

IA - théorème de Gödel - Roger Penrose

<https://intelligence-artificielle.developpez.com/actu/369727/-Le-theoreme-de-Godel-brise-le-mythe-le-plus-important-de-l-IA-L-IA-ne-sera-jamais-consciente-declare-le-mathematicien-Roger-Penrose-il-ajoute-que-le-nom-intelligence-artificielle-n-est-pas-le-bon/>

« Le théorème de Gödel brise le mythe le plus important de l'IA. L'IA ne sera jamais consciente », déclare le mathématicien Roger Penrose

il ajoute que le nom « intelligence artificielle » n'est pas le bon

Roger Penrose, lauréat du prix Nobel de physique, affirme que le théorème de Gödel remet en cause le mythe selon lequel l'IA est consciente ou le sera un jour. Selon lui, « la conscience n'est pas calculable » et l'IA, telle que nous la connaissons, ne parviendra jamais à une véritable intelligence. Il ajoute qu'il est important de comprendre la physique sous-jacente de la conscience, qui, selon lui, n'est pas calculable et implique la réalité quantique. Il réfute ainsi les déclarations selon lesquelles l'IA commence à montrer des signes de conscience. En outre, il a souligné que le nom « intelligence artificielle » n'est pas le bon, car il ne s'agit pas d'intelligence.

L'IA sera-t-elle un jour consciente ? Roger Penrose répond par la négative

Sir Roger Penrose, né le 8 août 1931 à Colchester, en Angleterre, est mathématicien, physicien et philosophe des sciences britannique. Il est professeur émérite de mathématiques à l'Université d'Oxford et est reconnu pour ses contributions majeures à la relativité générale et à la cosmologie. En 2020, il a reçu le prix Nobel de physique pour avoir découvert que la formation des trous noirs est une prédiction de la théorie générale de la relativité d'Albert Einstein.

Roger Penrose s'est également intéressé à la nature de la conscience humaine. Dans son livre "The Emperor's New Mind", il affirme que la conscience ne peut pas être expliquée uniquement par des processus computationnels et suggère que des phénomènes quantiques pourraient jouer un rôle dans l'activité cérébrale.

Lors d'une récente interview, Roger Penrose est revenu sur le sujet de la conscience dans le contexte de l'essor de l'IA. Il s'oppose d'emblée aux déclarations de certains leaders de l'industrie selon lesquels l'IA a une « conscience de soi ». Selon Roger Penrose, le théorème de Gödel détruit ce mythe.

Le théorème d'incomplétude de Gödel est un résultat fondamental de la logique mathématique qui dit que tout système logique suffisamment puissant admet nécessairement un énoncé qu'il ne peut ni démontrer ni réfuter. Le théorème a été publié en 1931 par Kurt Gödel, logicien et mathématicien autrichien.

Roger Penrose

Il ne s'agit pas d'intelligence artificielle. L'intelligence impliquerait la conscience. D'après l'argument de Gödel, vous voyez, c'est l'intrigue qui, je pense, a été perdue. Un ordinateur est un type très spécifique de structure mathématique. Il s'agit de mathématiques computationnelles. C'est une partie très limitée des mathématiques. Je ne pense pas qu'ils deviendront un jour conscients du fait qu'ils sont des ordinateurs au sens actuel du terme.

Mais je pense que la conscience n'est pas computationnelle. C'est toujours le cas avec l'IA. Si vous parlez à une IA, elle ne sait pas ce qu'elle fait. Les ordinateurs sont devenus si puissants qu'ils ne sa-

vent plus ce qu'ils font. Je pense que l'essentiel a été perdu. Je pense que Turing n'était pas si loin de la vérité. Je pense que les gens ont perdu le fil.

Le théorème de Gödel démontre qu'il existe des vérités mathématiques qui ne peuvent être prouvées par aucun ensemble de règles formelles. Selon Roger Penrose, cela implique que la conscience transcende le calcul, car elle implique de comprendre les raisons derrière les règles, et pas seulement de les suivre.

Il a déclaré que la conscience est un phénomène physique enraciné dans la physique non calculable, impliquant éventuellement la réalité quantique. Par conséquent, selon Roger Penrose, l'IA, qui s'appuie sur des règles de calcul, ne pourra jamais atteindre une véritable conscience.

Roger Penrose affirme que le nom « intelligence artificielle » n'est pas le bon

Outre les déclarations sur une éventuelle conscience de l'IA, Roger Penrose n'est pas non plus d'accord avec l'appellation « intelligence artificielle ». Selon le scientifique britannique, il ne s'agit en aucun cas d'intelligence, car l'intelligence implique la conscience. « Le nom n'est pas le bon. Il ne s'agit pas d'une intelligence artificielle. Ce n'est pas de l'intelligence. L'intelligence implique la conscience », a expliqué Roger Penrose lors de son interview.

Il a ajouté : « j'ai toujours été un fervent défenseur de l'idée que ces appareils ne sont pas conscients. Et ils ne seront pas conscients à moins qu'ils n'apportent d'autres idées. Ce sont toutes des notions calculables. Mais l'intelligence n'est pas la conscience. Elle dépend de la conscience. Vous voyez, cela dépend de la façon dont vous utilisez le mot. Je pense que lorsque les gens utilisent le mot "intelligence", ils font allusion à quelque chose qui est conscient ».

Roger Penrose

Il peut y avoir des choses complètement inconnues dont nous n'avons aucun concept pour le moment. C'est peut-être vrai. Je parle de choses... Vous voyez, si vous parlez d'IA, cela signifie, malheureusement, que c'est un mauvais terme parce qu'on parle d'intelligence artificielle. Selon moi, quand les gens utilisent ce terme, ils veulent dire des choses qui peuvent être faites par des ordinateurs. C'est ce qu'ils veulent dire. Ils ont des ordinateurs et leurs ordinateurs sont très puissants.

Et ce qu'ils entendent par IA, c'est l'utilisation d'ordinateurs. Il se peut qu'elle ne soit pas utilisée de manière tout à fait logique. Je n'ai pas étudié ces choses, donc je ne suis probablement pas à jour. Mais d'après ce que je vois, il s'agit d'examiner des masses de données et de rechercher des modèles ou d'essayer de les analyser de cette manière. Mais tout cela est informatique. Vous utilisez des ordinateurs.

Roger Penrose a déclaré que l'un des risques liés au développement de l'IA est que les gens pensent qu'il y a une conscience. Les déclarations de Roger Penrose ont suscité des réactions mitigées. Certains critiques affirment que l'IA n'est qu'un « algorithme » et n'est donc pas dotée de conscience.

D'autres rejettent les affirmations de Roger Penrose. « La plupart des théories de calculabilité ne s'appliquent pas à l'IA, y compris le théorème de Gödel. Les gens commencent par se tromper en parlant d'"algorithmes" d'IA. Ce terme s'applique plutôt à des concepts tels que la descente de gradient », note un critique.

« Mais les inférences des réseaux neuronaux qui en résultent ne sont pas un algorithme. Il ne s'agit pas d'une séquence définie d'opérations produisant un résultat défini. C'est comme une heuristique,

une procédure qui se rapproche d'un résultat correct, mais qui ne fournit aucune garantie mathématique ».

La conscience de l'IA est un concept à la fois fascinant et problématique. D'un côté, les défenseurs de la théorie affirment que doter les machines d'une conscience de soi pourrait les rendre plus empathiques, plus capables d'interagir avec les humains de manière authentique et de comprendre le monde qui les entoure. Cependant, il y a plusieurs raisons de considérer cela comme une entreprise potentiellement dangereuse et fondamentalement futile.

...