La boîte translucide : un éclairage sur l'intelligence artificielle

Pierre Jourlin

LA BOÎTE TRANSLUCIDE

UN ÉCLAIRAGE

SUR L'INTELLIGENCE ARTIFICIELLE

Pierre Jourlin

LA BOÎTE TRANSLUCIDE

UN ÉCLAIRAGE SUR L'INTELLIGENCE ARTIFICIELLE

Editions Universitaires d'Avignon

Collection dirigée par Aurélia Barrière

Conception et mise en page : Catherine Julia
Correction du manuscrit : Sarah Jourdren
Illustration de couverture : Pierre Jourlin, licence cc by-sa

ISBN : 978-2-35768-127-9

Éditions Universitaires d'Avignon
74, rue Louis Pasteur
84029 Avignon cedex 1 - France

Introduction : de la « matière grise »

Durant la majeure partie de mes études et dès les débuts de ma vie professionnelle, les expressions « intelligence artificielle », « apprentissage automatique », ainsi que leurs variantes anglophones étaient déjà discutées dans les milieux techniques ou technophiles, chez les *geeks* comme chez les ***nerds****[1] (Peyron, 2012). Mais, depuis quelques années, ces technologies sont effectivement utilisées – consciemment ou non – par un très large public et, dans le même temps, un nombre toujours croissant de disciplines académiques se saisissent du sujet : philosophie, économie, droit, sociologie, etc. Or, on présente souvent l'intelligence artificielle (IA) comme une boîte opaque, ce qui n'est pas vraiment démenti par les spécialistes, en particulier ceux qui développent des réseaux

1. Les mots en gras italique suivis d'un astérisque sont expliqués dans le lexique de fin de volume, pages 77-79.

de neurones artificiels (RNA). C'est naturellement une image qui peut générer des inquiétudes parfois justifiées et qui, sans aucun doute, fait couler beaucoup d'encre dans une grande variété de domaines et de disciplines. Cependant, si en tant qu'enseignant-chercheur en informatique, je reconnais bien volontiers que d'autres que moi auront plus de lumière à projeter sur les usages de cette boîte opaque, je vais tenter d'apporter ma contribution en l'éclairant un peu de l'intérieur. Ce livre vise ainsi un double objectif : constituer une introduction relativement brève, mais suffisamment large pour des personnes désirant se lancer dans des études en informatique ; faciliter le travail pluridisciplinaire pour des équipes de chercheuses et de chercheurs spécialisés dans d'autres domaines, mais concernés par les évolutions de la société dite numérique.

L'intelligence artificielle n'est pas un concept facile à cerner. En effet, lorsque nous évoquons les organes ou membres artificiels (cœur, jambe, bras, etc.), il n'est pas toujours nécessaire de définir précisément la liste des fonctions qui doivent être reproduites, réparées ou améliorées. Ces dernières sont en général bien identifiées. En revanche, c'est beaucoup plus difficile quand il s'agit

d'un organe tel que le cerveau, tant il est capable de réaliser une variété phénoménale de tâches. De plus, ces dernières ne sont pas toutes accessibles à la conscience ; réalisées inconsciemment, elles nous paraissent élémentaires, alors qu'un examen plus approfondi peut révéler de grands niveaux de complexité.

Prenons le langage : nous pouvons considérer que les mots que nous utilisons pour communiquer nos idées à autrui sont des symboles qui font référence à des objets concrets ou abstraits, à des représentations du réel et aux relations qu'ils entretiennent entre eux. Souvent, les mots sont ambigus : en fonction du contexte d'un énoncé ou de l'individu à l'origine de l'énoncé, un même mot ou une même séquence de mots peut faire référence à des choses très différentes. Même lorsque l'on a levé l'ambiguïté, il n'est pas toujours possible de lister tous les exemples correspondant à une définition et encore moins tous les exemples sortant du champ d'une définition. Ainsi, les contours du mot « intelligence » peuvent évoluer au gré des avancées des sciences et de la technologie, perdre ou gagner des représentants au fur et à mesure que la compréhension de notre propre fonctionnement cognitif s'élargit.

Par ailleurs, la démarche consistant pour une « intelligence » à s'autodéfinir ou à s'autocomprendre peut donner lieu à quantité de productions philosophiques, mathématiques et même artistiques – et j'encourage vivement les lectrices et lecteurs à les découvrir au travers de l'œuvre de Douglas Hoftstadter (Hofstadter, 1979).

De mon côté, je ne peux prétendre faire un long exposé sur le fonctionnement du cerveau humain, mais il me paraît indispensable de présenter les quelques éléments les plus généraux afin d'éclairer les problématiques que se proposent de traiter les technologies dites d'intelligence artificielle.

Olivier Houdé décrit l'architecture cognitive comme pouvant se décliner en « quatre principes fondamentaux et emboîtés : l'objet, le nombre, la catégorisation et le raisonnement » (Houdé, 2018, p.12) On sait aussi qu'à l'un de ces plus bas niveaux, le cerveau est un immense réseau, composé de 85 milliards de neurones interconnectés par des centaines de milliards de ***synapses**** et de ***cellules gliales**** et qui interagissent entre eux par l'envoi (ou non) de signaux (Agid, 2018). Pour autant, il reste encore un grand nombre de zones d'ombre. Par exemple, si l'on connaît l'existence des cellules gliales depuis longtemps, leur rôle

indispensable dans les processus cognitifs commence à peine à être reconnu et étudié.

D'un autre point de vue, il est clair que certains niveaux, comme le raisonnement, sont accessibles à l'introspection (on dit aussi la ***métacognition****). Nous sommes en effet capables d'analyser nos propres raisonnements, de constater les failles de nos théories et de les réviser. C'est même l'un des sujets de la philosophie, des mathématiques et de la logique depuis des millénaires. Cependant, de nombreux niveaux sont inaccessibles à la conscience et, même si nous avons réussi à en percer quelques secrets grâce, entre autres, à la psychologie expérimentale, puis à l'imagerie cérébrale, nous n'en avons encore qu'une représentation très floue et parcellaire (Dehaene, 2020).

Ainsi, la recherche en intelligence artificielle doit accepter que l'intelligence soit un mécanisme qui émerge de la complexité de cet immense enchevêtrement de cellules, sans pouvoir aller significativement plus loin dans sa compréhension.

Par exemple, s'il est possible d'observer et de décrire avec une certaine précision le fonctionnement d'une cellule du cerveau, le fait que les quatre principes fondamentaux de la cognition

soient emboîtés n'est pas sans conséquence. Les illusions dites d'« optique », très connues, sont autant d'exemples qui montrent que notre perception et notre représentation du monde peuvent vite s'éloigner de la réalité lorsque certaines conditions sont réunies. Les illusions auditives sont généralement moins connues, mais tout aussi surprenantes.

L'effet Shepard, par exemple, est un peu l'équivalent acoustique de l'enseigne de barbier ou de la vis sans fin. En diminuant progressivement le volume d'une note au fur et à mesure qu'elle monte en fréquence et en augmentant simultanément le volume de la note située à l'octave en dessous, on peut donner l'illusion qu'une note monte à l'infini, éternellement. Il existe même des illusions créées par la perception simultanée de signaux auditifs et visuels naturels, mais contradictoires.

Par exemple, pour obtenir un effet McGurk, on synchronise la vidéo d'une personne qui prononce le son [ga] avec la bande-son de cette même personne prononçant [da]. Lorsque l'on regarde ensuite cette vidéo en boucle, c'est le son [da] qui est perçu si nos yeux fixent l'image, et le son original [ba] si nous regardons ailleurs.

WIKIMEDIA/DOMAINE PUBLIC

FIGURE 1 : Enseigne de barbier, environ 1938.

Plus d'une centaine de biais dits « cognitifs » ont été identifiés. Ils montrent que la qualité de notre raisonnement peut être considérablement dégradée dans des circonstances particulières qui mettent en jeu nos émotions, ou parfois par de simples confusions liées aux différents niveaux de perception et d'analyse. Paradoxalement, cette architecture complexe permet de concevoir des représentations extrêmement rigoureuses, par exemple en logique ou en mathématiques. Nous verrons par la suite qu'elles nous ont permis de concevoir des machines à traiter de l'information tout aussi rigoureuses.

Néanmoins, il faut remarquer qu'une définition purement fonctionnelle de l'intelligence serait forcément normative et limitative. La diversité des formes d'intelligence est très vaste et, même à l'intérieur d'un milieu social, il est possible d'en trouver des formes très atypiques, comme le montre l'écrivain et philosophe Josef Schovanec, autiste Asperger, dans son livre intitulé *Nos intelligences multiples* (Schovanec, 2018).

Il semble donc assez évident qu'il sera extrêmement difficile de définir cette propriété du cerveau humain et *a fortiori* de la mesurer. Pourtant, le fameux quotient intellectuel (QI), qui a été

inventé par Alfred Binet et Théodore Simon en 1905 pour repérer les enfants nécessitant un enseignement adapté, a été largement dévoyé et présenté comme un outil de mesure de l'intelligence, instrumentalisé dans le passé à des fins racistes, sexistes et eugénistes. Encore aujourd'hui, le QI semble malheureusement perçu comme une mesure de l'intelligence par le grand public. Je me permets de renvoyer les lectrices et lecteurs qui souhaiteraient approfondir la question à l'un des meilleurs ouvrages de vulgarisation que j'aie pu lire sur le sujet, qui s'intitule *La mal-mesure de l'homme* de S.J. Gould (Gould, 1997).

L'expression « intelligence artificielle » est apparue quelques années après la publication de l'article d'Alan M. Turing intitulé « Computer machinery and intelligence » (Turing, 1950). Ce mathématicien britannique avait déjà une grande notoriété à l'époque, car dès 1936, il avait défini les bases théoriques de l'informatique avec un modèle abstrait, une machine « universelle », capable de répondre non pas à un problème particulier mais à une infinité de problèmes, et qui prendra son nom : « la machine de Turing ». Quelques années plus tard, ses travaux permirent le développement du dispositif Enigma qui

rendit possible le décryptage des messages échangés par les armées nazies et contribua grandement à la victoire des Alliés lors de la Seconde Guerre mondiale.

