

Alan Turing : les machines à calculer et l'intelligence

L'article «Computing Machinery and Intelligence» qu'Alan Turing publie en 1950 est un incontournable pour tout chercheur en informatique ou toute personne intéressée par l'intelligence artificielle. Facilement accessible sur Internet (le site <scholar.google.fr> donne accès à 218 versions¹) ou même en fac-similé sur le site TuringArchive.org², cet article présente une discussion sur ce qui, selon l'auteur, dotera les machines de capacités intelligentes et de ce qui permettra d'en convaincre tout un chacun.

Alan M. Turing (1912-1954) est un personnage qui aura marqué le XX^e siècle, d'abord par ses travaux pendant la première moitié de ce siècle, puis par leurs conséquences dans la seconde moitié. Mathématicien, il a formalisé les bases de ce qui est devenu l'ordinateur. Patriote, il a développé un calculateur pour décrypter les messages secrets du haut commandement allemand pendant la Seconde Guerre mondiale et a participé ainsi au maintien d'une Angleterre libre. Visionnaire, il a imaginé avant l'heure ce que serait l'intelligence artificielle. Tragédien, il se serait suicidé en croquant une pomme imprégnée de cyanure. Auteur d'une quinzaine d'articles scientifiques seulement³, il totalise tout de même près de 25 000 citations et continue toujours d'être largement cité, avec plus de 8 000 citations à son actif ces cinq dernières années⁴.

Outre un article majeur sur la morphogénèse en 1952, ses deux papiers les plus fameux dans le monde de l'informatique sont «On Computable Numbers, With an Application to the Entscheidungsproblem» en 1937, où il pose les bases de la machine de Turing et de la calculabilité, et «Computing Machinery and Intelligence» en 1950, où il pose la question de l'intelligence pour les machines et énonce le test de Turing. À eux seuls ces

trois articles comptabilisent plus de 20 000 citations (et le présent texte en ajoutera un au compteur!).

Relire aujourd'hui «Computing Machinery and Intelligence» n'est pas faire acte de nostalgie mais est au contraire une source indispensable pour qui s'intéresse à la cognition. Précisons que la (re)lecture de cet article nécessite d'avoir présent à l'esprit que si la presse utilisait déjà à l'époque l'expression *cerveau électronique* pour parler des ordinateurs naissants, ce n'est que six années après sa parution que sera construit le terme «intelligence artificielle». Sur l'article lui-même, il est tout d'abord délicieux de lire un texte rédigé par une personne d'autorité, expliquant aux béotiens ce que sont le calcul formel, l'algorithme et le fonctionnement des ordinateurs (en 1950, tout le monde était novice sur ces sujets). Ensuite, ce texte développe l'idée que la puissance de calcul des machines amène ou amènera nécessairement à se poser la question du statut de l'intelligence : si c'est par sa faculté de raisonnement que je crédite autrui d'une certaine intelligence, comment réagirai-je face à un calculateur électronique? Enfin, et surtout, Turing y expose un test destiné à mettre en évidence les traces d'un raisonnement intelligent, et poursuit par la description des neuf objections qui fuseront dès que sera posée la question : *Est-ce que les machines peuvent penser?*

Ce test, introduit sous l'appellation du *jeu de l'imitation*, est passé à la postérité sous le nom de «test de Turing». Dans un jeu de questions-réponses, un ordinateur est programmé pour tenter de se faire passer pour un humain. Si au bout de quelques minutes d'échanges, un vrai humain ne décèle pas la supercherie, l'ordinateur est réputé avoir *passé le test de Turing*. Dans cet échange, la personne n'aura pas vu l'aspect calculatoire

des réponses fournies, ou plus exactement, la programmation de la machine aura permis de simuler les réactions et le libre arbitre d'un véritable interlocuteur. Ce défi stimula beaucoup les informaticiens et plusieurs réalisations ont vu le jour, comme Eliza, développé par Weizenbaum en 1966. Depuis 1990, il existe même un prix annuel, le Loebner Price, décerné au programme informatique qui réalise le meilleur agent conversationnel du moment.

Face à cette situation de bluff potentiel, Turing développe des arguments pour contrer les objections qui ne manqueront pas d'émaner de la part de ceux pour qui l'intelligence serait une faculté uniquement humaine. Nous reprenons ici brièvement ces neuf objections.

