

Une nouvelle intelligence qui s'ignore

Autor(en): **Ulmi, Nic**

Objektyp: **Article**

Zeitschrift: **Horizons : le magazine suisse de la recherche scientifique**

Band (Jahr): **29 (2017)**

Heft 113

PDF erstellt am: **22.07.2024**

Persistenter Link: <https://doi.org/10.5169/seals-821699>

Nutzungsbedingungen

Die ETH-Bibliothek ist Anbieterin der digitalisierten Zeitschriften. Sie besitzt keine Urheberrechte an den Inhalten der Zeitschriften. Die Rechte liegen in der Regel bei den Herausgebern. Die auf der Plattform e-periodica veröffentlichten Dokumente stehen für nicht-kommerzielle Zwecke in Lehre und Forschung sowie für die private Nutzung frei zur Verfügung. Einzelne Dateien oder Ausdrucke aus diesem Angebot können zusammen mit diesen Nutzungsbedingungen und den korrekten Herkunftsbezeichnungen weitergegeben werden. Das Veröffentlichen von Bildern in Print- und Online-Publikationen ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. Die systematische Speicherung von Teilen des elektronischen Angebots auf anderen Servern bedarf ebenfalls des schriftlichen Einverständnisses der Rechteinhaber.

Haftungsausschluss

Alle Angaben erfolgen ohne Gewähr für Vollständigkeit oder Richtigkeit. Es wird keine Haftung übernommen für Schäden durch die Verwendung von Informationen aus diesem Online-Angebot oder durch das Fehlen von Informationen. Dies gilt auch für Inhalte Dritter, die über dieses Angebot zugänglich sind.

Une nouvelle intelligence qui s'ignore

Que vaut une découverte si on ne peut la comprendre? L'arrivée des machines intelligentes dans les laboratoires interroge l'essence même de la connaissance. *Par Nic Ulmi*

Formuler des hypothèses, faire des découvertes: ces termes sont au cœur de la démarche scientifique. Depuis le début des années 2000, des machines dotées d'intelligence artificielle ont commencé à s'engager sur ce terrain. Elles développent de nouvelles manières de produire des résultats et assument un rôle qu'on imaginait autrefois réservé aux seuls humains. De plus en plus performantes, mais dépourvues de la capacité de comprendre leurs propres réussites, ces machines laissent entrevoir une vision déroutante: une recherche scientifique automatisée, robotisée, industrielle.

Trois exemples pour commencer. A l'Université Tufts au Massachusetts, un réseau de neurones artificiels s'est attaqué à une vieille énigme de la biologie en formulant une hypothèse sur la régénération du planaire, un ver d'eau douce qui sait faire repousser sa queue ou sa tête lorsqu'il les perd. Un autre, à l'Université d'Adelaïde en Australie, a découvert le moyen optimal de produire un condensat de Bose-Einstein, un ensemble de bosons refroidi près du zéro absolu qui rend apparents des phénomènes quantiques macroscopiques. Et les machines de la startup Insilico Medicine à l'Université Johns Hopkins de Baltimore ont mis au point des molécules potentiellement utiles en oncologie.

«Nos esprits peuvent effectuer un travail d'abstraction sans effort.»

Roger Schank

Vétéran de l'intelligence artificielle (IA), Roger Schank ne se montre pas du tout convaincu. «Tout cela, ce n'est absolument pas de l'intelligence artificielle», proteste-t-il. Le chercheur, actif aux Universités Yale, Northwestern et Carnegie-Mellon, redoute la répétition d'un scénario connu. Au début des années 1970 et à la fin des années 1980,

deux vagues de hype autour de l'intelligence artificielle font place à des «hivers de l'IA»: les attentes excessives commencent à retomber, l'intérêt du public entre en hibernation, les fonds sont gelés. «La presse s'y donne maintenant à cœur joie, mais elle passera ensuite à autre chose. Le problème, c'est qu'entre-temps la recherche sur l'IA risque d'être tuée une fois de plus. Cela ne m'amuse pas.»