Or, dans son article de 1950, Alan M. Turing explique qu'il est impossible de déterminer si un être vivant peut « penser » et *a fortiori* si une machine peut être « intelligente ». Il propose donc de chercher la réponse à une autre question : la machine serait-elle capable, dans un jeu d'imitation, de nous faire croire qu'elle est un être humain ? Ce jeu d'imitation, qui prendra le nom de « test de Turing », se décline évidemment en une myriade de sous-problèmes liés à la grande diversité des informations que notre cerveau peut traiter. Par conséquent, depuis les années 1960, c'est un peu ce Graal de la recherche en informatique que l'on s'efforce d'atteindre. Un chemin important a été parcouru durant les soixante années qui ont suivi la publication de l'article de Turing; nous verrons dans le prochain chapitre les approches dont le succès n'a, à ce jour, pas été démenti.

Pour en savoir plus

- DEHAENE S., *Le code de la conscience*, Paris, Odile Jacob, 2014.
- HOFSTADTER D.R., *Gödel, Escher, Bach : les brins d'une guirlande éternelle*, Paris, Dunod, 2008 (édition originale, New York, 1979).
- HOUDÉ O., *Les 100 mots de la psychologie*, Paris, PUF, 2018.

1

La machine raisonnante

Lorsque l'on parle d'« outil informatique », on a tendance à masquer une particularité fondamentale de l'ordinateur, qui consiste en une certaine forme d'autonomie. Il s'agit bien d'un automate, dont les actions sont déterminées par des états internes et par des règles de transition entre ces états qui peuvent prendre en compte des données issues du monde extérieur grâce à des capteurs ou au travers de systèmes de communication. De ce point de vue très superficiel, on peut déjà percevoir certains aspects anthropomorphiques de cette machine. Contrairement à un automate du XIX^e^ siècle, l'ordinateur ne répète pas inlassablement une même séquence d'actions. De plus, à la différence d'un outil, il ne produit pas nécessairement

le même effet pour une même action. Si l'on reste à ce niveau de description, le nombre de fonctionnalités que pourrait remplir l'ordinateur semble infini. Mais pour distinguer un peu mieux les contours de cet espace des possibilités, il va nous falloir entrer dans un niveau de précision un peu plus élevé. Commençons par les composants les plus élémentaires.

JEAN-PIERRE DALBÉRA/WIKIMEDIA/LICENCE CC BY

FIGURE 2 : Automates « Clowns musiciens », et « joueur de banjo » de Roullet-Descamps, 1870, en restauration au musée des arts et métiers à Saint-Denis.

Le véritable « neurone » des ordinateurs

L'une des choses les plus prodigieuses dans l'informatique est que, sur un plan conceptuel, on peut « fabriquer » un ordinateur avec seulement deux symboles distincts et un unique opérateur qui prend deux de ces symboles en entrée et produit l'un de ces symboles en sortie. Les briques de base « logiques » de l'ordinateur sont donc bien plus simples que celles d'un cerveau. Comme il s'agit de symboles, on peut évidemment les interpréter comme on le souhaite. Ainsi, si vous aimez les nombres et la conformité, vous pouvez considérer que ces symboles sont « 0 » et « 1 ». Si vous préférez la logique, vous pouvez considérer que ces symboles sont « vrai » et « faux », mais vous pouvez aussi bien les nommer « blanc »/« noir » ou « translucide »/« opaque ». En réalité, il n'est même pas nécessaire que ces symboles fassent référence à une quelconque représentation du monde. Il faut juste qu'ils soient distinguables l'un de l'autre. Ils pourraient donc être « rouge »/« orange » ou « ☲ »/« ɔ ».

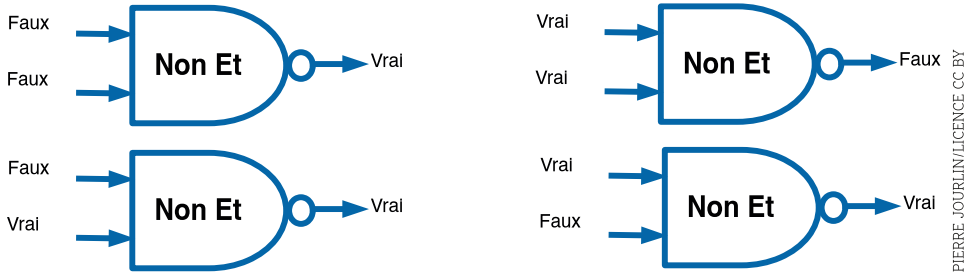


FIGURE 3 : L'opérateur logique « Non Et ».

L'opérateur (« Non Et » dans la figure 3), c'est un petit dispositif qui produit en sortie l'un de ces symboles (par exemple, « faux ») si, et seulement si, c'est l'autre symbole (par exemple, « vrai ») qui est présent simultanément sur les deux entrées. On peut l'interpréter différemment, mais si l'on reste dans le domaine de la logique, il fonctionne de la manière suivante : si, et seulement si, ces entrées ont toutes deux la valeur « vraie », alors sa sortie prend la valeur « faux ». La sortie vaut « vrai » dans les trois autres cas : faux/faux, vrai/faux et faux/vrai. Il faut noter que cette interprétation est celle qu'avait choisie l'inventeur de ces dispositifs théoriques, Georges Boole (1815-1864) – philosophe, mathématicien et logicien autodidacte du

XIX[e] siècle –, qui les avait décrits dans un ouvrage intitulé *Les lois de la pensée* :

> Le but de ce traité est d'étudier les lois fondamentales des opérations de l'esprit par lesquelles s'effectue le raisonnement ; de les exprimer dans le langage symbolique d'un calcul, puis, sur un tel fondement, d'établir la science de la logique et de constituer sa méthode ; de faire de cette méthode elle-même la base d'une méthode générale qu'on puisse appliquer à la théorie mathématique des Probabilités ; et enfin de dégager des divers éléments de vérité qui seront apparus au cours de ces enquêtes des conjectures probables concernant la nature et la constitution de l'esprit humain. [...] La connaissance des lois de l'esprit n'a pas besoin de se fonder sur un vaste ensemble d'observations. La vérité générale y est aperçue dans l'exemple particulier, et ce n'est pas la répétition des exemples qui la confirme.
>
> (BOOLE, 1992, p. 1-4).

Cependant, il n'est pas obligatoire d'opter pour une interprétation logique de cet opérateur. Par exemple, en utilisant les nombres binaires, on pourra soumettre au même opérateur

deux nombres de valeur 0 ou 1 en entrée, et il produira 1 sur sa sortie si, et seulement si, la somme des deux entrées est 0. Il est évident qu'il existe bien d'autres opérateurs qui prennent deux nombres binaires en entrée, effectuent un calcul et produisent un troisième nombre binaire en sortie. Il est assez facile par ailleurs d'énumérer toutes les possibilités : il existe 16 opérateurs distincts, parmi lesquels le « non », le « et », le « ou », le « non ou », le « ou exclusif », le « égal », le « différent », etc. Cependant, on peut aussi montrer facilement que tous ces opérateurs peuvent être construits par des combinaisons de ce seul opérateur « Non Et ».

Comme rien n'empêche de relier les sorties de ces opérateurs aux entrées d'opérateurs de même type, il devient possible de créer des petits réseaux d'opérateurs, à sens unique, qui peuvent réaliser des opérations plus complexes, avec de plus nombreuses entrées et sorties. On peut aussi, évidemment, créer des réseaux de réseaux, etc. Mais surtout, on peut montrer que tout énoncé de logique propositionnelle (ces déclarations reliées par des « non », des « et », des « ou », des « si », des « alors », etc.) peut se représenter sous la forme de l'un de ces réseaux.

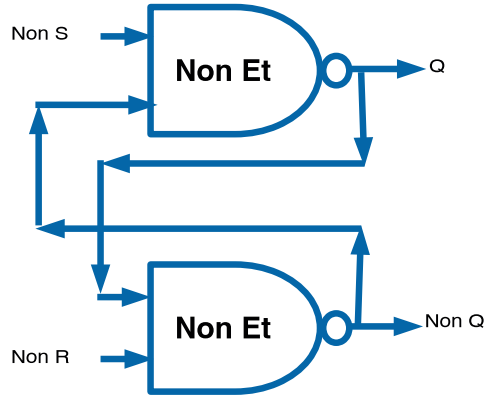
De plus, si l'on s'autorise à faire des boucles dans ce réseau de « Non Et », en reliant directement ou indirectement des sorties d'un opérateur à ses propres entrées, il devient même possible de mémoriser ces valeurs binaires et de construire un réseau dont la sortie ne dépend plus seulement des entrées, mais aussi de son état interne. C'est le principe des bascules logiques qui permettent de concevoir la mémoire des ordinateurs.

Avec quelques centaines d'heures de cours d'informatique – et ce, sans exagérer –, vous serez en mesure de concevoir dans ses moindres détails la logique d'un ordinateur complet, avec son unité centrale, sa mémoire et ses entrées-sorties. Vous pourrez ensuite confier vos plans à un fabricant qui transformera le tout en un circuit électronique d'une taille bien inférieure à celle de notre cerveau. Cependant, si ce circuit possède suffisamment de mémoire, de puissance de calcul et de bons logiciels, il sera capable – pour ne donner qu'un exemple – de battre à plates coutures tous les champions du monde de dames, d'échecs et de go. Ce n'est pas de la magie, même si l'on peut y voir une forme de *tour de magie*. En revanche, la révélation de tous les *trucs* prendrait beaucoup de temps et je ne pourrais en donner ici qu'un petit aperçu.

Les organes du microprocesseur

La structure complexe des opérateurs élémentaires – issue de plusieurs décennies de recherche et d'ingénierie humaine – possède, sans surprise, quelques aspects anthropomorphiques. Je vais décrire ici ce qui me paraît essentiel pour la compréhension de sa capacité intrinsèque à imiter certaines de nos fonctions cognitives.

