

LA IA: ¿PERO RAZONARÁ IGUAL QUE NOSOTROS?



Dr. VÍCTOR THEOKTISTO

La Inteligencia Artificial (IA, o AI en inglés), fundamentada en su potencial para ocasionar transformaciones profundas y reacomodos sociales que diriman el rumbo de la humanidad, se ha convertido en el más reciente campo de batalla de una continuada competición hegemónica que rivaliza con la tecnología nuclear y espacial. Y sin embargo, no estamos seguros de que sea una inteligencia tal como la entendemos. Se está glosando sobre los beneficios (reales, sustanciales y revolucionarios) en la solución de problemas de la humanidad, y solo dejando advertencias genéricas sobre las consecuencias de empujar las fronteras hacia una explotación poco ética y hasta criminal de la tecnología.

INTRODUCCIÓN:

¿ES LA INTELIGENCIA ARTIFICIAL COMO LA INTELIGENCIA HUMANA?

El padre de la psicología, William James, la describió en *Principles of Psychology* (1890) como "la cosa más misteriosa del mundo": ¿Cómo puede ser física la conciencia? ¿Cómo es posible que unas pocas libras de gelatina gris den lugar a nuestros propios pensamientos y a nuestro yo?.

de las siguientes habilidades: la capacidad de usar el lenguaje; de formar conceptos; de resolver problemas que ahora sólo pueden resolver los humanos; de mejorarse así mismos". McCarthy reconoció que fracasaron porque "la IA es más difícil de lo que pensábamos". Y al parecer sigue siendo más difícil de lo que se podría pensar.

El objetivo último de la IA, lograr que una máquina tenga una inteligencia de tipo general similar a la humana, es uno de los más ambiciosos que se ha planteado la ciencia. Por su dificultad, es comparable a explicar el origen de la vida, el origen del universo o conocer la estructura de la materia. Hablamos de inteligencia artificial como si tuviéramos ya la respuesta de lo que constituye nuestra inteligencia "natural", y allí donde comienzan las lagunas y los interrogantes, porque no sabemos (todavía) que nos hace humanos: compartimos casi el 99 % de nuestro material genético con chimpancés, y las diferencias en cognición son absolutamente enormes en órdenes de magnitud a ese 1% de diferencia, extensible a gorilas y orangutanes. Lo mismo con otros animales considerados inteligentes como elefantes, delfines, cuervos, pulpos o loros.

Todo comienza con la famosa pregunta de Alan Turing en la década de 1950, "¿Pueden las máquinas pensar?". Ya de por sí, el concepto de "razonar por sí misma" está lleno de trampas.

El hombre que acuñó el término Inteligencia Artificial, John McCarthy, postulaba en 1956 que "si un grupo de científicos cuidadosamente seleccionados trabajan juntos en ello durante un verano, lograrían un progreso significativo hacia la creación de máquinas con una o más

de los trabajos de la psicología evolutiva y del desarrollo apuntan que la mente no es una cosa, sino muchas, según postula uno de los pioneros en el campo:

¿Qué truco de magia nos hace inteligentes? El truco es que no hay truco. El poder de la inteligencia proviene de nuestra gran diversidad, no de un principio único y perfecto.

-Marvin Minsky,

La sociedad de la mente

Minsky argumenta que debemos ver la cognición como una "sociedad mental", con decenas o cientos de "agentes" distintos, cada uno especializado en diferentes tipos de tareas. Por ejemplo, beber una taza de té requiere la interacción de un agente de SUJETAR, un agente de BALANCEO, un agente que sienta SED y un cierto número de agentes de MOVIMIENTO.

Las siguientes definiciones, extraídas de un texto clásico sobre IA de Winston [1], así como del estudio de la inteligencia biológica, tienen claridad explicativa sobre los efectos reales y potenciales de la inteligencia, especialmente en las máquinas. "La inteligencia es la capacidad de hacer lo correcto en el momento adecuado, en un contexto en el que no hacer nada (no cambiar el comportamiento) sería detrimental con respecto a tomar una acción específica". Por lo tanto, la inteligencia requiere (a) Capacidad de

Las siguientes definiciones, extraídas de un texto clásico sobre IA de Winston [1], así como del estudio de la inteligencia biológica, tienen claridad explicativa sobre los efectos reales y potenciales de la inteligencia, especialmente en las máquinas. "La inteligencia es la capacidad de hacer lo correcto en el momento adecuado, en un contexto en el que no hacer nada (no cambiar el comportamiento) sería detrimental con respecto a tomar una acción específica". Por lo tanto, la inteligencia requiere (a) Capacidad de



Víctor Theoktisto, Ph.D., Computación de la Universitat Politècnica de Catalunya BarcelonaTECH (UPC), España. Posee un M.Sc. en Ingeniería Industrial de la Universidad de Texas en Austin. Profesor Titular, Vicerrector Académico Universidad Simón Bolívar. miembro del Directorio de la Comisión Nacional de Informática y Telecomunicaciones (CONATI)

percibir contextos de acción, (b) Capacidad para actuar y (c) Mecanismos de asociación para que determinadas acciones se ejecuten en determinados contextos.