Le steak et le coiffeur

De quoi s'agit-il, si ce n'est pas de l'intelligence? «Les exemples que vous citez sont en fait des programmes de pattern matching (filtrage de motif, ndlr), répond Roger Schank. C'est le procédé utilisé par Facebook pour identifier votre visage sur des photos.» Autrement dit, les machines génèrent des motifs (le schéma d'une molécule, le plan de régénération d'un ver planaire) qu'elles comparent ensuite à des régularités identifiables dans des bases de données préexistantes, passant ces dernières au peigne fin. «Mais la manière dont les scientifiques font des découvertes est tout autre: elle commence par le fait d'être désorienté. L'incompréhension que vous ressentez face à un phénomène vous conduit à formuler une hypothèse, puis à la tester. C'est ce processus, précisément, qu'on appelle science.»

Roger Schank aime raconter une histoire pour illustrer la différence entre la découverte humaine et l'apprentissage automatique. Il l'appelle «Le steak et la coupe de cheveux»: «Je discutais avec un collègue à Yale, me plaignant que je ne parvenais jamais à me faire servir un steak saignant: la viande arrivait toujours sur ma table trop cuite. Pourquoi diable? me demandais-je. Pour toute réponse, le collègue me dit: j'ai vécu en Angleterre, autrefois, et là-bas aucun coiffeur n'arrivait à me faire une coupe en brosse. Eureka! Chacune des deux histoires explique l'autre: en passant à un niveau d'abstraction supérieur, on se rend compte qu'elles sont identiques. Nous avons affaire, dans les deux cas, à des per-

«La manière dont les scientifiques font des découvertes commence par le fait d'être désorienté.»

Roger Schank

sonnes qui sont bien capables de fournir le service qu'on leur demande, mais qui ne parviennent pas à obtempérer, car elles jugent nos requêtes trop extrêmes.» Moralité de l'histoire? «Nos esprits savent effectuer ce travail d'abstraction sans effort. Ils sont en mesure de regarder une chose comme si elle était une autre. Ils sont motivés par le fait que nous avons naturellement des objectifs: rassasier notre appétit, ou satisfaire une curiosité mise en branle par une situation qui nous rend perplexes.»

C'est ainsi que fonctionne notre cerveau. Mais la nouvelle vague de recherches et d'applications dans le domaine de l'IA, née à la fin des années 1990, ne vise plus à reproduire le modèle de cognition de l'esprit humain. Ce tournant mène vers un type de connaissance inédit, où le savoir est produit par apprentissage automatique à partir des mégadonnées. Cette bifurcation a tiré l'IA de sa léthargie et a permis les suggestions d'Amazon, à Siri, à la victoire d'AlphaGo sur l'un des meilleurs joueurs de go au monde, ou encore à prédire l'expression des gènes d'une bactérie, selon une étude de l'Université de Pennsylvanie de 2016.

Cette technologie peut-elle véritablement produire de la science? Ou faut-il pour cela un besoin existentiel de comprendre, un désir de connaître, une «libido sciendi», comme aurait dit saint Augustin - autant de choses qu'une machine ne saurait avoir?

La théorie du fun

On se tourne vers Jürgen Schmidhuber, codirecteur de l'Institut Dalle Molle de recherche en intelligence artificielle (Idisia) à Lugano. «Lorsque l'IA aura grandi, elle l'appellera peut-être «papa», écrivait à propos de lui le New York Times en novembre 2016. Le chercheur a notamment développé en 1997 l'approche de la «long short-term memory», utilisée aujourd'hui par les programmes de reconnaissance du langage parlé. Il considère qu'un même système de récompense peut être à l'œuvre dans le cerveau des scientifiques, des artistes et des bébés, ainsi que dans les réseaux de neurones

artificiels. Ces différentes formes d'intelligence peuvent éprouver une forme de plaisir lorsqu'elles arrivent à identifier, dans le désordre du monde, des récurrences et des régularités qui leur étaient inconnues.

«Imaginez un programme qui doit modéliser une séquence d'images montrant cent pommes en train de tomber, explique le chercheur. Sans connaître la force de la gravité, il aura besoin d'un nombre considérable de bits pour encoder les données. Mais après avoir découvert cette régularité, il pourra l'utiliser pour faire des prédictions et moins de bits lui seront nécessaires. Cette différence entre avant et après, donc cette compression des données, mesure la profondeur de la nouvelle connaissance acquise par réseau. Cela déclenche un signal de récompense: un moment de joie intérieure pour le réseau, si vous voulez.»