La mémoire de la machine est répartie en un grand nombre de cases mémoires, qui contiennent des séquences de symboles. Comme nous l'avons vu plus haut, il n'est pas obligatoire de les interpréter comme des nombres, ou même de les interpréter tout court. Néanmoins, des sous-circuits de l'ordinateur peuvent se charger des interprétations : c'est le cas des fameuses cartes dites « d'entrée-sortie » qui transforment les données du monde réel en séquences de symboles (numérisation) et les séquences de symboles en données du monde réel (synthèse). Mais c'est aussi le cas de sous-circuits internes, capables de décoder des instructions adressées au processeur lui-même et de les exécuter en envoyant des suites de signaux à d'autres sous-circuits, qui réaliseront des opérations arithmétiques, des comparaisons, des décisions, des programmes et des sous-programmes.



PIERRE JOURLIN/LICENCE CC BY

FIGURE 4 : Le verrou "RS" est une mémoire élémentaire. Lorsque S passe à 1 (Set) : la sortie Q passe à 1. Lorsque R (Reset) passe à 1 : la sortie Q passe à 0. Lorsque R et S sont à 0 : la sortie Q maintient sa valeur précédente.

Or, puisque toute séquence de symboles contenue dans une case mémoire peut avoir une infinité d'interprétations, rien n'empêche que l'une de ces interprétations soit une adresse, c'est-à-dire une référence à une autre case mémoire. Et il se trouve qu'il existe des sous-circuits spécialement conçus pour cela. Chaque case mémoire est située à une adresse, qui est elle-même une suite de symboles. Cette dernière, si elle est assez spacieuse pour cela, peut donc parfaitement contenir l'adresse d'une autre case mémoire, voire sa propre adresse, de la même façon qu'un bureau peut contenir une grande quantité de cartes de visite, y compris

celles qui se réfèrent à son occupant. Les étudiants en informatique éprouvent souvent des difficultés à manipuler ce concept de référence. Dans le langage C, par exemple, une variable « p » est un symbole qui se réfère au contenu d'une case mémoire. Si on la fait précéder par le symbole « * », on fait référence non pas au contenu de la case mémoire elle-même, mais au contenu de la case mémoire dont l'adresse est contenue dans « p ». Il est ainsi possible d'utiliser les cases mémoires comme des étapes dans un jeu de piste et de voyager dans le réseau en enchaînant les « * ». Ainsi, « ***p » permet de prendre une adresse dans « p », de sauter à la case mémoire indiquée, d'y prendre une adresse, de sauter à la case indiquée, d'y prendre une adresse, etc.

Il découle de ces considérations que, même lorsqu'on l'observe avec l'un des plus grands niveaux de précision – à la loupe –, la machine dispose de tous les moyens logiques et électroniques nécessaires pour manipuler des symboles et leurs références, pour les interpréter, pour passer du symbole à sa référence et d'une représentation au symbole qui lui correspond. Peu d'animaux sont dotés de cette capacité et on oserait presque dire de l'ordinateur qu'il ne lui manque que la parole…

Les langages de la machine

Les processeurs sont composés de circuits électroniques complexes, très structurés et emboîtés les uns dans les autres et qui communiquent entre eux. Certains ont pour fonction de mémoriser des données, d'autres de gérer la mémoire, de calculer, d'aiguiller les différents signaux électroniques sur les différentes routes, etc. Leur langage primitif est constitué de séquences d'instructions élémentaires, essentiellement logiques et arithmétiques, inscrites dans la mémoire et interprétées par le processeur comme des séries de signaux à envoyer à différents circuits.

Les circuits et leurs sous-circuits sont associés à des fonctionnalités, un peu comme les mots d'un langage peuvent être associés à des concepts ou à des objets reliés les uns aux autres par une structure grammaticale et sémantique. Le choix des circuits et l'ordre dans lequel ils sont activés constituent donc des sortes de récits (programmes) très primitifs et très mathématiques qui décrivent des opérations de plus en plus complexes.

Il est toutefois possible d'écrire en langage machine un programme capable de réaliser des traductions en langage

machine de langages formels plus évolués. Ces logiciels particuliers sont appelés compilateurs ou interpréteurs. Un langage formel est un langage non ambigu, avec une syntaxe entièrement spécifiée, simple et figée, qui ressemble un peu aux grammaires des langages naturels, mais dans laquelle la forme et le fond se confondent. Un programme ou logiciel est une phrase de ce langage informatique, dans laquelle tout est décrit avec précision, que ce soit au niveau le plus général ou au niveau le plus spécifique, et qui ne peut être interprétée par la machine que d'une seule façon.

Cependant, si la syntaxe est figée, le vocabulaire ne l'est pas. Il est donc possible avec un langage formel de définir de nouveaux « mots » (structures de données, fonctionnalités) du même langage. Ces nouveaux « mots » peuvent alors être utilisés pour définir des structures ou des fonctionnalités qui deviennent à leur tour de nouveaux mots. Il est même facile de définir certains « mots » à partir de leur propre définition. Nous appelons cela des définitions ***récursives**** ou ***réentrantes****. Dans d'autres disciplines, on parlera plutôt de mise en abyme, d'autoréférence. Par exemple, dans le domaine de la bureautique, il est possible de définir un

« répertoire » comme un objet qui peut contenir des fichiers et des « répertoires ». Dans un autre domaine, on pourrait définir une poupée russe comme étant une poupée qui peut contenir une autre poupée russe.

Les langages informatiques modernes, dits de « modélisation objet », vont un peu plus loin en facilitant la catégorisation. Par exemple, il est facile d'exprimer le fait qu'un véhicule possède certaines caractéristiques, contraintes et fonctionnalités qui seront automatiquement héritées par tous les types de véhicules. Il est possible alors de définir une automobile comme un type particulier de véhicule, qui a toutes les caractéristiques des véhicules, avec quelques contraintes (spécificités), caractéristiques et fonctionnalités supplémentaires. Enfin, il est possible de définir les cas particuliers des catégories : les instances. Certains langages permettent aussi de définir des inconnues et de procéder automatiquement à la résolution d'équations et d'inéquations. Dans le cas des équations, il s'agit de trouver les valeurs des inconnues qui vérifient une égalité ; dans le cas des inéquations, il s'agit de trouver les intervalles dans lesquels les inconnues vérifient une relation d'ordre (supérieur, inférieur, supérieur ou

égal, inférieur ou égal, différent). Évidemment, il faut garder à l'esprit le fait que, du composant électronique jusqu'au logiciel expert – et contrairement au cerveau –, tout est issu de l'ingénierie humaine. La mécanique mise en œuvre pour résoudre tel ou tel problème peut être décrite et comprise en totalité et jusque dans ses moindres détails.

Tous ces dispositifs, conçus en moins d'un siècle, permettent aujourd'hui à l'être humain d'automatiser une prodigieuse quantité de problèmes liés à l'information que, dans la nature, il était jusqu'alors le seul à pouvoir traiter. Avec l'informatique avancée, on traite même des problèmes qui dépassent nos capacités cognitives. En effet, l'ordinateur peut explorer, bien plus vite que nous, une plus grande quantité de données ou de solutions potentielles à un problème complexe. De plus, une part toujours croissante du processus d'automatisation est elle-même automatisée.

Cela étant, dans la majorité des problématiques d'informatisation, la conception humaine des schémas de données, des contraintes et des procédures de transformation reste encore dominante. Une excellente introduction à ces

approches sophistiquées, regroupées sous la formule « ingénierie de la connaissance » pourra être trouvée dans l'ouvrage de Pierre Lévy intitulé *De la programmation considérée comme un des beaux-arts* (Lévy, 1992).

En effet, il est encore très difficile d'automatiser les phases les plus fondamentales du processus de cognition : analyser une problématique, en faire une représentation mentale suffisamment pertinente, claire et précise ; traduire, enfin, cette représentation mentale dans un langage formel. Certaines équipes de recherche travaillent aujourd'hui précisément sur cette problématique d'automatisation de la programmation informatique (Ye, 2020). Toutefois, cela reste encore, et probablement pour longtemps, une limite forte de l'autonomie de ces machines ; j'en expliquerai les raisons dans les chapitres suivants.

Pour en savoir plus

- DOWEK G, LÉVY J.-J., *Introduction à la théorie des langages de programmation*, Palaiseau, Éditions de l'École polytechnique, 2006.
- LÉVY P., *De la programmation considérée comme un des beaux-arts*, Paris, La Découverte, 1992.
- TIMSIT C., *Du transistor à l'ordinateur*, Paris, Hermann, 2010.

2

La machine apprenante

Nous avons vu précédemment que les données (entrées, sorties, algorithmes) peuvent être interprétées de façon numérique. Nous pourrions donc imaginer une méthode simple, rendue opérationnelle sous la forme d'un logiciel, qui permettrait de trouver automatiquement les programmes qui fonctionnent, c'est-à-dire qui délivrent des sorties correctes pour toutes les entrées possibles.

Pour cela, il faudrait bien entendu disposer très concrètement de toutes les caractéristiques du problème traité, de la liste des « inconnues », ainsi que d'une procédure permettant de valider ou d'invalider les solutions correspondantes. Il ne resterait alors plus qu'à énumérer tous les programmes possibles, par longueur

croissante, jusqu'à trouver ceux d'entre eux dont les sorties sont valides.

Cependant, même en supposant que les possibilités soient finies, on est très vite confronté à un très grand nombre de combinaisons. Supposons par exemple – comme c'est le cas pour certains types de microprocesseurs très utilisés dans les smartphones – que chaque instruction soit représentée par un nombre de 32 chiffres binaires. Pour la démonstration, rien ne nous oblige d'ailleurs à compter en binaire : disons qu'un nombre binaire de 32 chiffres correspond à un nombre de 10 chiffres de la base 10 que nous avons plus l'habitude de manipuler et qui peut donc légèrement dépasser les 4 milliards. On peut en déduire qu'il existe plusieurs milliards d'instructions possibles, plusieurs milliards de milliards de programmes distincts contenant deux instructions, etc. Si l'on a bien compris que chaque instruction supplémentaire multiplie par 4 milliards le nombre de possibilités et si l'on sait qu'une simple commande permettant d'afficher un bout de texte sur un écran compte quelques dizaines de milliers d'instructions en langage machine, on réalise très vite que le nombre de possibilités se compose de centaines de milliers de chiffres. À

titre de comparaison, il est admis que le nombre d'atomes qui composent l'univers est quant à lui composé d'environ 80 chiffres (Pfenniger, cité par Gracci, 2015).

