

Quand les robots seront-ils doués de bon sens ?

Publié le 13 septembre 2022

Christopher Gannatti, CFA

Global Head of Research

En 2022, nous avons entendu parler de plusieurs applications d'intelligence artificielle qui renforcent constamment différentes sortes de capacités :

- Les grands modèles de langage, illustrés par GPT-3.5, ont pris de l'ampleur et ont investi leurs capacités dans un nombre croissant de domaines, à l'image des langages de programmation informatique.
- DeepMind a étoffé sa boîte à outils AlphaFold en présentant les prévisions sur la structure de plus de 200 millions de protéines et en mettant ces prévisions gratuitement à la disposition des chercheurs².
- Nous avons même observé une expansion de l'« autoML »³, qui se rapporte aux outils d'apprentissage automatique (« machine learning ») 'low-code' capables de fournir à un plus grand nombre de personnes, dénuées d'expertise de la science des données ou des sciences informatiques, un accès à l'apprentissage automatique⁴.

Mais même si nous reconnaissons que des progrès sont en cours, les machines sont encore essentiellement utiles pour des tâches discrètes et disposent de peu de flexibilité pour réagir rapidement dans différentes situations.

À l'intersection des grands modèles de langage et de la robotique

Les grands modèles de langage sont intéressants dans de nombreux cas pour leurs propriétés émergentes. Ces modèles gigantesques ont des centaines de milliards, voire des milliers de milliards de paramètres. Ils peuvent produire comme résultat du texte écrit ou du code auto-rempli dans des applications de codage.

Or qu'en est-il si vous dites à un robot que vous avez faim par exemple.

En tant qu'être humain, si nous entendons quelqu'un nous dire qu'il a faim, nous pouvons intuitivement penser assez rapidement à différentes solutions autour de nous. À une certaine heure du jour, nous penserons peut-être aller au restaurant. Nous pouvons prendre notre smartphone et utiliser à la vente à emporter ou la livraison à domicile. Peut-être que nous commencerons à préparer un repas.

Un robot n'aurait pas nécessairement cette conscience de la situation s'il n'a pas été programmé à l'avance. Nous pensons naturellement aux robots capables d'exercer leurs fonctions spécifiques programmées à l'avance afin d'effectuer des tâches précises. Nous pensons peut-être à un robot capable d'agir en fonction d'une série d'instructions très simples : lui dire où se déplacer en utilisant certains mots clés, ce qu'il doit faire en ajoutant certains autres mots clés.

« J'ai faim » : une commande composée de deux mots sans instructions inhérentes serait supposée être impossible à appliquer.

Le Pathways Language Model (PaLM) de Google - le début d'interactions plus complexes entre les êtres humains et les robots

Les chercheurs de Google ont été capables de faire la démonstration d'un robot capable de répondre, dans un environnement clos il faut l'admettre, à la phrase « j'ai faim ». Il a été en mesure de localiser la nourriture, de s'en saisir et de la tendre à l'être humain⁵.

Le modèle PaLM de Google se cachait derrière la capacité du robot à prendre en compte les données de langage et à les traduire en actions. PaLM mérite d'être mentionné car il intègre la capacité d'expliquer en langage naturel comment il parvient à certaines conclusions⁶.

Comme c'est souvent le cas, les résultats les plus dynamiques sont généralement obtenus lorsque l'on est capable de conjuguer différentes méthodes d'apprentissage pour parvenir à des capacités supérieures. Bien entendu, PaLM ne peut pas en soit informer automatiquement le robot sur la manière d'attraper physiquement une barre de chocolat, par exemple. Les chercheurs démontreront grâce à un contrôle à distance comment effectuer certaines tâches. Mais PaLM a été utile car il a permis au robot de faire le lien entre ces actions concrètes apprises et les déclarations relativement abstraites faites par des êtres humains, comme la phrase « j'ai faim » qui n'a pas nécessairement de commande explicite⁷.