La première est celle de l'objection théologique : pas d'âme, pas de faculté de penser ! Turing avance qu'une distinction vivant/inanimé aurait été plus difficile à éluder que la distinction théologique humain/reste du monde. Pour lui, il n'est pas irrévérencieux de proposer que les machines puissent avoir une âme ; l'homme ne serait finalement que l'instrument d'une mutation voulue et approuvée par Dieu, destinée à donner un semblant d'âme aux machines. Et d'ailleurs, lorsque deux humains font un enfant, c'est bien Dieu, et non les parents eux-mêmes, qui au final lui accorde une âme.

La seconde objection est celle de « faire l'autruche » : il serait préférable de croire que les machines ne peuvent pas penser, car cela évite d'aborder les terribles conséquences qui adviendraient. L'Homme pourrait alors perdre sa suprématie sur Terre. Plutôt que de répondre à cette objection, Turing propose de consoler les tenants de cette vision.

La troisième est une opposition mathématique : certains résultats de la logique mathématique tendent à montrer les limitations de la puissance des machines à états discrets. Turing cite plusieurs travaux d'auteurs des années 1930 (dont les siens !) et avance que le problème vient de l'aspect fini du stockage de ces machines. La

réfutation avancée est toujours du genre « Est-ce que les machines programmées de telle ou telle façon réussiront à répondre correctement à telle question ? » Pour Turing, ce genre de recherche de la faillibilité des machines permet à certains de jubiler en affichant un sentiment de supériorité, mais ces limites s'appliquent également à l'esprit humain (qui ne se serait donc jamais trompé ?) ; par ailleurs, formuler une telle question permet déjà de travailler à repousser un peu plus loin les limites des machines.

Le quatrième argument a trait à la conscience : pour s'assurer qu'une machine pense, il faudrait se mettre à la place de cette machine pour pouvoir se sentir penser. Mais à l'extrême cela voudrait dire également que puisque je ne suis pas à la place d'autrui, je ne puis jamais m'assurer que quelqu'un d'autre que moi pense ! Pour sortir de cette attitude solipsiste et éviter de passer le reste de sa vie à regarder les autres de façon suspicieuse, il est donc préférable d'accepter la pertinence du test de Turing.

Le cinquième argument porte sur les diverses incapacités : il serait impossible de programmer une machine capable de X – « X » pouvant porter sur des capacités comme être gentil, ingénieux, beau, convivial, avoir de l'initiative, du sens de l'humour, discerner le bien du mal, faire des erreurs, tomber amoureux, apprécier les fraises à la Chantilly, charmer quelqu'un, apprendre de son expérience, utiliser les mots à bon escient, être le sujet de sa propre pensée, avoir une diversité de comportements comme un humain, réaliser quelque chose de vraiment nouveau. C'est très souvent l'argument de la conscience qui se cache derrière l'affirmation de telle ou telle incapacité. Turing en discute certaines, mais c'est avec l'augmentation de la taille mémoire des machines qu'il voit ces verrous sauter dans le futur.

La sixième objection a trait à Lady Lovelace⁵ : les machines n'ont aucune prétention à générer quoique ce soit, elles ne réaliseraient que ce que nous leur deman-

dons de faire. L'intelligence ne serait donc que dans la tête du programmeur. Selon Turing, on retrouve ici l'argument que les machines ne seraient pas capables de faire quelque chose de réellement nouveau, de nous surprendre. Or, Turing se dit justement très souvent surpris par les résultats de ses calculs programmés. Bien que ce soit certainement dû au fait qu'il ne pousse pas lui-même les calculs mentaux pour anticiper la complète réaction de la machine, il n'en demeure pas moins surpris.

Le septième argument porte sur la continuité du système nerveux : une petite erreur sur l'influx d'un neurone peut porter à de grandes divergences de comportement, et ce type d'influence ne serait pas reproductible sur une machine à états discrets. En ce sens, les processus continus et les processus discrets seraient fondamentalement différents. Pourtant, dans le cadre du test proposé, la distinction d'une telle organisation interne de la machine n'est pas possible. Et Turing donne comme exemple que l'on ne peut pas différencier un analyseur différentiel d'un calculateur numérique, à la seule vue des résultats qu'ils affichent.