Jürgen Schmidhuber formalise ce mécanisme de récompense dans sa théorie formelle du fun et de la créativité. Pour lui, il se manifeste chez le musicien surpris par une harmonie nouvellement créée, et expliquerait également l'humour: «Lorsque la chute d'une blague fait apparaître un schéma inattendu dans le récit, elle permet tout à coup d'en compresser les données. C'est alors qu'on rit.» Ce phénomène est crucial pour programmer des machines dotées de curiosité artificielle et susceptibles de faire des découvertes.

Pour cela, le réseau doit fonctionner comme un duo, explique le chercheur: «D'un côté, le contrôleur exécute des actions et mène des expériences générant des données; il est motivé par la maximisation des récompenses reçues. De l'autre côté, un modèle envoie une rétribution à chaque fois que la découverte d'une nouvelle régularité lui permet de compresser les données. C'est un tel système qu'il faut mettre en place si l'on veut construire un système scientifique artificiel.»

Une science aléatoire

Le réseau neuronal d'Insilico Medicine recourt à un tel système dual, explique

Polina Mamoshina, généticienne et informaticienne associée au projet. «Le générateur est programmé pour créer de manière virtuelle et aléatoire des structures moléculaires. Le discriminateur s'entraîne avec des bases de données à reconnaître des molécules capables d'inhiber la croissance des tumeurs. Initialement, le but du générateur est de piéger le discriminateur, ce qui amène ce dernier à faire de fausses identifications, tout en lui permettant de s'améliorer par la suite.»

«La chute d'une blague permet tout à coup de compresser les données. C'est alors qu'on rit.»

Jürgen Schmidhuber

Parmi les 60 molécules trouvées par le générateur et validées par le discriminateur, certaines existaient déjà et avaient été brevetées comme agent thérapeutique contre le cancer. «C'est une indication encourageante sur la précision du système, poursuit Polina Mamoshina. Nous allons maintenant entamer le processus de validation des autres molécules in vitro, puis in vivo.» Pour la chercheuse, cette approche révolutionne le domaine: au lieu de cribler à l'aveugle une grande quantité de composés, le but est de créer des médicaments sur demande.

En attendant les machines curieuses de Jürgen Schmidhuber et des systèmes à l'aise avec l'histoire du steak et de la coupe de cheveu de Roger Schank, l'apprentissage automatique et les mégadonnées sont en train de redéfinir la nature de la pratique scientifique. A l'Université de Bristol, le professeur d'intelligence artificielle Nello Cristianini prend le parti d'embrasser ces nouveaux outils, tout en appelant à délimiter leur champ d'application: «Je travaille avec la machine learning depuis vingt ans. J'ai le plaisir d'affirmer qu'il

fonctionne. La machine apprend, dans le sens où elle améliore ses performances avec l'expérience.»

Ce sont ces approches qui se cachent derrière une part énorme des bénéfices d'Amazon lorsque les bons livres se voient recommandés à la bonne personne. «Il faut souligner que ces algorithmes ne contiennent pas un modèle psychologique de chaque utilisateur ni un modèle de critique littéraire pour chaque livre, poursuit Nello Cristianini. Ils mènent un raisonnement purement statistique: des gens ayant telle caractéristique et tel comportement ont acheté des livres ayant telle particularité. Cela signifie une chose importante: on peut faire une prévision sans avoir une théorie.»

«Je travaille avec le machine learning depuis vingt ans. J'ai le plaisir d'affirmer qu'il fonctionne.»

Nello Cristianini

Mais peut-on vraiment transposer ce modèle à la science? «Il n'y a pas de raison philosophique pour que cela soit impossible, répond le chercheur. Un ordinateur pourrait générer des modèles de molécules et prédire leur toxicité. Ce qu'on gagne? La capacité de mettre au point un médicament *in silico*, sans être obligé de produire toutes les molécules possibles et de les tester sur des souris. Ce qu'on perd? Le moyen de savoir pourquoi le médicament agit.»

L'apprentissage automatique constitue une boîte noire, car il ne nous permet pas de comprendre le raisonnement suivi par la machine. Cette opacité se révèle particulièrement épineuse en dehors du monde

académique, note le chercheur, «lorsque des algorithmes déterminent l'accès à des droits: être admis dans une école ou éligible pour une assurance, bénéficier de la liberté conditionnelle...»