Les chercheurs de Google et Everyday Robots ont intitulé un article (« Fais comme je peux, pas comme je dis : les bases du langage de la robotique actuelle » (« Do As I Can, Not As I Say: Grounding Language in Robotic Affordance's ») ⁸. Dans le schéma 1, nous constatons le génie qui se cache derrière le titre de cet article. En effet, il est important de reconnaître que les grands modèles de langage peuvent récupérer leur texte 'inspiré' sur Internet, dont la plupart serait inapplicable à un robot particulier dans une situation donnée. Le système doit 'trouver l'intersection' entre ce qu'il est logique de faire d'après le modèle de langage et ce que le robot est capable de faire dans le monde physique. Par exemple :

- Différents modèles de langage peuvent associer le nettoyage d'un déversement de liquide avec toutes sortes de nettoyage. Ces modèles risquent de ne pas être capables d'utiliser leur extraordinaire formation pour se rendre compte qu'un aspirateur n'est peut-être pas la meilleure façon de nettoyer un liquide. Les robots peuvent également exprimer simplement leur regret qu'un déversement se soit produit.
- Si nous réfléchissons à l'intersection qui a le plus de probabilité d'avoir du sens, si un robot peut dans une situation donnée 'trouver une éponge' et si le grand modèle de langage indique que la réponse 'trouver une éponge' a du sens, conjuguer ces deux concepts peut permettre au robot d'essayer d'effectuer au moins une action corrective productive face à ce déversement.

Le modèle 'SayCan', s'il n'est certainement pas parfait et s'il n'est pas un substitut à une compréhension véritable, est un moyen intéressant de faire exécuter à des robots des actions qui ont du sens dans une situation sans être directement programmés pour répondre à une affirmation de cette manière précise.



Source : Graphique de Ahn et al. « Do As I Can, Not As I Say: Grounding Language in Robotic Affordances. » ARXIV. Soumis le 4 avril 2022, dernière révision le 16 août 2022.

Dans un certain sens, il s'agit de la partie la plus passionnante de ce domaine de recherche :

- Les robots ont tendance à avoir besoin de commandes courtes codées en dur. Comprendre davantage des instructions moins spécifiques n'est pas généralement possible.
- Les grands modèles de langage ont démontré leur capacité extraordinaire à réagir à différentes instructions, mais c'est toujours dans un contexte 'exclusivement numérique'.

Si la force des robots dans le monde physique peut être conjuguée avec leur capacité, au moins apparente, de comprendre le langage naturel qui provient de grands modèles de langage, vous avez là l'occasion d'une synergie notable qui est meilleure qu'un fonctionnement autonome.

Conclusion : les entreprises en quête de capacités robotiques

Au sein du secteur de l'intelligence artificielle, il est important de reconnaître la progression critique du concept à la recherche et à l'innovation et plus tard seulement à une utilisation sur le marché du grand public et à la rentabilité (ce qui reste à espérer). Des robots comprenant le langage naturel abstrait sont loin d'être une activité génératrice de revenus sur le marché du grand public.

Pourtant, nous constatons que des entreprises prennent des mesures en faveur d'une utilisation toujours croissante de la robotique. Amazon est souvent sous le feu des projecteurs en raison de son utilisation possible de robots dans ses centres de distribution, mais plus récemment, le groupe a annoncé son intention d'acquérir iRobot, le fabricant du système d'aspirateur Roomba. À l'avenir, les robots aux capacités de plus en plus sophistiquées auront un rôle à jouer dans la société.

Dans l'environnement actuel marqué par des pressions croissantes sur les salaires, les entreprises étudient de plus en plus ce que les robots et l'automatisation peuvent apporter à leurs activités. Il est important de ne pas surestimer où nous en sommes en termes d'avancées en 2022. Les robots ne sont pas en mesure de reproduire des comportements à 100 % humains. Mais nous devrions assister à des progrès remarquables dans les prochaines années.

