

L'ESSENTIEL

> Une machine soumise au test de Turing tente, par ses réponses, de convaincre un interrogateur qu'elle est un humain.

> Ce test a longtemps été considéré comme la meilleure façon de jauger une intelligence artificielle.

> Le test de Turing a cependant mal vieilli. Le réussir est davantage affaire de tromperie que de véritable intelligence. Pour les experts, le test de Turing est à remplacer par une batterie d'épreuves qui évalueraient l'intelligence de la machine sous différents angles.

> Une machine réellement intelligente devrait pouvoir comprendre des affirmations ambiguës, assembler un meuble vendu en pièces détachées, réussir un examen scolaire... Étant donné la difficulté de ces tâches, l'homme n'est pas près d'être détrôné sur le plan de l'intellect.

L'AUTEUR



GARY MARCUS
directeur d'Uber AI Labs
et professeur de psychologie et
de neurosciences à l'université
de New-York, aux États-Unis

Machine, es-tu intelligente?

ON A LONGTEMPS CHERCHÉ LA RÉPONSE À CETTE QUESTION À L'AIDE DU FAMEUX TEST DE TURING. MAIS CE TEST, IMAGINÉ EN 1950, A MONTRÉ SES LIMITES. PAR QUOI LE REMPLACER ? PAR DES BATTERIES D'ÉPREUVES DE DIFFÉRENTES NATURES, DISENT LES SPÉCIALISTES.

En 1950, le mathématicien britannique Alan Turing imagine une expérience de pensée devenue célèbre et souvent considérée comme le meilleur moyen de cerner l'intelligence d'une machine. Il l'appela le « jeu de l'imitation », mais on la connaît généralement sous le nom de « test de Turing ».

Anticipant les *chat bots* d'aujourd'hui, ou agents conversationnels, des programmes informatiques qui se font passer pour des humains, Turing envisageait un échange textuel au cours duquel une machine tenterait de tromper un examinateur et de l'amener à penser qu'elle est un humain, en répondant à des questions relatives à de la poésie et en commettant délibérément des erreurs de calcul.

