

## Brève histoire de l'Intelligence Artificielle

Olivier Boisard,

Ingénieur-conseil et enseignant à l'École Centrale de Lille

[olivier@olivier-boisard.net](mailto:olivier@olivier-boisard.net), [www.olivier-boisard.net](http://www.olivier-boisard.net)

Mars 2020

### Résumé :

*L'invention de la machine de Turing, en 1936, rend possible le développement de l'informatique et de l'Intelligence Artificielle (IA). Les premiers ordinateurs seront toutefois plutôt destinés au calcul numérique ou à la gestion de bases de données. Il faudra attendre les années 1970 pour que l'IA prenne véritablement son essor, avec l'apparition de nouveaux langages de programmation et de logiciels tels que les systèmes experts. Aujourd'hui, ce sont les réseaux neuronaux qui s'imposent, et cette réussite soulève des questions sur l'usage de l'IA, qui dépassent très largement le seul cadre de l'informatique.*

Des machines destinées à simuler les comportements humains ont été imaginées bien avant la création de la robotique et de l'informatique. Mais il s'agissait alors d'automates, dont chaque action, ou chaque réaction dans un contexte donné, devait être préalablement pensée et intégrée explicitement dans des mécanismes complexes. Le XVIII<sup>ème</sup> siècle, et plus encore le XIX<sup>ème</sup> - parfois désigné comme l'*âge d'or des automates* - ont vu se multiplier ces machines étranges et spectaculaires.

Un exemple emblématique de cet âge d'or est le célèbre automate joueur d'échec - le « *Turc Mécanique* » - présenté au public durant des décennies en Europe puis en Amérique, et dont on raconte qu'il provoqua en 1809 la colère de Napoléon - pourtant réputé bon joueur d'échec - furieux d'avoir dû concéder une défaite cuisante.

Mais, comme le démontre Edgar Poe dans la nouvelle *Le joueur d'échec de Maelzel*<sup>1</sup>, cet automate ne pouvait être qu'une imposture, et derrière des mécanismes d'apparence se cachait nécessairement un joueur humain ...

Pourtant, dans sa volonté de démontrer rationnellement l'impossibilité d'une telle machine, Edgar Poe ne voit pas assez loin. Aujourd'hui, le moindre ordinateur a la capacité de faire tourner un logiciel dont les performances, pour limitées qu'elles soient aux yeux d'un grand joueur d'échec, sont suffisantes pour mettre en difficulté nombre d'amateurs.

Alors, où se situe la différence entre un automate et une machine jouant « intelligemment » ? Est-ce une simple question de puissance de calcul, désormais permise par la miniaturisation extrême des micro-processeurs ? Ou, plus fondamentalement, est-ce une différence de nature entre des mécaniques qui se bornaient à exécuter des tâches prédéfinies, et des circuits capables de faire appel à des modes de fonctionnement bien plus subtiles ?

### De l'automate à la machine de Turing

Clairement, c'est bien d'une différence de nature qu'il s'agit : lorsqu'en 1936 Alan Turing conçoit sa célèbre *machine de Turing*, sont réunis tous les éléments qui donneront naissance à l'informatique contemporaine, grâce à des engins n'ayant plus rien à voir avec des automates.

Ce qui caractérise ces machines, au demeurant assez simples à décrire sur le plan théorique, est leur capacité à manipuler des symboles ayant indifféremment le statut de *données* (la *data*, autrement dit l'information qui est manipulée), ou celui de programme (le *process*, autrement dit ce qui dicte à la machine la façon de manipuler la *data*). En d'autres termes, cela revient à considérer une machine capable de modifier elle-même son propre mécanisme, soit, si cette modification est effectuée en vue d'une fin particulière, de développer *une capacité d'apprentissage*.

Un automate joueur d'échec est inconcevable car son concepteur devrait préalablement anticiper toutes les parties du jeu - leur nombre est certes fini, mais incommensurable - et traduire chacune par un mécanisme particulier. Alors qu'une machine de Turing correctement conçue a la capacité de jouer en s'adaptant à chaque nouvelle configuration de l'échiquier, sans avoir à anticiper toutes les parties possibles : comme un humain, son jeu ne sera peut-être pas parfait, mais chaque erreur se transformera en expérience acquise, permettant de mieux jouer la partie suivante. En soi, l'apprentissage et la capacité d'adaptation sont déjà deux qualités caractéristiques de l'intelligence !