La fin de la théorie?

Ce type d'apprentissage, est-ce encore de la science? «Il ne faut pas se leurrer, poursuit Nello Cristianini. Si le machine learning fonctionne pour faire de l'argent avec des prévisions correctes, cela aboutira à redéfinir progressivement ce qu'on entend par science. Les fonds pour la recherche suivront les applications pratiques, et les autres approches risqueront d'avoir des difficultés de financement.» Doit-on craindre «la fin de la théorie», comme le proclamait Chris Anderson? En 2008, le responsable du magazine *Wired* affirmait que «le déluge des données rend la méthode scientifique obsolète».

«Nous devons nous demander quel est le but des théories, répond Nello Cristianini. Pour moi, la réponse est extrêmement claire. Faire une belle théorie de la mécanique ou de la thermodynamique, comprendre un morceau de notre univers, cela représente une valeur culturelle infinie. Nous avons besoin de connaître le fonctionnement du monde et de nous-mêmes.» Une bonne théorie a une valeur concrète: «Une prévision de type boîte noire n'est pas suffisante lorsque l'enjeu est très élevé, par exemple lorsqu'on envoie une sonde sur Mars ou planifie une opération chirurgicale. Dans ces cas, on veut savoir exactement ce qui se passerait si on changeait tel ou tel paramètre. Cela implique un raisonnement contrefactuel, qu'on ne peut aborder que par des théories.»

D'un côté, des machines probabilistes. De l'autre, des êtres mus par un désir de comprendre, enraciné dans leurs fonc-

tions biologiques et générateur de théories. Les premières remplaceront-elles les seconds? «Ce qu'on pourra confier aux machines, ce seront des tâches spécifiques», avance Nello Cristianini. Exemples? «Etant donné une séquence d'acides aminés, dessine-moi la protéine. A partir d'un génome humain complet - trois milliards de lettres -, trouve-moi les 20 000 gènes qui le composent... Mais déterminer la signification et l'importance de la découverte, cela restera un tâche humaine.»

Le même partage du travail peut s'appliquer dans les sciences sociales: «Je travaille avec des historiens pour mesurer les transitions historiques dans la culture anglaise et italienne à partir de la lecture systématique des journaux. Personne ne peut lire 500 millions d'articles: la machine les parcourt donc pour nous. C'est pourtant l'historien qui explique ensuite pourquoi tel ou tel résultat est important.»

Complémentarité: Jürgen Schmidhuber en est également convaincu. «Dans les laboratoires scientifiques comme ailleurs, les machines rempliront des tâches fastidieuses dont, au fond, les humains ne voudraient pas. Bien sûr, cela va supprimer des postes de travail. Il faudra donc réagir en tant que société pour distribuer les gains. Cela passera par un revenu inconditionnel, des impôts sur les robots ou autre chose.»

Les questions sociales soulevées par la robotisation des sciences restent aussi ouvertes que les questions épistémologiques. Une prédiction automatique est-elle un savoir scientifique? Nello Cristianini relève ce défi: «Je viens tout juste d'engager deux philosophes des sciences pour commencer à réfléchir à cette question.»

Nic Ulmi est journaliste indépendant à Genève.



He starts to shake.

H (CONT'D)

It may never be forgiven, but that is just too bad. I have to leave, but I'm not free of the world.

C

Yes. Perhaps I should take it from here. I'm not going to do something.

H

You can't afford to take this anywhere. It's not a dream. But I've got a good time to stay there.

C

Well, I think you can still be back on the table.

H

Mmm. It's a damn thing scared to say. Nothing is going to be a thing but I was the one that got on this rock with a child and then I left the other two.

He is standing in the stars and sitting on the floor.

Le programmeur Ross Goodwin et le réalisateur Oscar Sharp ont conçu Benjamin, le «premier générateur automatique de scénarios». Ils l'ont nourri avec des scripts d'une centaine de films de science-fiction. Le résultat est un film déroutant de huit minutes: *Sunspring*. Il s'avère difficilement accessible à une intelligence humaine.

Image: thereforefilms.com/sunspring