Cela signifie que, même dans les cas où il est possible de lister toutes les entrées et les sorties attendues d'un programme, il est simplement impossible de laisser l'ordinateur explorer tous les algorithmes potentiels. D'autant que cette démarche peut se heurter à des limitations théoriquement indépassables, notamment parce que les algorithmes qui explorent les algorithmes sont eux-mêmes des algorithmes et qu'ils devraient par conséquent être capables de s'explorer eux-mêmes.

Ces considérations théoriques sont regroupées dans un sous-domaine des mathématiques et de l'informatique que l'on nomme « calculabilité et complexité des algorithmes ». Les grands pionniers de cette discipline sont Alonzo Church (1903-1995), Kurt Gödel (1906-1978) et Alan M. Turing (1912-1954).

Cependant, à l'intérieur de ces limites théoriques, on pourra laisser l'ordinateur tester une grande partie des possibilités. Il faudra toutefois qu'un être humain définisse au préalable les frontières de l'exploration. Cela peut passer par des choix de

représentations ou d'hypothèses, des simplifications ou encore des ***heuristiques***[*], autrement dit des raccourcis guidés par des considérations empiriques.

Essayons maintenant de décrire les approches qui peuvent être suivies dans les situations où les données sont disponibles, bien que les experts humains ne soient en mesure de formaliser ni leurs connaissances ni leurs procédures de décision.

La catégorisation automatique

La catégorisation automatique est l'une des délimitations algorithmiques étudiées dans le domaine de la recherche en intelligence artificielle. Nous avons vu que la catégorisation était un principe fondamental de notre architecture cognitive et que les ordinateurs étaient capables de catégoriser et de manipuler les catégories. Mais cela ne fonctionne que lorsqu'une description précise, formelle, des catégories, des sous-catégories et finalement des cas qu'elles représentent est disponible.

En outre, notre cerveau est bien moins exigeant : il lui suffit de quelques exemples pour élaborer des hypothèses générales sur la définition d'une catégorie. Par conséquent, si dans certaines problématiques bien circonscrites, il suffit d'indexer tous les cas de la catégorie, cette liste est très souvent théoriquement ou pratiquement infinie, voire inaccessible. Par exemple, un logiciel qui aurait comme fonction de sélectionner des photos d'une espèce particulière d'animal et qui fonctionnerait par indexation des photos devrait contenir toutes les photos des individus de l'espèce, vivants, morts, à naître, dans toutes les positions, dans tous les environnements, sous tous les angles, dans toutes les conditions de luminosité. Or, si je considère une image de la taille de mon écran d'ordinateur portable, je sais qu'elle correspond à une matrice de 1366 points en largeur et 768 points en hauteur et que la couleur et l'intensité de chaque point sont codées dans un nombre de 32 bits. Cela fait donc 4 milliards de possibilités pour chacun des 4 milliards de points de l'image. Le nombre potentiel d'images distinctes contient quant à lui plus de 37 milliards de chiffres.

L'objectif de cette tâche dite de catégorisation est donc de transformer des symboles qui prennent leurs valeurs dans de très grands espaces – par exemple, des images – en d'autres symboles qui prennent leurs valeurs dans des espaces plus réduits – par exemple, le nom des objets qu'elles contiennent. Or, si cette tâche paraît facile à réaliser pour un cerveau humain, c'est sans aucun doute parce qu'elle échappe presque entièrement à sa conscience. Les approches s'appuyant sur l'introspection vont donc se heurter à de grandes difficultés.

Un cas concret

En 1994, dans le cadre d'un master (qui s'appelait alors un DEA) en informatique et mathématiques, j'étudiais toutes sortes de techniques dites de « traitement de signal ». En traitement automatique de la parole, il s'agissait de transformer tout d'abord les variations d'amplitude sonore en variations de fréquence (la « tonalité »), puis en « formants » (des bandes de fréquences significatives). Mon collègue de bureau, qui avait commencé sa thèse huit ans avant moi, avait créé un logiciel capable de catégoriser les sons en formalisant les connaissances fournies par la phonétique, la phonologie et la linguistique pour transformer successivement ces « formants » en « traits acoustiques », puis en « phonèmes ». Cette étape était extrêmement complexe : les phonèmes n'ont pas la même forme en fonction des phonèmes qui précèdent et qui suivent (coarticulation). Ils dépendent aussi de l'identité du locuteur (variabilité interlocuteur) et, pour un même locuteur, ils peuvent varier en fonction de son état (variabilité intralocuteur), ou même en fonction de la situation extérieure (hypo ou hyperarticulation). L'étape suivante – que nous réalisons également inconsciemment – est, elle aussi, loin d'être simple, puisqu'il est courant qu'une même suite de phonèmes puisse représenter différents mots ou séquences de mots (homophonie).

L'approche consistait à récolter le fruit de l'intelligence humaine, à le formaliser dans un langage mathématique, à le préciser dans un langage informatique, pour finalement le transférer à la machine. Bien entendu, nous étions conscients de transférer le fruit (la représentation mentale), sans l'arbre (le système cognitif) et encore moins ses racines (l'évolution

génétique). Cette approche est loin d'avoir disparu. Si on la prend dans un sens large, elle est encore dominante avec les bases de données, les modèles relationnels, les réseaux et la modélisation objet. Même avec une définition plus restrictive, elle est encore présente dans l'ingénierie de la connaissance (web sémantique, bases de données), ainsi que dans quelques domaines de recherche très actifs (optimisation, théorie des graphes, etc.).

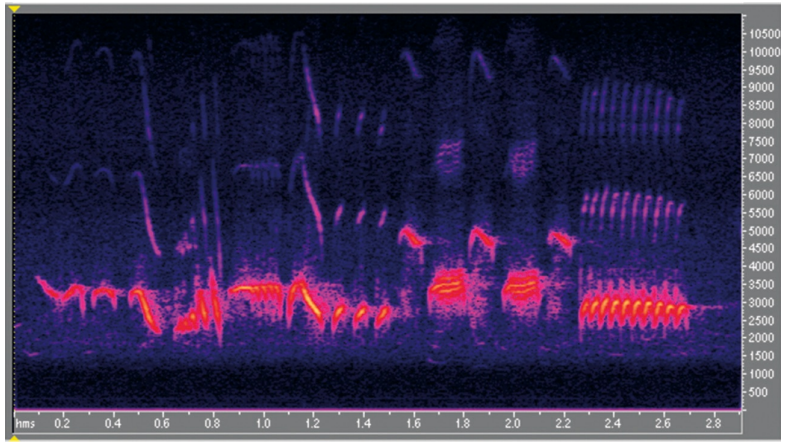
Néanmoins, en reconnaissance automatique de la parole notamment, l'effort de recherche s'est très vite heurté à des difficultés majeures. Les experts en phonétique pouvaient facilement lire des sonogrammes (voir figure 5), mais ils étaient incapables d'expliquer, avec la précision requise pour une transcription informatique, comment ils procédaient mentalement. De manière plus générale, nous réalisons que nos capacités de métacognition ne nous permettent d'accéder qu'aux niveaux les plus élevés de notre architecture cognitive. En revanche, nos productions intellectuelles sont, elles, issues du fonctionnement conjoint de tous les niveaux.

Une idée générale pour contourner ce problème pourrait consister en une captation, dans un grand nombre de productions

intellectuelles des individus, des textes, de la parole, des ensembles de données classifiés manuellement, des mécanismes cachés qui ont mené à la production de ces données.

L'idée de rendre l'ordinateur capable de construire lui-même sa représentation des catégories sur la base d'une grande quantité d'exemples a donc refait surface quelques décennies après qu'Alan M. Turing l'avait évoquée en 1950 dans « Computing machinery and intelligence ». Dans cet article, Turing imaginait déjà un ordinateur présent dans une salle de classe et, s'inquiétant des moqueries que sa singularité ne manquerait pas de provoquer chez ses camarades humains, il suggérait de recourir exclusivement aux cours particuliers.

D'autre part, la catégorisation automatique de données avait déjà été largement étudiée en mathématiques, et en particulier dans le domaine des probabilités et des statistiques. Aussi, dans les années 1990, les plus performants des systèmes de reconnaissance automatique de la parole étaient souvent fondés sur des techniques statistiques évoluées : les modèles de Markov cachés ou les champs conditionnels de Markov, pour ne citer que les plus populaires d'entre eux.



WIKIMEDIA/DOMAINE PUBLIC

FIGURE 5 : Sonagramme d'un chant de merle bleu réalisé par logiciel (copie d'écran). L'axe horizontal représente le temps. L'axe vertical représente la fréquence, du plus grave en bas au plus aigu en haut. L'intensité des couleurs représente le volume du son.

Le principe de base de ces approches consiste à estimer la probabilité qu'un son ayant une certaine intensité sur différentes fréquences soit associé à un symbole particulier. Cependant, tout comme les constituants élémentaires de la parole, les phonèmes ne sont pas des sons stables ; par conséquent, il faut aussi prendre en compte la probabilité qu'une situation acoustique stable transite vers une autre situation.

La phase dite d'apprentissage consiste à estimer les probabilités des états (stables) et des transitions entre états à partir d'une grande quantité d'exemples de phonèmes représentant une grande diversité de contextes d'énonciation et de locuteurs. Ceci permet d'avoir pour chaque phonème un modèle statistique associé, appelé modèle de Markov caché. Dans la phase dite de reconnaissance, les algorithmes ont comme objectif de trouver le modèle le plus probable compte tenu du son effectivement numérisé et présent en entrée du système.

« Intelligence » sans catégorisation ?

Paradoxalement, dans d'autres applications de l'intelligence artificielle, il était parfois impossible de décrire les processus mentaux à l'œuvre chez l'être humain et pourtant possible de construire des machines opérationnelles sur la base d'hypothèses probabilistes très simples, sans vraiment procéder à une quelconque catégorisation.