Sources

1 Generative Pre-trained Transformer 3

2 Source : Ewen Callaway. « 'The Entire Protein Universe': AI Predicts Shape of Nearly Every Known Protein. » Nature. Volume 608. 4 août 2022.

3 Apprentissage automatique automatisé (« Automated machine learning »)

4 Source : Tammy Xu. « Automated techniques could make it easier to develop AI. » MIT Technology Review. 5 août 2022.

5 Source : Will Knight. « Google's New Robot Learned to Take Orders by Scraping the Web. » WIRED. 16 août 2022.

6 Source : M. Knight, 16 août 2022.

7 Source : M. Knight, 16 août 2022.

8 Source : Ahn et al. « Do As I Can, Not As I Say: Grounding Language in Robotic Affordances. » ARXIV. Soumis le 4 avril 2022, dernière révision le 16 août 2022.

9 Source: Hart, Connor. "Amazon Buying Roomba Maker iRobot for \$1.7 Billion." Wall Street Journal. 5 August 2022.

Important Risks Related to this Article

Informations importantes

Communications commerciales publiées dans l'EEE Ce document est publié et approuvé par WisdomTree Ireland Limited, une société autorisée et réglementée par la Central Bank of Ireland.

Communications commerciales émises dans des juridictions en dehors de l'EEE Ce document est publié et approuvé par WisdomTree UK Limited, une société autorisée et réglementée par la Financial Conduct Authority du Royaume-Uni.

WisdomTree Ireland Limited et WisdomTree UK Limited sont toutes les deux désignées comme « WisdomTree » (le cas échéant). Notre Politique sur les conflits d'intérêts et notre Inventaire sont disponibles sur demande.

Réservé aux clients professionnels uniquement. Les informations figurant dans ce document sont fournies à titre informatif et ne constituent pas une ore de vente, ou une sollicitation d'achat de titres ou d'actions. Ce document ne doit pas être utilisé comme fondement d'une décision d'investissement. La valeur des investissements peut fluctuer et vous êtes susceptible de perte tout ou partie du montant investi. La performance passée ne constitue pas nécessairement une indication des performances futures. Toute décision d'investissement doit être fondée sur les informations figurant dans le prospectus approprié et sur des conseils indépendants en matière d'investissement, fiscaux et juridiques.

L'application des réglementations et lois fiscales peut souvent conduire à des interprétations diérentes. Tous les points de vue ou opinions exprimés dans cette communication représentent les points de vue de WisdomTree et ne doivent pas être interprétés comme des conseils réglementaires, fiscaux ou juridiques. WisdomTree ne donne aucune garantie ou représentation quant à l'exactitude des vues ou opinions exprimées dans cette communication. Toute décision d'investissement doit être fondée sur les informations contenues dans le prospectus approprié et après avoir sollicité des conseils indépendants en matière d'investissement, fiscaux et juridiques. Ce document n'est pas et ne doit en aucun cas être interprété comme une publicité ou une ore publique d'actions ou de titres aux États-Unis ou dans toute province ou tout territoire des États-Unis. L'introduction, la transmission et la distribution (directes ou indirectes) de l'original ou d'une copie de ce document sont interdites aux États-Unis.

Bien que WisdomTree s'efforce d'assurer l'exactitude du contenu de ce document, WisdomTree ne peut en garantir l'exactitude. Les fournisseurs de données tiers sollicités pour obtenir les informations contenues dans le présent document ne donnent aucune garantie ou représentation de quelque sorte en rapport avec ces données. Lorsque WisdomTree exprime ses propres opinions concernant le produit ou l'activité du marché, ces opinions sont susceptibles de changer. WisdomTree, ses alliés et leurs dirigeants, directeurs, partenaires ou employés respectifs déclinent toute responsabilité pour toute perte directe ou indirecte découlant de l'utilisation de ce document ou de son contenu.