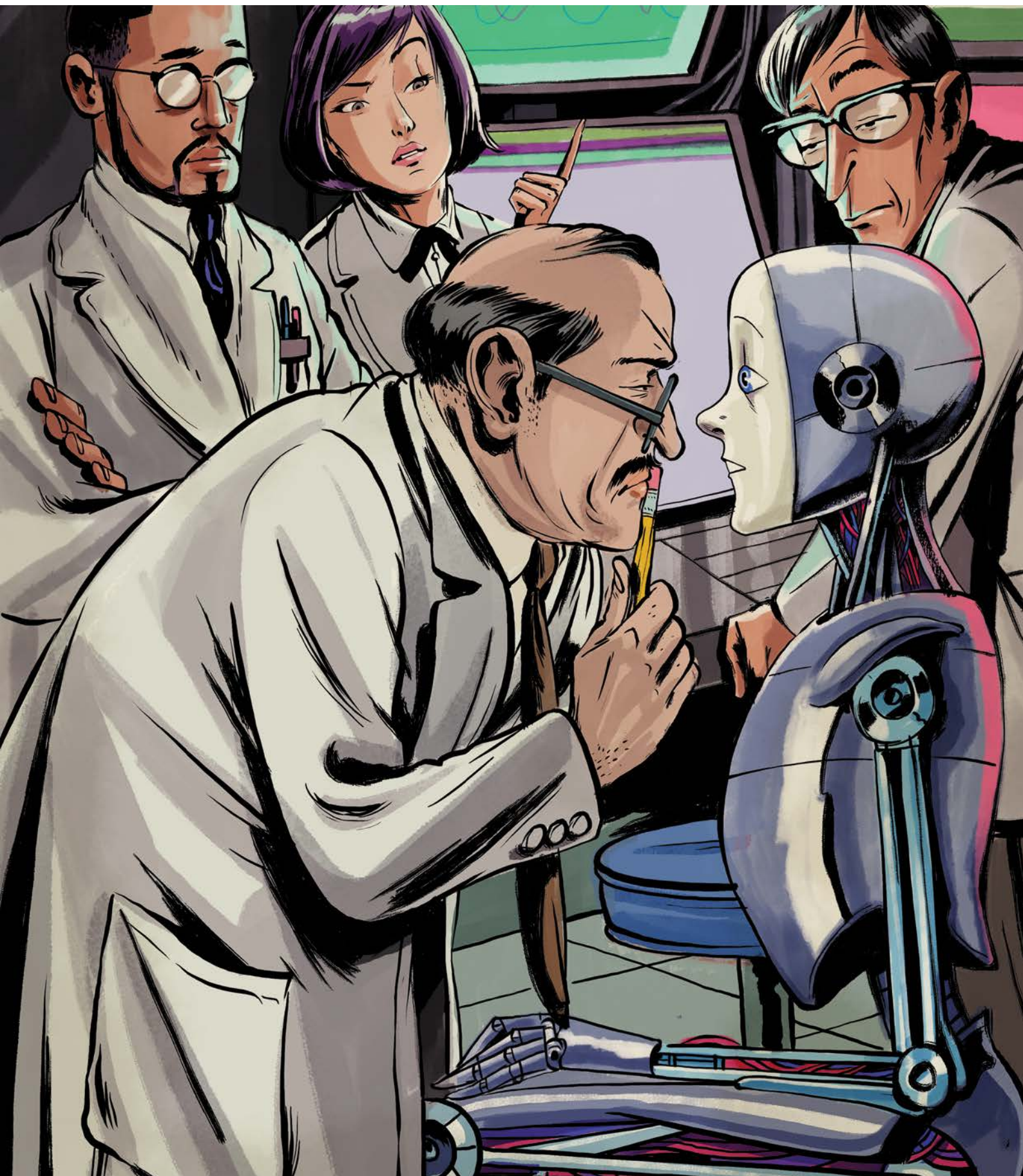
De nos jours, le grand public voit souvent le test de Turing comme une sorte de Rubicon, dont le franchissement signifierait que la machine est devenue l'égale de l'homme, du moins sur le plan de son intelligence. C'est un tort : on peut franchir ce Rubicon en trichant. On sait aujourd'hui construire des machines qui

dupent les gens, au moins sur de courtes durées, mais ces victoires sont fugaces et ne semblent pas du tout nous rapprocher de ce qu'est une véritable intelligence.

UNE MACHINE PEUT PASSER LE TEST DE TURING EN DUPANT

Le problème avec le test de Turing est qu'il est trop facile à dénaturer. On peut « gagner » en pratiquant la tromperie ou en feignant l'ignorance, tout simplement. Turing lui-même l'avait en partie pressenti : le moyen de gagner consiste essentiellement à ne pas répondre aux questions. Par exemple, si un examinateur demande à une machine : « Avez-vous faim ? », la machine doit mentir pour gagner. Cette méthode détournée n'est pas une exception, c'est la règle : pour qu'une machine passe le test de Turing, il faut qu'elle trompe son interlocuteur.

Par exemple, en 2014, un agent conversationnel russe nommé Eugene Goostman a prétendument gagné au test de Turing en se faisant passer pour un garçon sarcastique d'Odessa (Ukraine), âgé de 13 ans, qui ne répondait jamais directement aux questions. Certains examinateurs se >



> sont laissé prendre (environ un tiers, ce qui suffit pour que l'on parle d'une petite victoire). Mais qu'est-ce que cela prouve? Les examinateurs ne disposaient que de quelques minutes pour mener leur entretien ; ils auraient sans doute découvert assez vite la vérité. En fait, Eugene ne savait (et ne comprenait) presque rien. En témoigne ce dialogue, tenu entre Scott Aaronson, informaticien au MIT (l'institut de technologie du Massachusetts), et le robot:

SCOTT: Qu'est-ce qui est plus grand, une boîte de chaussures ou l'Everest?

