R

ENDEZ-VOUS

IA CONNEXIONNISTE ET IA SYMBOLIQUE : L'INÉVITABLE ASSOCIATION ?

L'essentiel des progrès récents en IA reposent sur une approche appelée «connexionnisme». Mais l'IA symbolique, qui a dominé le domaine pendant des décennies, conserve des atouts qui la font revenir dans la course.

L'AUTEUR

JEAN-PAUL DELAHAYE
professeur émérite à l'université de Lille et chercheur au laboratoire Cristal (Centre de recherche en informatique, signal et automatique de Lille)

Les modèles massifs de langage (*large language models*, LLM, en anglais) témoignent d'une révolution dans le domaine de l'intelligence artificielle (IA), en marche depuis 2017 – date de la création de l'architecture «transformers» pour les réseaux de neurones. Cette révolution a éclaté aux yeux du grand public avec la mise à disposition, en novembre 2022, de l'agent conversationnel ChatGPT, de l'entreprise américaine OpenAI. Le principe général de fonctionnement des modèles massifs de langage est assez simple: il consiste à prolonger des morceaux de textes, appelés «prompts» ou «invites», de la manière jugée la plus probable par un réseau de neurones entraîné sur une grande quantité de données. On parle d'«intelligence artificielle générative», car elle produit du contenu – ici, du texte, mais il existe des IA génératives qui produisent des images, du son… Dans le cadre d'un chatbot s'appuyant sur un LLM, le prompt est constitué de quelques lignes soigneusement rédigées par des ingénieurs spécialisés – et auxquelles les utilisateurs n'ont pas accès – et d'une partie de l'historique de la conversation entre l'internaute et le chatbot. Pour formuler des réponses correctes aux questions qui leur sont posées, les LLM exploitent une connaissance statistique de nombreux textes liés à la requête, ainsi que de ce que sont des phrases correctes dans la langue de son interlocuteur. Cette connaissance statistique provient de l'assimilation, par un gros réseau de neurones artificiels, d'un corpus géant de textes: on parle d'«apprentissage profond». Cela donne au robot conversationnel la capacité de créer des phrases nouvelles auxquelles nous donnons du sens, des histoires cohérentes, des réponses précises apparaissant le plus souvent bien informées et reposant sur ce qui ressemble à la compréhension d'une multitude de sujets. On nomme «connexionnistes» de telles approches à base de neurones artificiels.

Pour éviter certains dérapages (propos illégaux ou violents, prises de position politiques, discours à caractère sexuel…) et pour affiner les performances des IA, la phase initiale d'entraînement du réseau de neurones est complétée d'une phase d'ajustement et de filtrage, ou parfois de spécialisation à des domaines particuliers: on parle d'«ajustement fin» (*fine tuning*, en anglais). Cependant, l'essentiel des capacités des LLM provient bien de l'assimilation de la statistique générale d'un ensemble de textes.

Les prouesses des LLM ont surpris le monde entier. Diverses expériences, dont celles récemment menées par Cameron Jones et Benjamin Bergen, de l'université de Californie à San Diego, montrent d'ailleurs que le fameux test de Turing est maintenant satisfait: il est devenu, en pratique, impossible de discerner un humain d'une machine fonctionnant avec un LLM au cours de dialogues écrits limités à quelques minutes.

Pourtant, quand j'ai demandé le 23 juillet 2025 à ChatGPT (https://chatgptfree.ai) le

Jean-Paul Delahaye a également publié : **Aux frontières des mathématiques : Kurt Gödel et l'incomplétude** (Dunod, 2025).

produit de 1234 par 567, il m'a répondu 700578. En réalité, $1234 \times 567 = 699678$. J'ai donc signalé au chatbot que son résultat était faux; il l'a reconnu et m'en a donné un autre, 700878, tout aussi incorrect que le premier. D'autres LLM font parfois mieux, mais en leur soumettant des calculs un peu plus difficiles, comme par exemple celui de 17^{10} ou de 19^{20}, on finit presque toujours par obtenir des réponses fausses. Cette médiocrité des LLM en calcul a de quoi surprendre, puisqu'on conçoit depuis longtemps des programmes qui calculent correctement, même avec des nombres de plusieurs centaines ou milliers de chiffres.