Según esta amplia definición, las plantas, los caracoles, las alarmas de vehículos y los termostatos son inteligentes. Perciben y reaccionan al contexto: las plantas reaccionan a la dirección de la luz y los termostatos a la temperatura. Un sistema inteligente entonces lo calificamos como cognitivo si es capaz de modificar (y aumentar) su inteligencia, lo que las plantas y los termostatos (al menos mecánicos) no pueden hacer. Los sistemas cognitivos pueden aprender nuevos contextos y acciones, y nuevas asociaciones entre los anteriores y los

recientes.

La inteligencia, tal como se define aquí, es transformación de la información, un subconjunto estricto del cálculo. La computación es un proceso físico, no matemático: requiere tiempo, espacio, recursos y energía. La definición operativa que ha permitido desarrollar el campo de la IA se basa en tomar como axioma que la inteligencia es el subconjunto de la computación que convierte el

contexto en acción.

Como ha subrayado Judea Pearl [2], **la comprensión de la causalidad es el aspecto primordial de la cognición humana.** En un mundo simple donde tuviéramos pleno conocimiento de todo, la única causalidad que necesitaríamos sería la física. Podríamos predecir qué afecta a qué, haciendo simulaciones; si aplicamos una fuerza de X magnitud, ¿qué sucede a continuación? Pero en el mundo real ese tipo de simulación detallada no es realista; hay demasiadas partículas que seguir y muy poco tiempo, y nuestra información es demasiado imprecisa.

En su lugar, solemos utilizar aproximaciones; cosas que pretendemos están relacionadas causalmente, aunque no entendamos por

qué. Tomamos café porque sabemos que nos eleva el nivel de atención; no necesitamos entender la bioquímica de la cafeína. El conocimiento causal es la base de muchas de nuestras acciones. Por lo tanto para simular inteligencia tenemos que modelar el único instrumento que conocemos que utiliza relaciones causales para hacer inferencias: nuestro propio cerebro. Nuestro objeto de estudio es también la herramienta. Ya de partida arrancamos con serios problemas.

El cerebro

humano es enormemente complejo y diverso, con más de 150 áreas cerebrales claramente identificables, unos 86.000 millones de neuronas (con miles de tipos diferentes); billones de sinapsis (una neurona está conectada con más de 10.000 otras más); cientos de proteínas distintas dentro de cada sinapsis individual, además de un número indeterminado de hormonas y neurotransmisores. Con alta complejidad y cantidad de elementos, las neuronas reales como procesadores de información biológica tienen

arborizaciones dendríticas complejas con propiedades eléctricas y químicas no triviales, y presentan conductancias iónicas que producen efectos no lineales. Además, la mayoría de las células cerebrales no son neuronas, sino células gliales, que no sólo regulan el funcionamiento de las neuronas, sino que también muestran potenciales eléctricos, generan ondas de calcio y se comunican entre sí, lo que parece indicar que desempeñan también un papel importante en los procesos cognitivos. Las neuronas del cerebro no solo se comunican por vía canales iónicos con otras neuronas, sino que se ha descubierto que también emiten fotones (biofónica). Esto

implica que los efectos cuánticos de superposición, coherencia-decoherencia, tuneo, y entrelazamiento (entanglement) juegan un rol importante en la interacción entre células cerebrales, y que haya resurgido

la hipótesis del Premio Nobel de Física Roger Penrose [3] [4] que la conciencia (humana o de otro tipo) reside en esas interacciones cuánticas.