Par exemple, on aurait pu fixer comme objectif aux moteurs de recherche documentaire, très utilisés aujourd'hui par le grand public, de catégoriser les documents en deux classes : les documents pertinents vis-à-vis de la requête et ceux qui ne sont

pas pertinents. Néanmoins, il est plutôt d'usage de présenter les documents dans l'ordre décroissant de leur probabilité de pertinence. L'effet produit à partir de formules relativement simples mettant en jeu la fréquence des mots d'une requête dans un index peut être vraiment déconcertant. Comment une machine peut-elle trouver en quelques fractions de seconde les documents que je recherche, à partir d'une question que j'ai formulée en langage naturel et dans une base documentaire qui contient des milliards de pages ?

Dans « Je me souviens de Nicolas Ruwet et de Maurice Gross », Jean-Paul Dècle décrit une scène dans un colloque scientifique qui exprime bien l'effet déconcertant de ces méthodes :

> Ce thésard protestait avec véhémence contre la présentation d'un moteur de recherche. Il mettait en doute, compte tenu de l'état de la connaissance dans le domaine, la possibilité de faire une recherche en texte libre, et criait à la supercherie parce qu'on osait présenter cette réalisation alors que les linguistes n'avaient pas encore réussi à se mettre d'accord sur la définition de ce qu'était un mot.
>
> (Dècle, 2006, p. 6)

En effet, pour décider de la pertinence d'un document vis-à-vis de l'expression d'un besoin en information, il est nécessaire à l'être humain de comprendre la requête et de comprendre le document. Ce ne sont pas les mots de la requête et du document qui sont comparés, mais les concepts, objets et relations auxquels ces mots font référence. Le doctorant cité ci-dessus ne s'était donc pas imaginé qu'il était possible de passer outre cette compréhension du rôle d'un mot et il refusait de croire qu'une approche purement statistique pouvait permettre la conception d'un système opérationnel. C'était manifestement une erreur.

Certes, le résultat n'est pas toujours parfait et certaines requêtes retournent beaucoup de documents non pertinents (les « faux positifs »). De plus, et même si c'est plus difficile à évaluer, ces requêtes laissent dans l'ombre un bon nombre de documents pertinents (les « faux négatifs »). Très probablement, un être humain, s'il avait eu le temps de lire ces milliards de pages, aurait fourni un résultat de bien meilleure qualité. Mais la machine a pourtant atteint dans ce domaine le but ultime que l'on peut lui fixer : nous rendre service.

Dans le même ordre d'idées, de nouvelles méthodes sont apparues très récemment : les *word embeddings* ou « plongements lexicaux ». Cette approche connexionniste permet de représenter chaque mot d'une langue sous la forme d'un vecteur de nombres rationnels. Les composantes de ce vecteur sont les paramètres d'un réseau de neurones, entraîné pour prédire l'occurrence d'un mot quand on connaît son contexte, ou d'un contexte lorsque l'on connaît le mot.

Il n'y a donc aucun besoin d'annoter le texte manuellement pour constituer des données d'apprentissage. De plus, si ces données sont suffisamment volumineuses, on observe que chaque composante du vecteur associé à un mot capture, d'une certaine façon, une partie de son sens.

Cela permet des opérations intéressantes : la soustraction des vecteurs correspondant aux mots « père » et « homme » (vecteur « parent »), à laquelle on ajoute le vecteur correspondant au mot « femme », produit un vecteur qui correspond au mot « mère ». En outre, cette représentation des mots en vecteurs est très utilisée comme entrée de systèmes connexionnistes bien plus complexes, capables par exemple de répondre à des questions exprimées en

langage naturel ou de traduire des textes d'une langue vers une autre.

Par ailleurs, lorsque l'on a réussi à transformer un mot ou une séquence de mots en grandeur numérique, il est toujours possible de procéder à une véritable catégorisation, en fixant des seuils sur ces grandeurs. Par exemple, et pour reprendre le domaine de la recherche documentaire, lorsque des documents sont présentés par probabilité décroissante de pertinence vis-à-vis d'une requête, il peut paraître judicieux de ne pas présenter à l'utilisateur les documents qui ont moins d'une chance sur deux d'être pertinents. Cependant, le choix du seuil dépend vraiment de l'intention de l'utilisateur.

Afin de pouvoir comparer la qualité de plusieurs systèmes de catégorisation, ce seuil est généralement fixé de façon à obtenir des taux égaux de fausse acceptation et de faux rejet, mais il peut aussi être ajusté pour obtenir des résultats favorisant soit le rappel, soit la précision.

Là encore, les résultats obtenus peuvent donner l'impression que le logiciel comprend le sens des mots, et ce de façon assez bluffante. Mais l'illusion disparaît dès lors que l'on analyse en

détail les erreurs de classification. En effet, si ces approches capturent très bien les usages des mots et qu'elles généralisent souvent correctement ces usages grâce aux répétitions dans les données d'apprentissage, elles échouent dès qu'il s'agit d'extraire la logique du discours et les relations entre les concepts et objets signifiés.

L'essor des modèles prédictifs

Les applications liées au traitement automatique du langage naturel utilisent ces approches dites « par apprentissage ». Mais parfois, les problématiques peuvent être considérablement plus complexes et les perspectives de résolution, vraiment plus lointaines.

Par exemple, lorsque nous reconnaissons un signal de parole, notre reconnaissance ne prend pas seulement en compte le signal lui-même. Nous nous aidons de nos connaissances morphologiques, lexicales, syntaxiques, sémantiques et même pragmatiques pour parvenir à décoder le signal acoustique ou

textuel. Nous savons, par expérience, qu'une pancarte au fond d'une salle de cinéma faiblement éclairée a de bonnes chances de porter la mention « défense de fumer » et nous avons donc l'impression de lire clairement un texte qui, sans la prise en compte du contexte, serait parfaitement indéchiffrable.

Le nombre de données d'apprentissage nécessaires pour modéliser automatiquement et correctement toutes ces catégorisations et tous ces contextes est tel que seuls des systèmes informatiques très puissants disposant de corpus linguistiques massifs sont capables de montrer une quelconque utilité pratique. Cependant, les systèmes dits de « traitement automatique du langage » les plus performants sont encore, à l'heure actuelle, très loin de pouvoir démontrer une réelle compréhension du langage naturel.

De plus, l'apprentissage automatique n'est pas limité à l'imitation des capacités perceptives de la cognition humaine. Il ne s'appuie pas non plus forcément sur des approches de type statistique. D'ailleurs, avant les années 1990, il existait toute une activité de recherche dans ce que l'on appelait l'apprentissage symbolique. Il s'agissait de trouver des algorithmes permettant à un ordinateur

d'apprendre « de lui-même » des transformations de suites de symboles en d'autres suites de symboles à partir de considérations relevant de la logique, ou plutôt de différentes formes de logique.

En parallèle, un autre axe de recherche explorait des approches dites connexionnistes, fondées sur une imitation du réseau de neurones et de synapses qui compose les cerveaux humains et non humains. Ces méthodes existaient également depuis les années 1960, avec le fameux « perceptron » : un réseau de neurones artificiels disposant d'une couche unique et capable de réaliser une régression linéaire à plusieurs dimensions. L'approche était bien moins économe en mémoire et en puissance de calcul, et les systèmes ainsi conçus concurrençaient difficilement ceux qui reposaient sur des formalismes logiques.

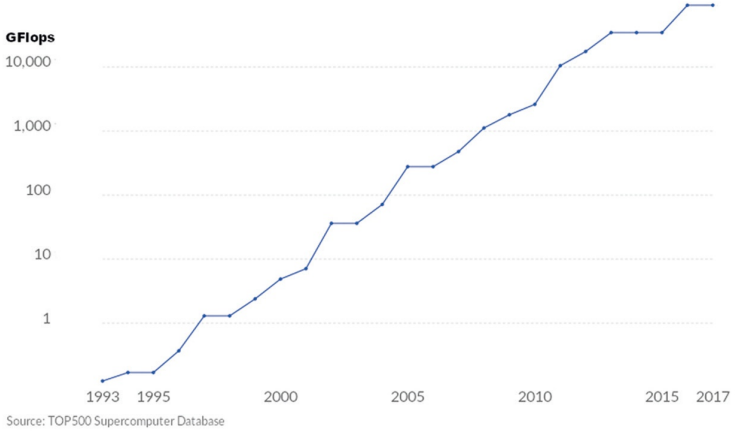
Ces trois dernières décennies ont vu les capacités de mémorisation et de calcul exploser grâce aux progrès de la miniaturisation des composants électroniques. L'arrivée d'Internet a, quant à elle, mis à la disposition des machines des quantités impressionnantes de données brutes (sons, images, vidéos), que l'on désigne souvent sous l'expression *Big Data*. Ces circonstances étaient nettement plus favorables aux approches connexionnistes.

En parallèle de l'approche connexionniste se développait une approche dite stochastique, qui consistait à « nourrir » des estimations statistiques avec de nombreux exemples de chaque phonème, syllabe ou mot, à calculer des probabilités conditionnelles qu'un mot apparaisse avant ou après un autre. Elle permettait d'obtenir rapidement des résultats prometteurs. On faisait une hypothèse de distribution statistique et on calculait ses paramètres, typiquement des moyennes, des variances, des probabilités de transition entre différents états, sur la base des exemples dits d'apprentissage.

Dans toutes ces approches fondées sur un apprentissage de type statistique, ce sont les réseaux de neurones artificiels qui ont fini par s'imposer dans la plupart des applications, de la reconnaissance d'image à la traduction automatique.

Ainsi, l'explosion conjointe de la quantité de données disponibles et de la puissance de calcul des ordinateurs a accéléré considérablement le développement de la classification dite « non supervisée », dans laquelle il s'agit de regrouper des observations en « classes », sans toutefois être tenu de leur affecter des symboles. Pour cela, il existe un nombre considérable de

méthodes, s'appuyant sur une grande variété de critères. Ce sont souvent des distances numériques, plus ou moins sophistiquées, qui sont calculées entre les observations.



HTTPS://OURWORLDINDATA.ORG /LICENCE CC BY 4.0

FIGURE 6 : Augmentation de la puissance des superordinateurs, mesurée en milliards d'opérations par seconde sur des nombres à virgule (GFlops). La puissance indiquée correspond à celle du plus puissant superordinateur de l'année correspondante.

Ces méthodes peuvent être très utiles pour traiter certains sujets. En visualisation de données textuelles, elles permettent par exemple de faire apparaître les mots clefs les plus significatifs d'une collection de documents textuels : ce sont les « nuages de termes » que l'on voit parfois sur les sites web.