On retrouve les mêmes limites dans la manipulation des tables de vérité, en logique. La plupart des LLM se trompent régulièrement quand on leur demande de reconnaître les formules logiques vraies, appelées «tautologies», comme «(*A* implique NON-*B*) implique (*B* implique NON-A)» ou «((*A* et *B*) implique *C*) équivalent à (*A* implique (*B* implique *C*))». Comme pour les calculs arithmétiques, les techniques permettant ces calculs de tables de vérité semblent hors des capacités des LLM, ou maîtrisée très imparfaitement.

Ces limites, au fond, n'ont rien de véritablement surprenant: les LLM sont des modèles de prédiction conçus pour générer du texte, pas pour effectuer des calculs précis. Leur architecture ne comporte pas, en général, de modules spécialisés pour les opérations mathématiques. Sauf exception, ils n'exploitent pas d'algorithmes de calcul comme le font les humains ou les calculatrices, mais prédisent la séquence de chiffres la plus probable en se basant sur les données avec lesquelles ils ont été entraînés.

AIDE AUX CALCULS

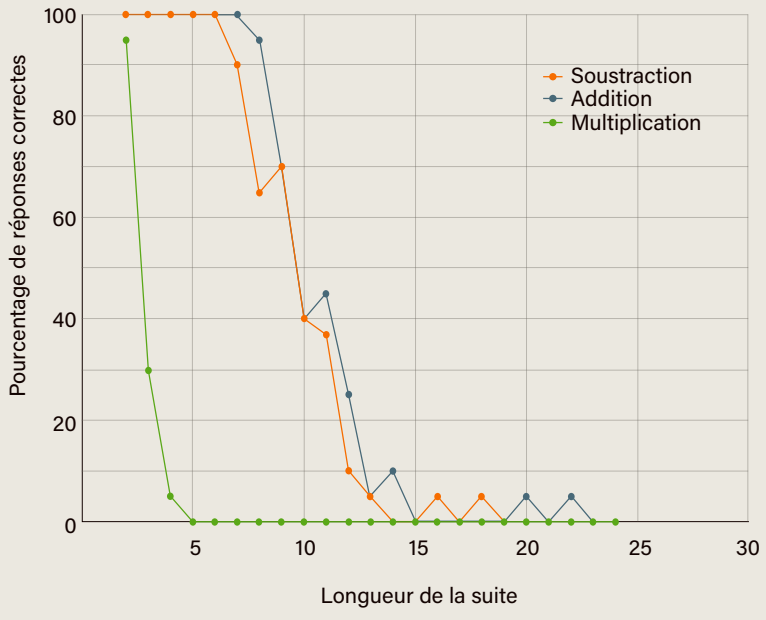
John Loeber, ingénieur en IA à San Francisco et fondateur de l'entreprise Limit, a été intrigué par ce problème et a mené des tests avec ChatGPT4 (*voir l'encadré 1*). Il a par ailleurs synthétisé un certain nombre de travaux ayant tenté d'élaborer des méthodes pour aider les LLM à mieux calculer. Nous répertorions ci-dessous quelques-unes des idées explorées.

Il faut d'abord noter que, quand un LLM analyse un texte, il commence par le séparer en *tokens*, qui sont en quelque sorte les syllabes du texte. Un bon découpage des nombres entiers en *tokens* – qui seront donc, ici, des paquets de plusieurs chiffres – est essentiel pour pouvoir effectuer des manipulations arithmétiques. Séparer les nombres entiers en partant de la gauche et en regroupant les chiffres, disons, par groupes de trois est un procédé naturel pour un LLM, puisque la lecture d'un texte se fait de gauche à droite. Ce n'est pourtant pas une bonne méthode, car 24324+1299334 donnerait alors `[243][24][+][129][933][4]`, ce qui ne se présente pas très bien même si l'on sait additionner les *tokens* deux à deux. On peut imaginer qu'il est préférable de séparer les nombres entiers de grande longueur en commençant par la droite plutôt que par la gauche. S'appuyant sur

