



¿Será consciente alguna vez la IA?

Ingeniería, 28/04/2021



A medida que la sociedad se acerca a la IA a nivel humano, los científicos debaten qué significa existir.

Imagínese que se somete a un procedimiento en el que cada neurona de su cerebro se reemplaza gradualmente por componentes electrónicos funcionalmente equivalentes. Digamos que el reemplazo ocurre una sola neurona a la vez, y que en el comportamiento, nada cambia en usted. Desde fuera, sigues siendo "tú", incluso para tus amigos más cercanos y seres queridos.

¿Qué le pasaría a tu conciencia?
¿Desaparecería gradualmente, una neurona a la vez? ¿Dejaría de existir repentinamente después del reemplazo de alguna partícula crítica para la conciencia en su corteza posterior? ¿O simplemente seguiría siendo usted, plenamente consciente de su experiencia vivida y de su

sensibilidad (y complacido o horrorizado de que su mente teóricamente pueda conservarse para siempre)?

Este famoso experimento de pensamiento sobre la conciencia, propuesto por el filósofo David Chalmers en su artículo de 1995 *Absent Qualia, Fading Qualia, Dancing Qualia*, plantea casi todas las cuestiones importantes que existen en el debate en torno a la posibilidad de la conciencia en la inteligencia artificial.

Si la perspectiva de comprender los orígenes de nuestra propia conciencia y la de otras especies es, como te dirán todas las personas que la estudian, desalentadora, replicarla en máquinas es ambiciosa hasta un grado absurdo.

¿Será la IA consciente alguna vez? Al igual que con todas las cosas relacionadas con la conciencia, la respuesta es que nadie sabe realmente en este punto, y muchos piensan que puede ser objetivamente imposible para nosotros entender si el

fenómeno resbaladizo aparece alguna vez en una máquina.

Tomemos el experimento mental que acabamos de describir. Si la conciencia es una característica única de los sistemas biológicos, incluso si el reemplazo robótico de su cerebro le permitiera funcionar exactamente de la misma manera que lo hacía antes del procedimiento, no habría nadie en casa en el interior y usted sería un caparazón zombie-esque de tu antiguo yo. Los más cercanos a usted tendrían todas las razones para tomar su conciencia como un hecho, pero estarían equivocados.

La posibilidad de que podamos inferir erróneamente la conciencia sobre la base del comportamiento externo no es una proposición absurda. Es concebible que, una vez que logremos construir inteligencia artificial general, del tipo que no es estrecho como todo lo que existe ahora mismo, que pueda adaptarse, aprender y aplicarse en una amplia gama de contextos, la tecnología se sentirá consciente para nosotros. independientemente de si realmente lo es o no.

Imagine una especie de Alexa o Siri con esteroides, un programa con el que puede conversar, que es tan hábil como cualquier ser humano para comunicarse con una entonación variada y un ingenio creativo. La línea se difumina rápidamente.

Dicho esto, puede que no sea necesario, deseable o incluso posible que la IA alcance o presente algún tipo de conciencia.

En *Life 3.0: Being Human in the Age of Artificial Intelligence*, Max Tegmark, profesor de física en el MIT y presidente del Future of Life Institute, lamenta: "Si le mencionas la" palabra C "a un investigador de IA, neurocientífico o psicólogo, es posible que pongan los ojos en blanco. Si son su mentor, podrían compadecerse de usted y tratar de convencerlo de que no pierda el tiempo en lo que consideran un problema desesperado y poco científico ".

Es una posición razonable, si no un poco desdenosa, a tomar. ¿Por qué molestarse siquiera en tomar en cuenta el problema de la conciencia? Los titanes de la tecnología como Google e IBM ya han logrado avances impresionantes en la creación de algoritmos de autoaprendizaje que pueden pensar y superar (aunque en circunstancias estrictamente definidas) cualquier cerebro humano, y los programas de aprendizaje profundo en el campo de la medicina también lo están. médicos con mejor desempeño en algunas áreas de identificación de tumores y evaluación de análisis de sangre. Estas tecnologías, aunque no son perfectas, funcionan bien y solo mejoran en lo que hacen.

Douglas Hofstadter, el fenomenólogo pionero que escribió *Gödel, Escher, Bach: An Eternal Golden Braid*, ganador del premio Pulitzer, se encuentra entre los que piensan que es absolutamente necesario molestarnos, y por una buena razón.

En una entrevista de 2013 con *The Atlantic*, Hofstadter explica su creencia de que, en gran medida, estamos perdiendo el punto si no tomamos en cuenta aspectos como la naturaleza de la inteligencia consciente. Haciendo referencia a *Deep Blue*, el famoso programa de ajedrez desarrollado por IBM que venció a Gary Kasparov en 1997, dice: "Está bien, [...] *Deep Blue* juega muy bien al ajedrez, ¿y qué? ¿Eso te dice algo sobre cómo jugamos al ajedrez? No. ¿Te dice cómo Kasparov concibe y entiende un tablero de ajedrez?"

La perspectiva de Hofstadter es fundamental. Si estos algoritmos hipercapaces no se construyen con una comprensión adecuada de nuestras propias mentes que les informan, una comprensión que todavía es muy incipiente, ¿cómo podríamos saber si alcanzan la inteligencia consciente? Más urgentemente, sin una comprensión clara del fenómeno de la conciencia, ¿la carga hacia el futuro con esta tecnología creará más problemas de los que resuelve?

En *Inteligencia artificial: una guía para pensar en humanos*, Melanie Mitchell, una ex estudiante de posgrado de Hofstadter, describe el miedo al desarrollo imprudente de la inteligencia artificial que su mentora expresó una vez en una sala llena de ingenieros de Google en una reunión de 2014 en la sede de la compañía en Mountainview. California.

"Me parece muy aterrador, muy preocupante, muy triste, y me parece terrible, espantoso, extraño, desconcertante, desconcertante, que la gente se apresure ciega y delirando hacia adelante en la creación de estas cosas".

Esa es una buena cantidad de adjetivos desagradables para unir. Pero cuando un lenguaje como ese proviene de alguien que el filósofo Daniel Dennet dice que es mejor que nadie en el estudio de los fenómenos de la mente, te hace apreciar la gravedad potencial de lo que está en juego.