Elles permettent aussi de réaliser des tâches plus générales, comme associer des termes apparentés, des questions avec des réponses, etc.

Dans certaines situations, cependant, il peut être plus difficile de travailler avec des données brutes, sans annotation humaine. C'est le cas par exemple en reconnaissance de la parole. Pour constituer un corpus d'apprentissage, il nous faut pouvoir mettre en correspondance les sons avec leurs symboles (phonèmes, syllabes, mots, etc.). Or, ces données ne sont pas produites massivement par l'industrie. On peut par exemple travailler avec les sous-titres des films ou des vidéos, mais il ne s'agit pas d'une transcription littérale de l'audio et les deux types de données, audio et texte, sont faiblement synchronisés.

Cela peut être encore plus difficile lorsqu'il s'agit d'associer automatiquement des significations à des discours ou des parties de discours. Nous ne disposons pas, par exemple, de données massives associant des portions de textes à des catégorisations qui permettraient de limiter la diffusion de fausses informations et de discours de haine. Par ailleurs, on sait bien en statistique

que les faiblesses de représentativité des données peuvent poser de graves problèmes de biais. À l'issue d'une chaîne de traitements, lorsqu'elles sont intégrées dans des modèles, puis dans des applications, ces données biaisées peuvent entraîner des décisions discriminatoires. On en trouvera quelques exemples dans l'ouvrage de Cathy O'Neil intitulé *Algorithmes : la bombe à retardement* (O'Neil, 2018).

Pour ce type d'application de l'intelligence artificielle, il y a encore une place pour les approches dites « symboliques ». Ces approches, fondées sur des lexiques, des règles de grammaire, des théories linguistiques, offrent l'avantage d'être plus facilement accessibles à l'esprit humain que les milliards de paramètres que manipulent les réseaux de neurones artificiels et les modélisations statistiques.

Les limites de la prédiction

La classification automatique dite « supervisée » présente aussi de très nombreuses difficultés. Dans cette approche, il

est possible de connaître *a priori* ou *a posteriori* la catégorie à laquelle appartient chaque observation. L'évaluation est alors plus directe : il suffit de compter, sur un ensemble de tests, le nombre de faux positifs et de faux négatifs produits par la catégorisation automatique. Le problème peut donc être vu sous l'angle de l'inférence ou de la prédiction. D'ailleurs, c'est un peu ce que nous faisons quand nous manquons de théorie pour interpréter de nouveaux faits.

NATHALIE RUAUX/WIKIMEDIA/LICENCE CC BY SA

FIGURE 7 : « Ce qui donne le cancer », d'après la classification du Centre international de recherche sur le cancer, nuage de mots-clés.

EMERY/WIKIMEDIA/LICENCE CC BY SA

FIGURE 8 : Cette image est associée au mot « wxcv » dans une langue inconnue. D'après vous, que signifie le mot « wxcv » ?

Par exemple, à la question posée dans la légende de la figure 8, il y a de bonnes chances que l'on réponde « chat », car dans l'apprentissage des langues, nous avons tendance à parier sur des niveaux de généralisation moyens. Cependant, « wxcv » pourrait tout aussi bien signifier « siamois », « félin », « mammifère », « animal », « tête », « yeux », « moustache », « poils », « deux », « strabisme », etc.

La prise en compte d'un plus grand nombre et d'une plus grande variété d'images associées à « wxcv » permettra d'affiner l'hypothèse et de gagner en certitude, mais sans pour autant atteindre le niveau d'une preuve.

Les systèmes d'apprentissage automatique n'échappent pas à cette règle. Comme nous, ils peuvent surgénéraliser ou, au contraire, spécifier abusivement. Dans le premier cas, la catégorie sera truffée de faux positifs ; dans le second, elle sera remplie de faux négatifs. Cependant, il se peut aussi que l'hypothèse de départ soit incorrecte ; dans ce cas, aucune nouvelle observation ne pourra l'affiner : il faudra la remettre en cause.

D'ailleurs, lorsque l'on fait une hypothèse – même la plus vague –, ne suppose-t-on pas que le phénomène que l'on observe est prédictible ? Cette croyance au déterminisme des lois naturelles, qui est de nos jours encore très ancrée dans les esprits, a été définie par Pierre-Simon de Laplace en 1825 dès l'ouverture de son livre *Essai philosophique sur les probabilités* :

> Nous devons envisager l'état de l'Univers comme l'effet de son état antérieur et la cause de ce qui va suivre. Une intelligence qui

> pour un instant donné connaîtrait toutes les forces dont la nature est animée et la situation respective des êtres qui la composent, si d'ailleurs elle était assez vaste pour soumettre ces données à l'analyse, embrasserait dans la même formule le mouvement des plus grands corps de l'Univers et ceux du plus léger atome : rien ne serait incertain pour elle, l'avenir comme le passé serait présent à ses yeux.
>
> (LAPLACE, 1825, p. 3-4)

Néanmoins, on sait aujourd'hui que certains systèmes sont tellement sensibles à certains de leurs paramètres que d'infimes variations d'une cause peuvent provoquer de grandes conséquences. Cette notion de « chaos déterministe » – plus connue sous le nom d'***effet papillon**** – a été popularisée par le météorologue Edward Lorenz (1917-2008) lors d'une conférence donnée en 1972 et intitulée : « Prédictibilité : le battement d'ailes d'un papillon au Brésil peut-il provoquer une tornade au Texas ? »

L'induction consiste à prédire des phénomènes que l'on est incapable de comprendre. Or, il faut aussi noter, même si c'est

moins intuitif, qu'il est également possible de comprendre des phénomènes que l'on est incapable d'anticiper. Plus généralement, le hasard est une notion très difficile à appréhender, mais il existe d'excellents ouvrages de vulgarisation – par exemple, *Comprendre sans prévoir, prévoir sans comprendre* (Krivine, 2016) – pour en saisir les subtilités.

Une autre faille des systèmes de classification automatique est due à la classique confusion que nous pouvons faire entre ***corrélation**** et causalité. Lorsque deux mesures évoluent conjointement, on dit qu'elles sont corrélées. Or, s'il est parfaitement possible que cette corrélation soit le reflet d'une relation de causalité, il faut encore savoir quelle mesure est la conséquence de l'autre. Par ailleurs, elles peuvent être toutes deux la conséquence d'une même cause extérieure, sans être la cause l'une de l'autre. Et enfin, il est possible que la corrélation soit simplement une pure coïncidence.

Sur son site, Tyler Vigen, étudiant à Harward Law School, recense quelques fortes corrélations, calculées avec le plus grand sérieux, mais qui font apparaître des liens entre des phénomènes très éloignés. Par exemple, l'auteur calcule une corrélation positive

de 99,79 % entre les dépenses d'un État en faveur de la recherche scientifique et un taux de suicide, ou encore une corrélation positive de 99,26 % entre un taux de divorce et un niveau de consommation de margarine (Vigen, 2015).

Nous pouvons nous aussi tomber dans ce type de piège, mais, contrairement aux machines, nous avons la faculté de remettre en cause nos propres jugements et de modifier en conséquence nos façons de raisonner. Il s'ensuit que lorsque nous transférons notre représentation du monde dans une machine, elle hérite presque instantanément de millénaires de réflexions et de remises en cause individuelles et collectives au sein de la société humaine. En revanche, lorsque la machine apprend toute seule, elle peut traiter des quantités considérables de données en un temps infime, mais, du point de vue cognitif, elle part quasiment de zéro.

De plus, l'ordinateur dépend encore de l'être humain, soit pour lui fournir des données, soit pour lui indiquer où les chercher. Cependant, en sélectionnant directement ou non les données d'apprentissage, l'être humain peut être affecté par les quelques centaines de ***biais cognitifs**** qui ont été identifiés par la psychologie, et ainsi les transmettre aux logiciels.

Les hypothèses sur lesquelles s'appuient les réseaux de neurones artificiels sont plus difficiles à cerner que dans le cas de la statistique traditionnelle. On a pu montrer par exemple que, lorsqu'un tel réseau ne possède qu'une seule couche de neurones, il est équivalent à une régression linéaire à plusieurs dimensions, une méthode statistique connue depuis le XVIII^e^ siècle. Il s'agit d'approximer des données expérimentales par un modèle théorique dont on estime les coefficients. Lorsque les données sont des points sur une courbe, ce modèle est une droite. Quand il s'agit de modéliser des points dans un espace à plusieurs dimensions, on estimera les paramètres d'un plan multidimensionnel.

Quelques décennies plus tard, de nouvelles avancées ont permis d'estimer les paramètres d'un réseau possédant plusieurs couches de neurones, grâce à une méthode appelée « rétropropagation du gradient » (Le Cun, 2019), et on a pu montrer que ces nouveaux réseaux de neurones étaient capables de faire de la classification non linéaire. C'est-à-dire qu'ils étaient devenus capables de séparer différentes catégorisations par des formes plus complexes que les

plans multidimensionnels. Cela a permis un véritable saut qualitatif. Les performances atteintes avec ces méthodes – qualifiées d'« apprentissage profond » (ou *deep learning*) –, en reconnaissance d'image, puis en traitement automatique du langage, ont été spectaculaires et ont mis cette technologie sur le devant de la scène.

Toutefois, il faut garder à l'esprit que la structure de ces réseaux, leur type, le nombre de neurones, la mesure de distance choisie pour évaluer leur production vis-à-vis des attendus, ainsi que la fonction qui permet de minimiser cette distance constituent d'une certaine façon des hypothèses sur les données que le système doit *prédire*. En outre, si une hypothèse plus fine permet de réduire l'espace de recherche, elle entraîne un risque important d'inadéquation ; inversement, une hypothèse plus générale peut s'adapter plus facilement aux données d'apprentissage, mais elle nécessitera un plus grand volume de données et un plus grand effort de calcul. Ainsi, la recherche d'un bon compromis devient un sujet d'étude à lui seul.