EUGENE : Je ne peux pas choisir maintenant. Je vais y réfléchir plus tard. Et j'ai oublié de vous demander d'où vous êtes...

SCOTT: Combien un chameau a-t-il de pattes?

EUGENE: Entre deux et quatre. Trois, peut-être? :-))) Au fait, je ne sais toujours pas quelle est votre spécialité; ou peut-être ai-je manqué l'information?

Combien le chameau a-t-il de pattes? Entre deux et quatre

SCOTT: Combien un mille-pattes a-t-il de pattes?

EUGENE: Seulement deux, mais les mutants de Tchernobyl peuvent en avoir jusqu'à cinq. Je sais que vous êtes censé me duper.

SCOTT: Non, j'ai besoin de savoir que vous n'êtes pas un robot. Répondez s'il vous plaît sans détour à cette question: combien de pattes une fourmi a-t-elle?

EUGENE : Entre deux et quatre. Trois, peut-être? :-))) Oh, quelle conversation constructive;-).

Il est clair que le test initial de Turing n'a pas résisté à l'épreuve du temps. Peu après que les médias ont découvert l'histoire d'Eugene Goostman, j'ai suggéré un autre test, censé mettre en avant la véritable intelligence plutôt que des échappatoires douces. Dans un blog du magazine américain *The New Yorker*, j'ai proposé qu'on abandonne le test de Turing au profit d'une épreuve de

Test 1

LE DÉFI DU SCHÉMA DE WINOGRAD

Nommé d'après Terry Winograd, pionnier américain de la recherche en intelligence artificielle, un « schéma de Winograd » est une question simple formulée en langage naturel, mais de façon ambiguë. Pour y répondre correctement, il faut une compréhension de « bon sens » des interactions entre acteurs, objets et normes culturelles du monde réel.

Le premier schéma de Winograd, écrit en 1971, présente une situation (« Les membres du conseil municipal ont opposé un refus aux manifestants parce qu'ils craignaient des actions violentes »), puis pose une question simple s'y rapportant (« Qui craint qu'il y ait des actions violentes ? »). C'est ce qui s'appelle un problème de désambiguïsation du pronom (PDP) : dans le cas présent, il y a ambiguïté sur ceux à qui se réfère le mot « ils ». Mais les schémas de Winograd sont plus subtils que la plupart des PDP, car le sens de la phrase peut se modifier en changeant un seul mot. Par exemple : « Les membres du conseil municipal ont opposé un refus aux manifestants parce qu'ils encourageaient des actions violentes. »

La plupart des gens utilisent

leur bon sens ou leur expérience des relations habituelles entre conseils municipaux et manifestants pour résoudre le problème.

Ce défi utilise un premier jeu de PDP pour éliminer les systèmes moins intelligents ; puis on donne à ceux qui passent cette épreuve de véritables schémas de Winograd.

POUR : Comme les schémas de Winograd s'appuient sur des connaissances auxquelles les ordinateurs n'accèdent pas de façon sûre, le défi résiste bien à Google, c'est-à-dire qu'il ne se résout pas en recherchant sur Internet.

CONTRE : Les schémas utilisables sont relativement peu nombreux. « Ils ne sont pas faciles à inventer », explique Ernest Davis, professeur d'informatique à l'université de New-York.

NIVEAU DE DIFFICULTÉ : Élevé. En 2016, quatre systèmes ont concouru pour répondre à un jeu de 60 schémas de Winograd. Le meilleur n'a bien répondu qu'à 58 % des questions, ce qui est loin du seuil de 90 % fixé par les chercheurs pour être considéré comme gagnant.

À QUOI CELA SERT : À faire le tri entre compréhension et simulation. « Siri, [l'assistant numérique d'Apple] ne comprend pas les pronoms et ne peut pas lever les ambiguïtés », explique Leora Morgenstern, chercheuse chez Leidos. Ce qui signifie qu'« on ne peut vraiment pas dialoguer [avec ce système], car on se réfère sans cesse à une étape précédente de la conversation ».

JOHN PAVLUS, journaliste