## L'ordinateur, machine à calculer ou machine logique ?

Parmi les précurseurs de l'informatique et de l'intelligence artificielle (IA) - terme inventé assez tardivement, en 1956, par John McCarthy et Marvin Minsky - il faut également citer des personnalités telles que Claude Shannon, créateur de la théorie de l'information, dont un des apports majeurs est le choix du *bit*, prenant les valeurs « 0 » ou « 1 », symboles de base manipulés par cette machine de Turing sophistiquée qu'est l'ordinateur.

Notons que les bits sont représentés par des chiffres - permettant de former des nombres *en base 2* - ce qui a conduit longtemps à considérer, de façon réductrice, l'ordinateur comme une puissante machine à calculer. C'est d'ailleurs ce que signifie la racine latine *comput* - de *Compuctus*, *Calcul* - sur laquelle est construit le terme anglo-saxon *computer*.

Pourtant, le 0 et le 1 peuvent tout autant signifier « faux » et « vrai ». Le cœur de tous les microprocesseurs est constitué d'une multitude de portes logiques - on les chiffre en milliards dans les ordinateurs actuels - qui ne font en définitive rien de plus qu'effectuer des traitements logiques élémentaires « ET », « OU », et « NON », selon des principes qui trouvent leur origine dans la logique aristotélicienne et qui seront formalisés dans l'algèbre de Boole au XIX<sup>ème</sup> siècle. Autrement dit, *autant que des calculs, un ordinateur effectue des traitements logiques, et possède donc une capacité de déduction.*

Machine à calcul ou machine logique ? C'est à l'utilisateur d'en décider, et peut-être est-ce là le fond de la question philosophique soulevée par l'IA : l'ordinateur est fondamentalement une machine de Turing qui se contente de manipuler des symboles abstraits, *comme des choses en soi vides de toute signification auxquelles seul l'esprit humain attribue un sens.*

Dès 1905, soit près de trente ans avant les travaux de Turing, le grand mathématicien Henri Poincaré formulait cette idée de façon prémonitoire, et avec un certain humour, imaginant une machine logique capable d'effectuer des raisonnements mathématiques « où l'on introduirait les axiomes par un bout pendant qu'on recueillerait les théorèmes à l'autre bout, comme cette machine légendaire de Chicago où les porcs entrent vivants et d'où ils ressortent transformés en jambons et en saucisses.<sup>2</sup> »

## Des débuts prometteurs

Dans les années 50, le film de science-fiction *Forbidden Planet* met en scène *Robby* le robot. Comme son cousin le « droïde de protocole » C3PO de *Star Wars* apparu sur les écrans vingt ans plus tard, c'est l'archétype de la machine sympathique, capable de communiquer par la parole avec les humains, voire même d'exprimer quelques émotions.

C'est à cette même époque, en 1957, que le linguiste Noam Chomsky publie sa thèse<sup>3</sup>, décrivant les *grammaires génératives, ou grammaires formelles* : en substance, celles-ci consistent à modéliser une langue en un ensemble d'axiomes d'une part (les catégories syntaxiques telles que le verbe, le nom, l'article, etc., ainsi que tous les mots du dictionnaire), et de règles syntaxiques d'autre part, qui ne sont rien de plus que les règles de grammaire qu'on apprend à l'école. Il sera ainsi possible, en appliquant ces règles aux axiomes, de déduire toutes les phrases grammaticalement correctes de la langue.

Sous cet angle, il existe une analogie parfaite entre la structure d'une langue et une théorie mathématique. Celui qui parle ne se contente donc pas de faire de la prose sans le savoir, *il déduit aussi des « théorèmes » à partir d'« axiomes »*. A l'inverse, s'il reconnaît une phrase écrite dans sa langue (soulignons qu'il n'est pas ici question du *sens* de cette phrase, mais seulement de sa forme grammaticale), *il effectue mentalement l'équivalent de la démonstration d'un théorème !*

On pourrait dire qu'à la fin des années 1950, tous les concepts sont réunis pour concevoir *Robby* le robot parlant, et des experts sans doute un peu trop optimistes prédisent à l'époque son apparition pour la fin de la décennie suivante. Cela prendra en réalité plus de 60 ans. Ce n'est qu'aujourd'hui, avec les *assistants personnels intelligents*, qu'une machine grand public peut engager un semblant de dialogue avec des humains - et encore, pour étonnante qu'elle soit sur le plan technologique, cette communication trouve-t-elle très vite ses limites ...