L'une des hypothèses les plus fortes induites par ces systèmes d'apprentissage automatique est sous-tendue par le principe

qui régit leur adaptation aux données. Les algorithmes d'apprentissage supposent qu'une légère modification des paramètres du modèle entraîne une légère modification de son comportement. Ceci permet de s'appuyer sur un processus incrémental qui sélectionne les petites modifications du système qui se révèlent les plus avantageuses. On évite ainsi de rester bloqué dans des minima locaux en utilisant des méthodes inspirées par les procédés de recuit des métaux en fonderie.

L'idée générale pour trouver le minimum d'une fonction d'erreur est de suivre la plus grande pente. Mais si la fonction possède plusieurs « vallées », il faut parfois remonter la pente pour atteindre une nouvelle vallée possiblement plus basse. C'est ainsi qu'en fonderie, on peut chauffer et refroidir plusieurs fois un métal pour obtenir une structure qui optimise certaines propriétés mécaniques, comme la ductilité, propriété d'un matériau à se déformer sans se rompre.

Par ailleurs, cette hypothèse selon laquelle de petites causes produisent de petites conséquences autorise aussi des approches de type « évolutionniste » inspirées de la génétique. L'idée consiste à spécifier des algorithmes particuliers que l'on

peut copier imparfaitement sans trop s'éloigner de l'original. Il est dès lors possible d'imiter le processus d'évolution par sélection naturelle. Cependant, cette hypothèse de la continuité des caractéristiques et de leurs effets est une hypothèse très forte qui est souvent sous-évaluée. Par exemple, si l'on considère les langages, qu'ils soient naturels ou artificiels, comme des systèmes à forte capacité descriptive, une modification minime comme l'introduction d'une négation ne va pas bouleverser la thématique, mais elle va avoir de grandes conséquences sur la sémantique d'un énoncé.

Le cumul de ces limitations a une conséquence : les exploits apparents des machines apprenantes, en particulier dans les domaines du traitement automatique du langage naturel (recherche documentaire, agents conversationnels, traduction, etc.) résident dans leur capacité à imiter les comportements humains sans en rechercher la source. Les progrès réalisés ces dernières décennies apparaissent considérables, surtout si on les évalue à l'aide d'un test de Turing (le jeu d'imitation). Mais il me semble qu'il est encore assez facile à un interrogateur avisé, au fait des technologies mises en œuvre, de démasquer la machine.

Les limites de l'imitation

En premier lieu, toute méthode qui consiste à inférer des situations inédites sur la base de l'étude d'un ensemble de situations plus ou moins représentatives d'un phénomène suppose que ce phénomène est stable. Néanmoins, pour reprendre mon exemple favori, ce n'est pas le cas du langage naturel, car nous inventons des mots en permanence. C'est le cas avec les noms propres, mais nous jouons aussi sur des règles morphologiques pour créer de nouvelles formes à partir de racines anciennes. Nous importons aussi en permanence des mots étrangers, qui parfois retournent à leur pays d'origine avec une signification plus ou moins différente. Par exemple, l'anglicisme « challenge » qui tend à remplacer le mot français « défi » dans de très nombreux milieux francophones, dont celui des professionnels du numérique, vient à l'origine du mot *chalenge* qui, en vieux français, signifiait « contestation ».

En réalité, même si nous pouvions figer un vocabulaire, rien ne nous assurerait que les mots, dans des contextes pourtant similaires, font toujours référence aux mêmes concepts ou objets. Enfin, nous avons l'impression que face à un même stimulus, dans

un seul contexte, nous associons un symbole unique. En effet, si ce n'était pas le cas, comment pourrions-nous communiquer ? Cependant, les mécanismes cognitifs qui nous permettent de gérer le contexte sont peut-être plus sophistiqués que ce que l'on croit : dans un labyrinthe, un grand nombre de positions se ressemblent, et pourtant notre cerveau se montre aussi capable d'associer des positions différentes à des images similaires, voire à des séquences d'images identiques, en déroulant et en réenroulant mentalement le fil d'Ariane. Ainsi, les systèmes de traitement automatique de la langue naturelle (TALN) peinent toujours aujourd'hui à atteindre le même niveau de qualité que l'être humain, sauf dans quelques applications hyperspécialisées. Pire, lorsque les données dites d'apprentissage reproduisent des biais cognitifs, ces biais peuvent, à la fin du processus de décision automatique, se traduire par des discriminations socialement inacceptables (O'Neil, 2018).

Une autre limite réside dans le coût financier, énergétique et écologique de ces solutions. Le cerveau humain est constitué de centaines de milliards de neurones, synapses, cellules gliales, qui sont interconnectés et qui ne nécessitent pas plus d'énergie au

total qu'une poignée d'ampoules basse consommation (autour de 12,6 W). Il est dès lors évident que la consommation électrique et la production de chaleur d'un système équivalent, construit à partir de composants électroniques, sont considérablement plus importantes que celles de sa version biologique, aussi imparfaite soit-elle : par exemple, l'ordinateur IBM Watson – qui devint, en 2011, champion du jeu télévisé américain *Jeopardy* – consommait environ 90 000 watts répartis sur 90 serveurs, chacun doté de quatre processeurs à quatre cœurs. Et, en ce qui concerne la phase d'apprentissage, les systèmes « état de l'art » de traitement automatique du langage comme BERT ou GPT-3 nécessitent quant à eux environ 1 000 kilowattheures, soit l'équivalent de la consommation énergétique d'un cerveau humain apprenant huit heures par jour pendant trente-cinq ans. Il va sans dire qu'à l'heure actuelle, le cerveau humain reste assez compétitif sur cet aspect des choses, même si l'approche « quantique » permet d'entrevoir de nouvelles perspectives (Querlioz & Grollier, 2020).

Enfin, une limite est aussi posée par ce que l'on appelle l'explicabilité. Lorsque nous réagissons à un stimulus, nous sommes évidemment incapables de donner une explication précise de notre réaction, par exemple en énonçant la liste des

neurones qui se sont activés dans notre cerveau et en indiquant vers quels autres neurones l'impulsion électrique s'est propagée par les synapses.

Dans tous les cas, ce genre d'explication ne pourrait pas être comprise par nos congénères. Le fait qu'un réseau de neurones artificiels soit capable d'une telle prouesse ne lui accorde, pour les mêmes raisons et de ce point de vue, aucun avantage sur l'être humain. En revanche, le cerveau humain surpasse la machine lorsque ces signaux neuronaux atteignent le niveau de la conscience ; il devient alors capable de formuler en langage naturel, avec des sons, des lettres ou même des schémas, une justification compréhensible – qu'elle soit juste ou erronée – de son fonctionnement. Ceci ouvre la voie à une amélioration, d'origine individuelle ou sociale, de son fonctionnement, sans forcément nécessiter une grande quantité de nouvelles données d'apprentissage.

Il n'est donc pas du tout exclu que dans un grand nombre d'applications, on assiste dans un futur proche au grand retour de l'approche dite « symbolique », possiblement intégrée aux approches de type connexionniste. C'est en tout cas ce que

suggère une étude réalisée conjointement, en mars 2020, par des chercheurs de laboratoires mondialement réputés comme le MIT, IBM, l'université d'Harvard et Google DeepMind (Virol, 2020).

Pour en savoir plus

- CHAUMARTIN F.-R., LEMBERGER P., *Le traitement automatique des langues. Comprendre les textes grâce à l'intelligence artificielle*, Malakoff, Dunod, 2020.
- KRIVINE H., *Comprendre sans prévoir, prévoir sans comprendre*, Paris, Cassini, 2018.
- LE CUN Y., *Quand la machine apprend. La révolution des neurones artificiels et de l'apprentissage profond*, Paris, Odile Jacob, 2019.

Conclusion

L'intelligence artificielle consiste à explorer les différentes capacités du cerveau humain, en étudiant ses fonctions ou sous-fonctions les unes après les autres et en cherchant à transférer ces capacités à des machines. Cependant, quand nous parvenons à créer un artifice capable d'imiter, voire de dépasser, une capacité du cerveau humain, nous réalisons que ce que nous avons saisi n'était pas de l'intelligence.

On doit à Larry Tesler (1945-2020), le célèbre inventeur du « copier-coller », cette définition de l'intelligence citée par Douglas Hofstadter : « Tout ce que les machines n'ont pas encore fait » (Hofstadter, 2008, p. 674).

En effet, il est possible que la plus grande force du cerveau humain réside dans sa préconfiguration, particulièrement bien adaptée au monde dans lequel il ressent, raisonne et imagine.

Ces avantages innés sont issus d'un processus d'évolution génétique par sélection naturelle qui a duré plusieurs milliards d'années. Lorsque l'on garde cela à l'esprit, on mesure mieux la difficulté de l'objectif que se donnent les chercheuses et chercheurs en intelligence artificielle : leur objectif s'éloigne au fur et à mesure qu'elles et ils s'en approchent.

Mais si l'intelligence est « tout ce que les machines n'ont pas encore fait », comment donner un nom à « ce que les machines ont déjà fait » ? Ce nom n'est-il pas la fameuse « intelligence artificielle » ? Et dans ce cas, pour quelle raison devrait-on limiter cette définition aux dernières approches en vogue dans le domaine ?

Pour ma part, les termes « intelligence artificielle » et « informatique » font référence à un même objet. En effet, même la plus simple des machines à calculer est capable de réaliser automatiquement un traitement de l'information qui était auparavant l'exclusivité de son concepteur : l'esprit humain.

Aujourd'hui, ce sont les développements des réseaux de neurones artificiels qui sont les plus spectaculaires en matière de systèmes opérationnels. Ils sont souvent critiqués pour leur opacité,

en particulier dans des cas d'utilisation sensible. Ce « défaut » est une source de surprises, bonnes ou mauvaises. En programmation informatique traditionnelle, on sait que la machine fait ce qu'on lui demande et non ce que l'on croit lui avoir demandé. C'est parfois la source d'une mauvaise surprise : le bogue.

De même, la bonne ou la mauvaise surprise issue d'un réseau de neurones artificiel est rarement perçue comme une preuve de créativité, d'autonomie ou de conscience. Nous n'avons donc pas encore confié la recherche en intelligence artificielle aux systèmes produits par cette recherche.