UNE EXPÉRIENCE DE CALCUL

1

John Loeber, ingénieur et entrepreneur en IA, a mené une expérience avec le modèle massif de langage GPT-4. Il a créé au hasard des suites d'entiers compris entre 1 et 100. La longueur des suites variait de 2 à 24. Il a demandé au modèle de faire l'addition des nombres de chacune des suites. Pour chaque longueur de suite, il a réalisé 20 essais et mesuré le pourcentage de réponses correctes. Cela fournit la courbe bleue dans le graphe ci-contre. On constate que dès que les suites comportent plus de 8 termes, le pourcentage de bonnes réponses chute et arrive rapidement à 0. John Loeber a opéré le même test pour les soustractions et les multiplications, et les résultats sont fournis par les courbes orange et verte du graphe.



cette remarque, une première approche pour améliorer les performances de calcul d'un LLM consiste, à chaque fois qu'un nombre est présenté au réseau de neurones pendant sa phase d'apprentissage, à inverser l'ordre des chiffres, pour provoquer une séparation en *tokens* plus commode pour opérer des calculs. Ainsi, 24324+1299334 donne [324][24][+][334][299][1]. Ce procédé améliore les résultats obtenus, mais les progrès ne sont pas généraux, et dès que les entiers dépassent cinq ou six chiffres (cela varie selon les LLM) les erreurs redeviennent fréquentes. Cette méthode n'est pas suffisante pour qu'un LLM déduise de corpus d'entraînement les algorithmes généraux d'addition et de multiplication, et qu'il se mette à les appliquer de manière satisfaisante.

Geoffrey Hinton, spécialiste d'apprentissage profond, lauréat du prix Nobel de physique en 2024

« Il y a longtemps, en sciences cognitives, un débat opposait deux écoles de pensée. L'une, menée par Stephen Kosslyn, soutenait que la manipulation d'images dans l'esprit se résumait à un ensemble de pixels qu'on déplaçait. L'autre, plus proche de l'IA conventionnelle, disait : "Non, non, c'est absurde. Ce sont des descriptions hiérarchiques et structurelles. On a une structure symbolique dans l'esprit, et c'est elle qu'on manipule." Je pense qu'ils commettaient tous deux une erreur. [...] Ce qui se trouve dans le cerveau, ce sont de grands vecteurs correspondant à l'activité neuronale. [...] L'apprentissage profond finira par pouvoir tout faire. »

G. Hinton, Interview donnée à la *MIT Technology Review*, novembre 2020.

On a donc exploré d'autres idées. On peut, par exemple, faire en sorte que les *tokens* ne comportent qu'un seul chiffre, et ajouter des 0 pour que les nombres manipulés dans des calculs aient la même longueur. On peut aussi opérer une phase d'apprentissage supplémentaire en présentant un très grand nombre d'opérations correctes au réseau de neurones. Il est également possible d'associer à chaque chiffre d'un long nombre entier le numéro de sa position : 98247 devient ainsi (7,1)(4,2)(2,3)(8,4)(9,5).

Avec ces méthodes, quitte à les combiner, on décroche de meilleurs résultats. On obtient par exemple des systèmes capables de calculer correctement tous les produits et sommes de nombres entiers de moins d'une dizaine de chiffres. Cependant, jamais ces méthodes fidèles au principe de l'apprentissage profond ne semblent aboutir à une véritable maîtrise des algorithmes d'addition ou de multiplication entre entiers.

De manière assez remarquable, on observe très souvent un saut qualitatif lors de la phase d'apprentissage : une brusque transition entre le début de l'apprentissage, où les capacités de calcul sont médiocres, et un instant où le LLM devient soudainement beaucoup plus performant. Ces sauts qualitatifs correspondent à l'émergence d'une forme de compréhension des manipulations arithmétiques, malheureusement jamais vraiment suffisante pour des nombres de grande taille.