Si la recherche fondamentale en IA se poursuit dans les années 1960, elle ne trouve pas encore de véritable application dans l'industrie. Celle-ci s'intéresse alors prioritairement aux deux capacités premières de l'ordinateur : gérer de l'information dans des bases de données, avec le développement des « Business Machines », et effectuer des calculs mathématiques complexes, avec au premier rang des applications dans les domaines scientifique ou militaire.

Il faudra attendre les années 1970 pour que se produise une nouvelle avancée majeure, permise par l'apparition de nouveaux langages de programmation. Ceux-ci - d'ailleurs souvent basés sur des approches linguistiques - donneront naissance à la *Programmation Orientée Objet* (POO), aujourd'hui d'un usage généralisé : alors que les premiers programmes informatiques étaient composés de lignes numérotées, interconnectées par des instructions « GoTo » ou « GoSub » qui rendaient leur maintenance difficile et source de bugs multiples, la POO offre un mode de programmation plus souple - on pourrait même dire élégant - ouvrant plus facilement la voie à la programmation de logiciels d'Intelligence Artificielle.

### Déceptions des années 1980

Les *Systèmes Experts* se développent durant les années 1980. Ces logiciels, visant notamment le tout nouveau marché des micro-ordinateurs personnels, sont conçus pour proposer des *outils d'aide à la décision*. Ils reposent sur des principes de logique formelle, dont l'exemple le plus élémentaire est le syllogisme.

Tous les syllogismes, en effet, n'ont pas la transparence de celui par lequel on démontre que Socrate est mortel ... Au XIX<sup>ème</sup> siècle, Lewis Carroll s'efforçait de trouver de tels exemples de déductions vraies sans être évidentes :

« *Tous les lapins non gourmands sont noirs ;  
Aucun vieux lapin n'est exempt de gourmandise.  
Donc :  
Certains lapins noirs ne sont pas vieux.*<sup>4</sup> »

Plus sérieusement, un système expert est capable de gérer un grand nombre de propositions, manipulées par des règles de déduction susceptibles de conduire à des conclusions non-intuitives. Une application possible est l'*aide au diagnostic*, chaque proposition portant par exemple sur la description d'un symptôme, et la conclusion caractérisant une pathologie. Le système expert ne se contentera pas d'établir un diagnostic, il le fera en ne posant au patient que le minimum de questions nécessaires.

Pourtant, malgré quelques réussites notables, les systèmes experts n'ont pas su répondre aux attentes qu'ils avaient suscitées, au point même de provoquer une remise en cause, aux yeux de certains observateurs, de l'intérêt pratique de l'Intelligence Artificielle.

La puissance des ordinateurs de l'époque - *a fortiori* des micro-ordinateurs - était probablement insuffisante pour effectuer dans de bonnes conditions ces traitements complexes. Plus fondamentalement, sur le plan théorique, pouvait-on aussi s'interroger sur la difficulté - voire dans certains cas l'impossibilité ? - à retranscrire explicitement, dans toute sa subtilité, le savoir-faire d'un expert en un ensemble de règles de déduction.

### Nouvel essor des années 1990

En décembre 1991, le *High Performance Computing Act*<sup>5</sup>, adopté par le Congrès des États-Unis, allait marquer l'essor des *autoroutes de l'information*, autrement dit *internet*. Ce réseau avait été créé plus de vingt ans auparavant, mais était resté confiné jusqu'alors - hors applications de défense - essentiellement dans le milieu universitaire et dans quelques entreprises de pointe.

Internet va avoir deux incidences sur le développement de l'IA : tout d'abord, la capacité d'interconnecter des ordinateurs, et donc d'augmenter potentiellement leur capacité globale de traitement. Mais aussi - et surtout - de faciliter l'accès à des quantités considérables de données - la *data*, qui conduira plus tard au *big data* - nécessaires à l'apprentissage de nouveaux algorithmes.

Dans le même temps, la puissance des micro-processeurs se développe, et des logiciels toujours plus performants conduisent notamment, de façon spectaculaire, le super-ordinateur *Deep Blue* à battre en 1997 le champion du monde d'échec Garry Kasparov.

Mais une autre révolution se prépare, qui prendra véritablement son essor dans les années 2000, avec les *réseaux neuronaux*.

### Succès des réseaux neuronaux

Les réseaux de neurones ne sont pas nouveaux. Les premières publications scientifiques sur le sujet datent de 1950, et sont inspirées tout naturellement par l'observation du neurone biologique - même s'il s'agit là d'une métaphore qui n'a probablement que peu à voir avec la complexité de cette cellule.