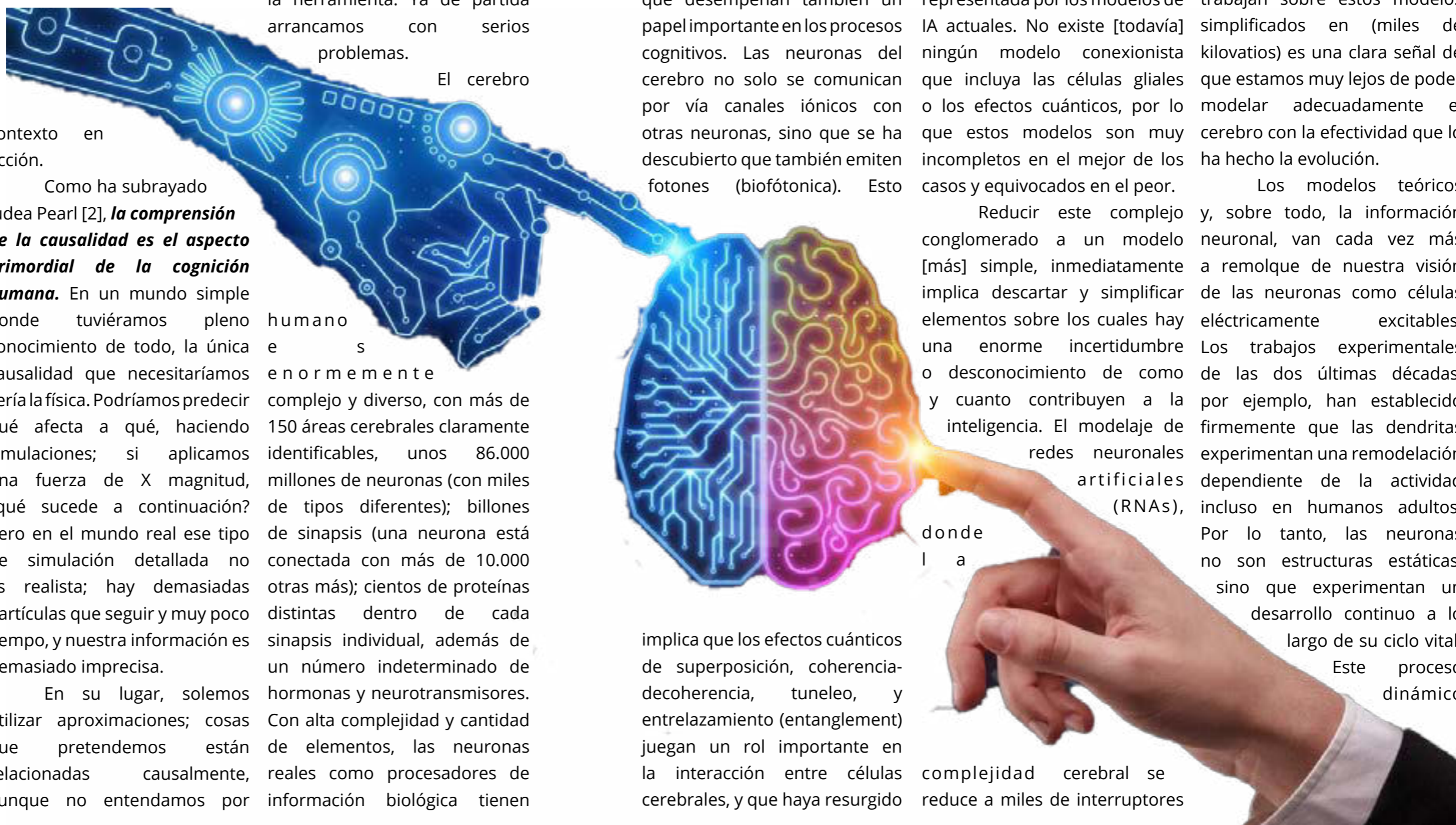
Todo esto con un consumo eficiente de energía del cerebro equivalente a un bombillo de 25 vatios, y una velocidad individual de computo bastante más baja que la de cualquier calculadora. En resumen, la enorme complejidad del cerebro está lejos de ser representada por los modelos de IA actuales. No existe [todavía] ningún modelo conexionista que incluya las células gliales o los efectos cuánticos, por lo que estos modelos son muy incompletos en el mejor de los casos y equivocados en el peor.

Reducir este complejo conglomerado a un modelo [más] simple, inmediatamente implica descartar y simplificar elementos sobre los cuales hay una enorme incertidumbre o desconocimiento de como y cuanto contribuyen a la inteligencia. El modelaje de redes neuronales artificiales (RNAs), donde la

complejidad cerebral se reduce a miles de interruptores

digitales simulados con cientos de conexiones ponderadas entre ellos, en un ensamble de múltiples procesadores digitales masivamente paralelos, son un modelo pobre y quizás errado de representar los eventos bioquímicos y de naturaleza cuántica que ocurren en el cerebro para adquirir, procesar y producir información. El gasto de energía que las grandes implementaciones de IA consumen cuando trabajan sobre estos modelos simplificados en (miles de kilovatios) es una clara señal de que estamos muy lejos de poder modelar adecuadamente el cerebro con la efectividad que lo ha hecho la evolución.