UN DÉBAT ANCIEN

Une méthode plus radicale est parfois utilisée pour traiter l'arithmétique élémentaire : faire appel à un outil de calcul extérieur. C'est la méthode exploitée par les LLM Gemini et Claude, qui de ce fait répondent mieux que leurs concurrents quand on leur demande d'effectuer des calculs ou des manipulations de tables de vérité. Claude peut en particulier calculer avec un grand nombre de chiffres exacts… mais il a pour cela recours à des algorithmes de calcul arithmétique qui gèrent les grands nombres et maintiennent la précision de chaque chiffre, même lorsque les résultats intermédiaires deviennent extrêmement grands. Ce n'est donc pas l'entraînement du LLM, la phase d'apprentissage profond, qui lui a permis d'acquérir des capacités de calcul satisfaisantes.

Ces calculs arithmétiques extérieurs, ou plus généralement les appels à des compétences ou informations ne résultant pas de l'entraînement

Amit Sheth, directeur de l'Institut d'intelligence de l'université de Caroline du Sud

« Les LLM ont démontré des capacités sans précédent à suivre des instructions en langage naturel exprimées par des invites. Ces modèles se révèlent capables d'un large éventail de tâches dès qu'on leur fournit des questions correctement formulées. [...] Les LLM peuvent générer des programmes informatiques à partir de la description de fonctions souhaitées ou rédiger des dissertations bien structurées sur des thèmes variés. [...] Malgré leurs performances impressionnantes, les LLM se heurtent à plusieurs défis. (1) Maîtriser des instructions complexes. (2) Éviter les incohérences et surmonter les ambiguïtés et les nuances contextuelles. (3) Arriver à généraliser des tâches qui s'écartent des exemples d'entraînement. (4) S'affranchir du manque de mécanismes de raisonnement explicites. »

A. Sheth, V. Pallagani et K. Roy, *IEEE Intelligent Systems*, 2024.

« Nous ne pouvons pas construire de modèles cognitifs riches de manière satisfaisante et automatisée sans le trio d'une architecture hybride, de riches connaissances préalables et des techniques de raisonnement perfectionnées. [...] Pour construire une approche robuste de l'IA fondée sur la connaissance, nous devons intégrer de la manipulation de symboles. Beaucoup de connaissances utiles sont abstraites, on ne peut donc pas se passer d'outils représentant et utilisant l'abstraction, et aujourd'hui la seule technique connue capable de manier de telles connaissances abstraites de manière sûre est la manipulation de symboles. »

G. Marcus et E. David, ***Rebooting AI : Building Artificial Intelligence We Can Trust,*** Pantheon Books, 2019.

Gary Marcus, professeur émérite de psychologie et de sciences neuronales à l'université de New York

du réseau de neurones, soulèvent une question fondamentale : est-il permis d'espérer que l'apprentissage profond seul, un jour, permette aux IA d'acquérir une maîtrise parfaite des opérations entre entiers – quitte à introduire de nouvelles idées dans les méthodes d'apprentissage ou à accroître encore la masse de données mobilisées pour l'entraînement des réseaux de neurones ? Ou l'appel à des outils extérieurs est-il inévitable pour pallier les insuffisances de la méthode connexionniste ?

Ce débat est bel et bien d'actualité chez les spécialistes d'IA, mais il n'est, en réalité, pas nouveau. Il est lié à la distinction, ancienne, opérée entre « IA neuronale » ou « connexionniste » et « IA symbolique » ou « logiciste », et à la façon dont on considère l'opposition ou la complémentarité entre ces deux types d'IA.

Selon l'approche connexionniste, il faut tenter d'imiter le fonctionnement du cerveau en concevant des neurones artificiels et en les associant sous la forme de réseaux, que l'on configure par des méthodes d'apprentissage pour qu'ils acquièrent certaines aptitudes simulant une forme d'intelligence. L'article fondateur à l'origine de cette approche date de 1943 et est cosigné par Warren McCulloch et Walter Pitts, et c'est dans la lignée des travaux qu'il a engendrés que sont nés les LLM.