Le neurone de l'intelligence artificielle, dans son principe, n'a rien de complexe : son rôle se borne à recueillir des données qu'il va traiter par des fonctions mathématiques élémentaires (certaines sont de simples combinaisons linéaires) pour produire de nouvelles données ... Celles-ci, produites par un premier ensemble de neurones - une première *couche* - peuvent à leur tour être utilisées comme source pour une seconde couche, et ainsi de suite... Le réseau est alors constitué de multiples niveaux, de plus en plus performants et *profonds*, notion qui a conduit au terme actuel de *Deep Learning*.

L'intérêt pratique de ce réseau est la simplicité de mise en œuvre de sa phase d'apprentissage : préalablement alimenté par une très grande quantité de données, des dizaines de milliers de photos de chats par exemple, il intégrera dans ses neurones l'information utile pour détecter, sur tout nouveau cliché qui lui est soumis, la présence de l'animal ...

Contrairement aux systèmes experts, aucune compréhension du sujet n'est ici nécessaire : un programmeur humain n'aura pas à traduire le concept « chat » en d'hypothétiques règles logiques capables d'analyser le contenu d'une photo. Seule compte la quantité de données brutes utilisées lors de l'apprentissage.

Il faut cependant admettre que les réseaux de neurones, compte tenu de la fiabilité insuffisante de leurs premiers résultats, sont longtemps restés dans les laboratoires de recherche. Un logiciel capable de reconnaître une adresse postale manuscrite avec une fiabilité de 80 ou 90 % n'aurait qu'une utilité relative. Pour être opérationnel, il lui faut, au minimum, atteindre un seuil de 99,9 %...

Cet objectif deviendra accessible dans les années 2000. Parallèlement aux micro-processeurs (le *Central Processing Unit*, ou CPU), on exploite alors plus systématiquement les ressources de l'autre puce d'un micro-ordinateur, *a priori* destinée à gérer son interface graphique (le *Graphic Processing Unit*, ou GPU), dont on prend conscience que le mode de calcul est particulièrement bien adapté aux traitements neuronaux. Et à partir de 2005 apparaissent sur Internet des bases d'images référencées - de vastes collections de photographies d'animaux et d'objets de toutes sortes... -, disponibles en libre accès, et destinées à alimenter l'apprentissage de nouveaux algorithmes.

La base la plus célèbre est *ImageNet*, présentée en 2009 par l'université de Princeton, comptant plusieurs millions de photographies. Lors de la compétition ILSVRC (*ImageNet Large Scale Visual Recognition Competition*), où s'affrontent chaque année les algorithmes de reconnaissance d'images exploitant cette base, les meilleurs taux de réussite seront de 99,5 % en 2010, puis de 99,7% en 2011. *Ils passeront la barre des 99,9% en 2014...*

Aujourd'hui les réseaux neuronaux s'imposent, au point qu'on les assimile souvent - à tort - à l'Intelligence Artificielle. Les smartphones de dernière génération intègrent des circuits entièrement dédiés aux calculs neuronaux, utilisés pour déverrouiller l'appareil par reconnaissance faciale. En 2015, *AlphaGo* est le premier programme informatique - basé sur des algorithmes neuronaux - capable de battre un joueur humain au jeu de Go. En 2017, une variante de ce programme, *AlphaZero*, dépasse le niveau des meilleurs logiciels de jeu d'échec, avec - contrairement à son prédécesseur *Deep Blue* - pour seules et uniques données de base les règles du jeu, son expérience étant acquise en jouant contre lui-même trois jours durant un nombre incalculable de parties.

### **Conclusion : quelques questions soulevées par l'IA**

Avec de tels outils, l'IA permet désormais de concevoir des machines suffisamment puissantes pour intervenir, de façon directe ou indirecte, dans des processus de décision. Cela pose, par nature, des questions d'ordre éthique : de quel *droit* une machine peut-elle intervenir dans un tel processus ?

Les interrogations soulevées par l'usage des machines ne sont pas nouvelles, et constituent même depuis toujours un thème majeur de la science-fiction. Excepté qu'il s'agit aujourd'hui d'une réalité : *il est essentiel de « garder la main » sur la puissance de l'IA*, comme le rappelleront en 2015 des personnalités du monde scientifique ou de grandes entreprises technologiques, telles que le physicien Stephen Hawking ou l'ingénieur-entrepreneur Elon Musk, dans une *lettre ouverte* <sup>6</sup> mettant en garde contre certains risques de l'IA.