Le huitième argument est lié au comportement informel des humains : il n'est pas possible de décrire tous les comportements existants dans toutes les situations potentielles. Comment réagir par exemple à l'affichage simultané de deux feux de circulation, un rouge et un vert ? Turing rappelle qu'il y a là une confusion entre les lois biologiques qui régissent la vie en induisant des comportements (si vous pincez quelqu'un, alors il grimace) et les préceptes conscients avec lesquels chacun règle sa conduite de vie (si vous voyez un feu rouge, alors vous vous arrêtez). Il dit également que pour conserver un peu de rigueur scientifique, ces règles devraient être inférées à partir d'observation. Et justement, avec la machine qu'il a programmée à Manchester, il met au défi quiconque de construire des règles comportementales à partir de la seule observation des résultats affichés.

Le neuvième argument provient d'hypothétiques

capacités extrasensorielles, comme la télépathie, la clairvoyance, la précognition ou la psychokinésie. Même s'il raille un peu de telles croyances, Turing avoue que la machine ne pourrait alors jamais se faire passer pour quelqu'un ayant des capacités télépathiques. Une telle personne disposerait manifestement de moyens supplémentaires pour répondre aux questions qui lui sont posées, en sondant l'esprit du questionneur ou en anticipant les réponses attendues ; bref un statut qui placerait la personne sur un plan inaccessible à la machine et à son seul traitement possible de l'aléatoire. Si l'on admet que les capacités extrasensorielles existent, il faudrait revoir les conditions d'expérimentation du test de Turing !

À travers cette série d'objections et d'arguments, on voit bien que Turing était persuadé dès le début de l'informatique que les calculateurs numériques nous renverraient à la question de l'intelligence, et que ce serait à nous humain de décider si les capacités de stockage et de traitement des machines en feraient des entités sub-critiques en matière de connaissances, comme le sont les animaux, ou bien super-critiques comme le sont les humains. Et il chiffre cette taille critique à une capacité inférieure au peta-octet, remarquant que l'*Encyclopaedia Britannica* occupe un volume d'environ 2 giga-octets. Turing va même jusqu'à proposer une piste de recherche pour constituer une telle intelligence de façon artificielle. Il propose une approche en décomposant le problème en deux parties : celui de la constitution d'un noyau de base («the child program» ; faisant notamment appel aux essais-erreurs, à l'aléatoire, aux mutations) et celui de l'évolution et de l'autonomisation de ce noyau («the education process» ; s'appuyant sur des processus de récompenses ou de punitions).

Dans cet article de 1950, Turing expose ses idées pour doter les calculateurs numériques de processus d'apprentissage et aboutir un jour à ce que de telles machines puissent bluffer un humain en se faisant passer pour l'un des leurs. Turing justifie son test comme étant

le mode opératoire pour mesurer si ce bluff est ou non atteint. Mais il montre aussi la quête d'un individu face à la société. D'après ses contemporains, Turing passait son temps à se prendre pour une machine et au final il cherchait une description objective de la part d'humanité résidant en lui. Aussi, le test de Turing peut tout à fait être entièrement retourné et se comprendre comme

la recherche d'une réponse à la question: «Y a-t-il un être humain face à moi?»

Benoît Le Blanc
École nationale supérieure de cognitive, Bordeaux-INP
ISCC

NOTES

1. Cf. <scholar.google.fr/scholar?oi=bibs&hl=fr&cluster=15753885411960689369>, consulté le 14/02/2014.
2. Cf. <www.turingarchive.org/browse.php/B/19>, consulté le 14/02/2014.
3. Disponibles en ligne sur: <www.turingarchive.org/browse.php/B>, consulté le 14/02/2014.
4. Cf. <scholar.google.fr/citations?user=VWCHlwkAAAAJ&hl=fr>, consulté le 14/02/2014.
5. Ada Lovelace est considérée comme étant la première personne de l'histoire à avoir programmé une machine, en l'occurrence la machine analytique de Babbage (XIX^e siècle).

RÉFÉRENCES BIBLIOGRAPHIQUES

TURING, A., «On Computable Numbers, With an Application to the Entscheidungsproblem», *Proc. London Math. Soc.*, ser. 2, vol. 42, 1937, p. 230-265.

TURING, A., «Computing Machinery and Intelligence», *Mind*, vol. 59, n° 236, 1950, p. 433-460.

TURING, A., «The Chemical Basis of Morphogenesis», *Philosophical Transactions of the Royal Society of London*, ser. B, vol. 237, n° 641, 1952, p. 37-72.