L'approche symbolique, née dans les années 1950, propose quant à elle de se baser sur l'algorithmique et le raisonnement formel pour écrire des programmes manipulant des symboles, qui réussiront à simuler certaines capacités qualifiées d'intelligentes. L'IA symbolique est la seule, dans un premier temps, à obtenir des résultats significatifs : démonstrateurs de théorèmes, traducteurs automatiques de langue, logiciels de jeux, systèmes de calcul formel pour les mathématiques etc. Dans les années 1970 et 1980, elle conduit en particulier aux développements des « systèmes experts ». Ces outils logiciels fonctionnent en collectant de manière explicite auprès d'« experts » des connaissances exprimées sous forme de règles – par exemple, des énoncés du type « si *A* et *B* alors *C* ». Un algorithme, appelé « moteur d'inférences », exploite ensuite ces connaissances pour réaliser une tâche : opérer un diagnostic de panne, rédiger une prescription médicale, émettre des conseils en investissement, etc. Cette approche connaît quelques succès spectaculaires, en particulier dans le domaine médical où le célèbre système MYCIN, de l'université Stanford, réussit dès 1979 à égaler les meilleurs spécialistes humains des infections bactériennes. Un des grands succès de l'approche symbolique est aussi la victoire du programme Deep Blue sur le champion du monde en titre Garry Kasparov lors du fameux tournoi d'échecs de mai 1997.

LES SUCCÈS DU CONNEXIONNISME

Cependant, alors que jusqu'aux années 2000 la grande majorité des travaux en IA reposaient sur l'approche symbolique, les méthodes neuronales ont progressé. Ce sont elles qui permettent aujourd'hui aux ordinateurs de gagner contre les meilleurs experts humains non seulement aux échecs, mais aussi au jeu de go, sur lequel les approches purement symboliques et algorithmiques se cassaient les dents. Alors que pendant longtemps les techniques mises en œuvre pour faire de la traduction d'une langue à l'autre étaient de nature symbolique et fonctionnaient à base de dictionnaires, de règles de grammaire etc., elles ont été dépassées depuis une dizaine d'années par les méthodes statistiques et neuronales. Ce ne sont là que deux exemples d'une bascule plus générale entre IA symbolique et IA neuronale, dont l'avènement récent des LLM est une autre illustration.

Reste à savoir s'il est possible de se passer complètement de l'approche symbolique, ou si l'on doit au contraire chercher à la combiner avec l'IA neuronale. C'est un débat dont l'issue n'est pas tranchée aujourd'hui. Notons que, dans le cas des LLM, si l'on souhaite qu'ils aient la possibilité de manipuler des données comme la date et l'heure ou le cours actuel d'une certaine action boursière – ce qui est utile pour les

2 IA, SYSTÈME 1 ET SYSTÈME 2

L'IA neurosymbolique, résultat de l'association entre IA neuronale et IA symbolique, est à mettre en parallèle avec une théorie sur la manière dont fonctionne l'esprit humain. Selon Daniel Kahneman et d'autres psychologues, celui-ci dispose à la fois de mécanismes rapides – qualifiés parfois d'« intuitifs » – pour traiter les problèmes exigeant une réponse instantanée, et de mécanismes plus lents, faisant usage de mots et de symboles, qui correspondent à la pensée rationnelle. Le psychologue décrit en détail ces deux modes de fonctionnement en 2011 dans son livre *Système 1, système 2 : les deux vitesses de la pensée* (Flammarion).
L'IA neuronale serait l'analogue, pour les systèmes informatiques, du système 1, alors que l'IA symbolique y jouerait le rôle du système 2, responsable des raisonnements complexes et maîtrisés et garant de ce que produit l'autre IA.
Pour Daniel Kahneman, le système 1 fonctionne de manière automatique, involontaire, rapide et demande peu d'effort. C'est le système utilisé par défaut, car il est le moins coûteux en énergie pour l'humain. Grâce aux associations d'idées qu'il effectue, on lui doit notre faculté d'imagination et notre créativité. La lecture instantanée des expressions d'un visage humain résulte par exemple du système 1. C'est aussi lui qui nous permet de comprendre que, dans la phrase « Le professeur envoya le cancre chez le proviseur parce qu'il chahutait », le pronom « il » ne se réfère ni au professeur ni au proviseur.
Le système 2 exige en revanche une certaine concentration. Son fonctionnement analytique permet la résolution de problèmes structurés et complexes. Il est sensiblement plus lent que le système 1, et intervient par exemple quand ce dernier est confronté à un problème inattendu auquel il ne sait pas répondre – typiquement, un calcul – ou lorsqu'il est important de contrôler ce qu'il suggère.
Il correspond à une pensée maîtrisée, qui ne peut sans doute passe passer de symboles et de règles logiques précises.
Les défenseurs de l'IA neurosymbolique, en particulier le psychologue Gary Marcus, s'appuient sur cette analyse du fonctionnement de l'esprit humain pour défendre leur conception. Gary Marcus et Douglas Lenat, dans un article de 2023, résument en quelques mots cette idée : « Les humains possèdent des connaissances et des capacités de raisonnement qui s'apparentent au système 2 de Kahneman, mais l'IA générative actuelle, plus proche du système 1 rapide et automatique, n'en possède pas. Par conséquent, une partie importante de ce qui est évident pour l'homme reste peu fiable dans le cadre de l'approche des LLM. »