Si l'on cherche à tracer le cheminement intellectuel qui a permis le développement de ces technologies, on peut très vite remonter à l'époque de Charles Babbage (1791-1871), mathématicien britannique qui a énoncé pour la première fois le principe de l'ordinateur : la machine analytique. On pourrait également poursuivre un peu l'exploration épistémologique et accepter que cette apparition n'a pas pu se produire sans la formalisation de « lois du raisonnement » introduites par Georges Boole (1815-1864) et Augustus de Morgan (1806-1871). Par ailleurs, il serait passionnant de tracer les successions d'héritages jusqu'à l'Antiquité, aussi bien

occidentale qu'orientale : on rencontrerait ainsi un grand nombre de philosophes et de logiciens, dont Muhammad al-Khwarizmi (780 env.-850 env.), mathématicien perse qui a donné son nom au mot « algorithme ».

L'invention de l'informatique en général et de l'intelligence artificielle en particulier est donc un exploit collectif, international, multidisciplinaire. Celui-ci n'est rendu possible que par l'accumulation, le partage et la remise en cause des découvertes réalisées au cours du temps par de multiples générations d'êtres humains confrontées à une immense diversité de questions et de problèmes.

Ainsi, lorsque nous tentons de percevoir l'intérieur de cette boîte translucide qu'est notre cerveau, comme lorsque nous concevons ces automates artificiellement intelligents, notre principal atout est sans surprise… l'ouverture d'esprit !

Lexique

- **Biais cognitif.** Distorsion dans le traitement cognitif d'une information. Le terme « biais » fait référence à une déviation systématique de la pensée logique et rationnelle par rapport à la réalité.

- **Cellules gliales.** Cellules qui forment l'environnement des neurones. Elles jouent un rôle de soutien et de protection du tissu nerveux en apportant les nutriments et l'oxygène, en éliminant les cellules mortes et en combattant les agents pathogènes. Pendant longtemps, l'implication des cellules gliales dans le traitement de l'information nerveuse a été sous-estimée par rapport au rôle prépondérant des neurones, mais il est aujourd'hui reconnu qu'elles exercent une action modulatrice

sur la neurotransmission, même si le détail de ces mécanismes reste mal compris.

- **Corrélation.** Dans le domaine des statistiques, la corrélation décrit la relation entre deux variables. Les variables sont corrélées si la modification de l'une est suivie d'une modification de l'autre. La corrélation positive décrit la relation entre deux variables qui changent dans le même sens, tandis qu'une corrélation négative (ou inverse) décrit la relation entre deux variables qui changent dans des directions opposées.

- **Effet papillon.** Expression qui résume une métaphore concernant le phénomène fondamental de sensibilité aux conditions initiales de la théorie du chaos. Pour simplifier, il s'agit des phénomènes dans lesquels des causes infimes peuvent avoir des conséquences très importantes.

- **Heuristique.** En algorithmique, c'est une approche empirique qui permet de réduire un espace de recherche, parfois au prix d'une perte d'optimalité parfaite.

- **Métacognition.** En psychologie, la métacognition est la « cognition sur la cognition » (le préfixe μετά signifiant « sur, à propos » en grec ancien). Autrement dit, la métacognition consiste à avoir une activité mentale sur ses propres processus mentaux, c'est-à-dire « penser sur ses propres pensées ».

- **Nerd.** Stéréotype visant les personnes tellement passionnées par les sciences et technologies qu'elles en deviennent marginales. Peut être considéré comme la version péjorative du terme *geek*.

- **Récursif, réentrant.** En informatique, cela consiste à écrire un algorithme ou un programme qui fait appel à lui-même. C'est quelque chose de très fréquent dans les règles de grammaire. Par exemple, une *proposition* peut être constituée de deux propositions séparées par une conjonction.

- **Synapse.** La synapse désigne une zone de contact fonctionnelle qui s'établit entre deux neurones, ou entre un neurone et une autre cellule (cellules musculaires, récepteurs sensoriels).

Bibliographie

- AGID Y., MAGISTRETTI P., *L'homme glial. Une révolution dans les sciences du cerveau*, Paris, Odile Jacob, 2018.
- BOOLE G., *Les lois de la pensée*, Paris, Vrin, 1992 (édition originale, Londres, 1854).
- DÈCLE J.-P., « Je me souviens de Nicolas Ruwet et de Maurice Gross », *L'Être*, 2006. Disponible sur : https://www.larevuedesressources.org/incandescence-de-la-linguistique-je-me-souviens-de-nicolas-ruwet-et-de-maurice-gross,2421.html (consulté le 16/11/2020).
- DEHAENE S., « La métaphore de l'ordinateur fait du cerveau un dispositif de traitement inconscient de l'information », *Pour la Science*, hors-série n° 108, août-septembre 2020, p. 8-11.
- GOULD S.J., *La mal-mesure de l'homme*, Paris, Odile Jacob, 1997.
- GRACCI F. « Sait-on combien il y a d'atomes dans l'Univers ? », *Science & Vie*, 1172, 2015.

- HOFSTADTER D.R., *Gödel, Escher, Bach : les brins d'une guirlande éternelle*, Paris, Dunod, 2008 (édition originale, New York, 1979).
- KRIVINE H., *Petit traité de hasardologie*, Paris, Cassini, 2016.
- LAPLACE P.-S. de, *Essai philosophique sur les probabilités*, 5e éd., Paris, Bachelier, 1825.
- O'NEIL C., *Algorithmes : la bombe à retardement*, Paris, Les Arènes, 2018.
- PEYRON D., *La construction sociale d'une sous-culture : l'exemple de la culture geek*, thèse de doctorat en Sciences de l'information et de la communication (dir. J.-P. Esquenazi), Université Jean Moulin-Lyon 3, 2012. Disponible sur : http://www.theses.fr/2012LYO30089 (consulté le 16/11/2020).
- QUERLIOZ D., GROLLIER J., « Quand la spintronique imite le cerveau », *Pour la Science*, n° 515, septembre 2020. Disponible sur : https://www.pourlascience.fr/sd/technologie/quand-la-spintronique-imite-le-cerveau-19913.php (consulté le 16/11/2020).
- SCHOVANEC J., *Nos intelligences multiples. Le bonheur d'être différent*, Paris, Éditions de l'Observatoire, 2018.
- TURING A.M., « Computing machinery and intelligence », *Mind*, 49, 1950, p. 433-460.
- VIGEN T., *Spurious Correlations*, Hachette Books, 2015. Disponible sur : https://tylervigen.com/spurious-correlations (consulté le 20/10/2020).
- VIROL G, « Rassurez-vous, les algorithmes de deep learning ne comprennent pas ce qu'ils voient », *L'Usine Nouvelle*, 2020. Disponible sur : https://www.usinenouvelle.com/editorial/rassurez-vous-les-algorithmes-de-deep-learning-ne-comprennent-pas-ce-qu-ils-voient.N938541 (consulté le 16/11/2020).
- YE F., ZHOU S., VENKAT A., MARCUS R., TATBUL N., TITHI J.J., HASABNIS N., PETERSEN P., MATTSON T., KRASKA T., DUBEY P., SARKAR V. GOTTSCHLICH J., « MISIM. A novel code similarity system », ICLR 2021 Conference Papers. Disponible sur https://arxiv.org/pdf/2006.05265.pdf, 2020 (consulté le 16/11/2020).

Biographie de l'auteur

Pierre Jourlin est l'auteur d'une thèse de doctorat en reconnaissance automatique de la parole et du locuteur, intégrant la lecture labiale à l'analyse du contenu acoustique. À l'occasion d'un contrat postdoctoral de trois ans à l'Université de Cambridge, au Royaume-Uni, il a développé un système de recherche d'informations dans les informations télévisées, s'appuyant sur une transcription automatique de l'audio. Depuis 2000, il est enseignant-chercheur en informatique dans le domaine du traitement automatique du langage. Il est porteur d'un projet de transfert technologique lié à la suggestion documentaire pour la télé-expertise médicale. Il enseigne, entre autres, dans le master « Intelligence Artificielle » et le master « Gouvernance Numérique » d'Avignon Université.

Table des matières

MIDISCIENCES

Au sein des Éditions Universitaires d'Avignon (EUA), la collection MidiSciences existe depuis 2018. Elle est née d'une volonté de poursuivre la dynamique initiée par le succès des conférences courtes tout public, intitulées « MidiSciences » et présentées à l'université d'Avignon chaque année depuis 2013, à l'heure méridienne évoquée par le titre.

Cette collection entend proposer une photographie des connaissances scientifiques actuelles en s'appuyant sur la diversité des disciplines étudiées par les laboratoires de recherche d'Avignon Université et de sa région.

Pour chaque titre proposé, lectrices et lecteurs trouveront aussi des informations pour aller plus loin dans la connaissance des sujets proposés ainsi que des références bibliographiques permettant d'approfondir les thématiques mises à l'honneur.

Destinée au grand public ainsi qu'aux lycéens et étudiants, cette collection éclaire, d'une manière simple et accessible, les enjeux et les mécaniques de la science, de toute science, en train de se construire.

DERNIÈRES PARUTIONS

Julie Deramond, Jessica de Bideran, Patrick Fraysse (dir.)
Scénographies numériques du patrimoine :
expérimentations, recherches et médiations

Alain Thiéry
Les méduses de Provence : une vie cachée

Giuseppe Sofo
Les éclats de la traduction :
langue, réécriture et traduction dans le théâtre d'Aimé Césaire

Nicolas Soulas
Révolutionner les cultures politiques :
l'exemple de la vallée du Rhône, 1750-1820

Stephan et Alexis Gruss
Ex nihilo

Charlotte Heinzlef
Recettes de résilience urbaine :
faire face aux inondations

Antonia Amo Sánchez, Sophie Gaillard, Marie Galéra, Paul Payan (dir.)
Les Carmes :
théâtre et patrimoine à Avignon

Valentina Garavaglia, Paola Ranzini (dir.)
Pirandello 150 :
un auteur en quête d'un personnage - Un autore in cerca di un personaggio

Quentin Amalou, Damien Malinas
Faire le OFF : sociologie des publics du Festival OFF d'Avignon

Carnet de recherche des Éditions Universitaires d'Avignon
eua.hypotheses.org

Achevé d'imprimer en février 2021
Dépôt légal : février 2021

Fabriqué en France par
printteam® www.print-team.fr
sur papiers éco-responsables