Ces questions ne doivent pas être considérées comme un réquisitoire contre l'IA. Mais il serait irresponsable de les ignorer. On peut en relever quelques-unes :

- **L'IA produit-elle des boîtes noires ?** Est-il possible de justifier le résultat d'un traitement complexe, autrement-dit de reconstituer de façon précise et compréhensible tout le cheminement qui a conduit à sa production ? Contrairement aux systèmes basés sur des inférences logiques, il est difficile de reconstituer le mode de fonctionnement interne d'un réseau de neurones. Au-delà de l'efficacité

opérationnelle avérée de ces systèmes, des progrès scientifiques sont à faire pour mieux comprendre et maîtriser leur développement.

- **Quel est le statut de la data exploitée par l'IA ?** Le citoyen est-t-il véritablement conscient de l'usage qui peut être fait de la *data* qu'il produit et qui alimente des algorithmes d'IA ? Cette *data* est-elle du domaine de la vie privée, que la loi a vocation à protéger (en Europe, le *Règlement Général sur la Protection des Données*, ou RGPD, tente d'aller dans ce sens) ? Ou, selon une approche plus anglo-saxonne, l'individu est-il seul propriétaire de ses données et donc libre d'en disposer, y compris en les cédant s'il le souhaite à un opérateur privé pour bénéficier, dans une logique commerciale, d'un service gratuit ou à prix réduit ?
- **L'IA ne réduit-elle pas la réalité à des phénomènes quantifiables ?** La *data* qui alimente un logiciel d'IA ne risque-t-elle pas de limiter la perception de chaque domaine qu'elle aborde à des mesures chiffrées : l'évaluation de la qualité d'un service à une simple note de 0 à 10 attribuée par des consommateurs, la santé à un ensemble de mesures biométriques effectuée par un bracelet connecté, l'intelligence à un Quotient Intellectuel, ou l'individu à un « crédit social » actualisé par un système de reconnaissance faciale détectant dans la cité ses supposées « incivilités » ... ? Et, au-delà, ces données une fois massivement récoltées et traitées, ne risquent-elles pas d'imposer leur moyenne comme une norme, écartant parce qu'« a-normaux » les cas atypiques ou simplement originaux ?
- **Quelle place souhaite-t-on donner à l'IA dans la société ?** L'IA est - et sera de plus en plus - en capacité d'effectuer des tâches jusqu'alors assurées par des humains, et donc de supprimer des emplois, y compris des emplois qualifiés. C'est un choix de société : jusqu'à quel point souhaite-t-on effectuer cette substitution ?
- **Peut-on anticiper les développements futurs de l'IA ?** La recherche en Intelligence Artificielle n'a pas de raison de s'arrêter, et, succédant à l'IA dite *faible* d'aujourd'hui, c'est-à-dire ne pouvant se confondre - malgré son efficacité sur certaines tâches - avec un esprit humain, on peut s'interroger sur l'hypothèse de l'émergence, demain ou après-demain, d'une *IA forte* aux capacités bien plus élevées. Une telle IA est-elle concevable, et même souhaitable ?

Toutes ces questions ne portent pas tant sur la nature en elle-même de l'Intelligence Artificielle, que sur l'usage qui peut en être fait. En ce sens, elles rejoignent les interrogations qui ont toujours accompagnées, à juste titre, l'émergence de nouvelles technologies. Ces réflexions critiques doivent se poursuivre de façon constructive et responsable, sans jamais perdre de vue les bénéfices apportés par une IA maîtrisée, qui *sera d'une aide précieuse - pour ne pas dire nécessaire - face aux défis majeurs du XXIème siècle.*

---

<sup>1</sup> Poe E.A., traduction française Baudelaire C. Histoires extraordinaires. Nouvelles histoires extraordinaires. Michel Lévy frères, Paris 1856.

<sup>2</sup> Poincaré H. Science et Méthode, Bibliothèque de philosophie scientifique, Flammarion, 1908.

<sup>3</sup> Chomsky N. Syntactic Structures. Mouton & CO. 1957.

<sup>4</sup> Carroll L. Symbolic Logic. London Macmillan ed. 1896.

<sup>5</sup> US Congress. High-Performance Computing Act of 1991. <https://www.congress.gov/bill/102nd-congress/senate-bill/272>

<sup>6</sup> Texte collectif. Research Priorities for Robust and Beneficial Artificial Intelligence: An Open Letter. 2015. En ligne sur le site Future of Life : <https://futureoflife.org/ai-open-letter>