chatbots basés sur ces LLM –, l'usage de fonctionnalités extérieures au réseau de neurones est indispensable. D'ailleurs, donner à un LLM de tels accès revient simplement à le situer dans le monde réel, comme nous le sommes par nos corps et nos sens. Cependant, quand il s'agit de calculs arithmétiques ou de questions logiques et plus généralement mathématiques, il est *a priori* concevable que la méthode neuronale réussisse seule. Les tenants d'une IA «pure», fondée uniquement sur l'apprentissage profond, arguant des succès récents et spectaculaires de cette approche, ont longtemps espéré se passer d'appels à des fonctions spécialisées extérieures. Cette position a par exemple été celle du spécialiste de l'apprentissage profond Geoffrey Hinton, lauréat du prix Turing en 2018 et du prix Nobel de physique en 2024. Aujourd'hui, néanmoins, cet extrémisme neuronal se fait de plus en plus rare, sans doute à cause des difficultés persistantes que rencontre l'approche. Les propos tenus à ce sujet se sont faits plus prudents, et même si les spécialistes espèrent pouvoir aller le plus loin possible avec l'approche connexionniste, ils semblent presque tous accepter l'idée que le succès n'est pas assuré.

DE L'ESPOIR POUR L'IA SYMBOLIQUE

Le problème provient du fait que l'approche symbolique comporte certains avantages dont la contrepartie neuronale est difficile, voire peut-être impossible à trouver. Trois points cristallisent les avantages que l'approche symbolique détient aujourd'hui sur l'approche neuronale.

Tout d'abord, comprendre le fonctionnement des IA est bien plus facile en symbolique qu'en neuronal. En effet, on sait explicitement de quelles connaissances dispose une IA symbolique, ce qui permet d'identifier précisément ce qu'elle exploite pour fonctionner. Les IA à base d'apprentissage profond, à l'inverse, souffrent d'un cruel manque d'explicabilité. Le mécanisme d'un LLM, par exemple, reste essentiellement une boîte noire, dont les milliards de paramètres sont extrêmement difficiles à interpréter – même si, bien sûr, de nombreux travaux de recherche s'y consacrent.

De plus, la fiabilité des résultats donnés par une IA symbolique est souvent très bonne, puisqu'on sait ce qu'elle connaît et comment ses algorithmes traitent cette connaissance. La contrôler et l'ajuster est donc relativement facile. Au contraire, si une IA neuronale se trompe – on peut ici mentionner les fameuses et persistantes hallucinations des IA génératives – il est très délicat de la corriger. Autrement dit, le manque de fiabilité des IA génératives reste un grave problème mal résolu. À l'heure actuelle, personne ne sait, par exemple, comment corriger une IA génératrice d'images pour qu'elle représente correctement un reflet dans un miroir (*voir l'image ci-contre*) ; tout juste peut-on lui montrer des milliers d'images correctes de scènes se reflétant dans un miroir, en espérant que cela conduira le système à corriger ses erreurs.

