a battu le champion d'échec Gary Kasparov en 1996, et le système IBM Watson a gagné au jeu américain *Jeopardy* ! en 2011. Mais la victoire de *Watson* était son chant du cygne, les tentatives de le réutiliser dans différents domaines échouant les unes après les autres.

En silence, l'utilisation d'algorithmes efficaces d'apprentissage couplés à des capacités informatiques toujours croissantes allaient signer le retour

sur le devant de la scène de l'apprentissage machine, et en particulier des réseaux de neurones. Bien qu'éclipsés à la fin du XX^e siècle par les approches Bayésiennes, les graphes, les modèles de Markov et autres *Support Vector Machines*, les réseaux de neurones montaient en puissance. L'auteur utilisait à la fin des années 1990 des perceptrons multicouches en série pour prédire la structure des récepteurs de la nicotine dans le cerveau. De plus, des architectures complètement nouvelles, comme les réseaux à neurones de convolution, les auto-encodeurs ou les réseaux récurrents voyaient le jour, attendant juste l'existence de jeux de données qui démontreraient leur puissance.

En 1998, le modèle *LeNet-5* – une évolution du modèle présenté par Yann Le Cun en 1989 pouvant reconnaître les chiffres écrits à la main – fut le premier réseau de convolution déployé en vie réelle. L'utilisation des processeurs graphiques a ensuite permis l'entraînement de réseaux profonds à neurones de convolution comportant des centaines de milliers de neurones et des dizaines de millions de paramètres (notamment par Geoffrey Hinton, voir

plus haut). Du pilotage de véhicules autonomes à la détection des tumeurs, ces outils allaient bientôt égaler, voire dépasser, les capacités de l'être humain. L'arrivée des « grands modèles de langage » enfonça le clou, avec des succès médiatiques comme *AlphaGO*, battant les humains au jeu de Go en 2015, ou *AlphaFold2*, arrivant à prédire la structure tridimensionnelle des protéines à partir de leur séquence, achevant ainsi une quête scientifique de près de 50 ans (prix Nobel de chimie 2024). En 2022, *ChatGPT* a mis dans la main de tout un chacun un assistant d'IA capable de répondre à n'importe quelle question (de manière parfois discutable), très vite suivi par une pléthore de modèles génératifs produisant textes, images, musiques, et plus.

L'IA a désormais envahi la quasi-totalité des entreprises humaines, de façon parfois très visible, souvent invisible. La puissance et la généralité de ces modèles posent de nombreuses questions, en termes de sécurité, d'éthique, de morale, mais aussi de maîtrise de l'information et d'environnement. Ces outils augmentent notre productivité, mais menacent également des pans entiers de l'emploi et attisent les peurs. Ils ouvrent aussi une fenêtre sur les prochaines étapes, comme l'Intelligence Artificielle Générale, capable d'effectuer ou d'apprendre pratiquement n'importe quelle tâche cognitive propre aux humains. ■