Un troisième domaine dans lequel l'approche symbolique est bien plus satisfaisante que l'approche connexionniste est celui de la performance énergétique : l'électricité nécessaire à la conception des IA symboliques est tout à fait raisonnable, alors que l'entraînement d'un réseau de neurones est

L'IA neuronale, aussi appelée IA connexionniste, a obtenu de formidables succès ces dernières années. Elle produit cependant des systèmes dont la fiabilité est incertaine. Un exemple particulièrement simple et flagrant des limites de tels systèmes apparaît sur cette image générée avec Dall-E, un système d'IA générative d'images fondée sur un réseau de neurones artificiels. Le prompt utilisé pour créer cette image est : « Une femme se regarde dans une glace, réalisme photographique ». L'image est esthétiquement réussie, mais comporte un défaut étonnant : le reflet du personnage dans le miroir est erroné.

extrêmement énergivore. Aujourd'hui, on évalue l'électricité utilisée pour l'entraînement des LLM à environ 1,5 % de la consommation mondiale, et on anticipe que cette proportion doublera d'ici à 2030. Un rapport publié en 2025 par l'Agence internationale de l'énergie indique : « Un centre de calcul dévolu à l'IA consomme autant d'électricité que 100 000 familles, et ceux en construction aujourd'hui en consommeront vingt fois plus. » Certains chercheurs préconisent, sur la base de ce constat, de se passer totalement de l'IA générative à chaque fois que c'est possible.

L'APPROCHE NEUROSYMBOLIQUE

Face à ces constats, bien des spécialistes de l'apprentissage profond explorent activement des manières d'intégrer, dans les IA neuronales, des éléments symboliques, pour créer des systèmes d'IA performants et fiables. Ce domaine, fondé sur l'idée d'une complémentarité des deux approches, est celui de l'IA « neuro-symbolique ».

Gary Marcus, professeur émérite de psychologie et de sciences neuronales à l'université de New York, qui travaille à l'intersection entre les neurosciences et l'intelligence artificielle, défend depuis longtemps cette idée d'une nécessaire coopération. Selon lui, seule l'utilisation de symboles peut apporter des garanties à ce que produit le connexionnisme. Il évoque en particulier un parallèle entre la situation en IA et celle de l'esprit humain, où une distinction a été proposée par le psychologue Daniel Kahneman en 2016 entre un « système 1 » et un « système 2 » (*voir l'encadré 2*). De même, en IA, deux systèmes reposant sur des bases différentes seraient nécessaires.

De nombreux chercheurs abondent dans ce sens. C'est le cas, par exemple, d'Amit Sheth, Vishal Pallagani et Kaushik Roy, de l'université de Caroline du Sud, qui coécrivent dans un article de 2024 : « L'approche neurosymbolique améliore la fiabilité et la prise en compte du contexte de l'exécution des tâches, permettant aux LLM d'interpréter et de répondre de manière dynamique à un plus large éventail de contextes avec plus de précisions et de flexibilité. »

Cette troisième voie sera peut-être la solution pour pallier les difficultés des IA génératives dans la manipulation de notions mathématiques. On ignore, en revanche, si les progrès qui en découleront pourront véritablement servir la recherche en mathématiques. Les IA génératives ont d'ores et déjà servi à la démonstration de théorèmes intéressants – nous en donnions un exemple dans cette rubrique en juillet 2024. Elles sont, cependant, encore loin de pouvoir se substituer aux méthodes classiques de démonstration automatique et de validation des preuves qu'autorisent les assistants de preuves. ■

BIBLIOGRAPHIE

J. Loeber, Everything we know about LLMs doing arithmetic, billet de blog, 2024.

A. Sheth *et al.*, Neurosymbolic AI for enhancing instructability in generative AI, *IEEE Intelligent Systems*, 2024.

D. Lenat et G. Marcus, Getting from generative AI to trustworthy AI: What LLMs might learn from Cyc, *arXiv preprint*, 2023.

A. Sheth *et al.*, Neurosymbolic AI – Why, what, and how, *arXiv preprint*, 2023.

G. Marcus et E. Davis, *Rebooting AI: Building Artificial Intelligence We Can Trust*, Pantheon Books, 2019.

M. Waldrop, What are the limits of deep learning ?, *PNAS*, 2019.

D. Kahneman, *Système 1, système 2 : les deux vitesses de la pensée*, Flammarion, 2012.