Los modelos teóricos y, sobre todo, la información neuronal, van cada vez más a remolque de nuestra visión de las neuronas como células eléctricamente excitables. Los trabajos experimentales de las dos últimas décadas, por ejemplo, han establecido firmemente que las dendritas experimentan una remodelación dependiente de la actividad incluso en humanos adultos. Por lo tanto, las neuronas no son estructuras estáticas, sino que experimentan un desarrollo continuo a lo largo de su ciclo vital. Este proceso dinámico



tiene importantes consecuencias para el funcionamiento tanto de las neuronas como del organismo. Por ejemplo, en organismos (como las orugas que se convierten en mariposas) que muestran una drástica contracción del cerebro, parte de sus recuerdos aprendidos permanecen y sobreviven al proceso.

También se ha permitido explorar la

reproducir la propagación y el procesamiento de la señal dependiente del tiempo en representaciones simplificadas de los árboles dendríticos, pero se vuelve rápidamente intratable al aumentar la complejidad geométrica. Los modelos de hardware de las neuronas, con varias capas de subredes ocultas, cada vez más sofisticados, permiten explorar la

reproducir la propagación y el procesamiento de la señal dependiente del tiempo en representaciones simplificadas de los árboles dendríticos, pero se vuelve rápidamente intratable al aumentar la complejidad geométrica. Los modelos de hardware de las neuronas, con varias capas de subredes ocultas, cada vez más sofisticados, permiten explorar la

plasticidad para la inteligencia. Además, no comprende la cuestión fundamental de la neurociencia cognitiva: cómo la información presente en el mundo externo se codifica, a través de alguna combinación de evolución, desarrollo, remodelación y aprendizaje, en la arquitectura funcional de un sistema altamente encadenado.



demostrado que las subredes dendríticas pueden computar operaciones lógicas locales, incluida la operación O exclusivo (XOR) de los circuitos digitales. La idea de que los árboles dendríticos computan eficientemente sumas ponderadas de señales se remonta al siglo XIX, y no sólo sigue dominando el desarrollo de las redes neuronales artificiales (RNA), sino que también guía el pensamiento biológico. El modelado teórico explícito de los cables puede

codificación de fase, la modulación de frecuencia y otras formas de computación híbrida analógico-digital, pero no son fáciles de analogar a las redes neuronales biológicas. y no son componentes estándar de la teoría neurocientífica.

Conceptualizar las neuronas como cables biológicos funcionalmente sofisticados no proporciona la base para entender la comunicación, las estructuras abstractas intermedias necesarias y la

Los cerebros no se construyen, como las RNA (Redes Neuronales Artificiales), en ausencia de actividad funcional, ni "empiezan a aprender" desde un estado inicial por defecto de conectividad uniforme. Los cerebros no se "conectan" a una configuración inicial de cero, tras lo cual se encienden y se exponen al mundo. Los cerebros se desarrollan a partir de microentornos multicelulares

funcionalmente complejos [5]. El primordio local, con sus diversos tipos de células, su complejidad bioquímica y (micro)anatómica, y su red de señalización biomolecular y bioeléctrica, es el "mundo" con el que interactúa cada neurona, y es la única fuente de información no genómica sobre cada neurona. Algunos escenarios de estos entornos son del tipo "socialización", toma de decisiones y resiliencia colectiva, donde una asociación topológica de grupos de neuronas afines en sitios separados, pero cercanos, dedicados a sus respectivas tareas, pueden a veces unirse en una coalición, votando por un estímulo y tomando una decisión similar a la del comité.

Esta inmensa complejidad del cerebro también nos lleva a pensar que la llamada **singularidad**, es decir, las futuras superinteligencias artificiales que, basadas en réplicas del cerebro, superarán concreces la inteligencia humana en un plazo de unos veinticinco años, es una predicción con poca base científica. La verdadera preocupación en estos momentos es decidir si la IA está orientada a aumentar humanos o a reemplazarlos.

¿QUE ESPERAMOS QUE PRESENTE UNA IA A MEDIDA QUE LA VAMOS CONSTRUYENDO?

Con expertos de todo campo, se ha llegado a expectativas comunes por las que podemos de alguna manera evaluar si

se está logrando alcanzar un grado de inteligencia con el que nos podamos identificar. En la actualidad, el aprendizaje profundo tiene dificultades con la inferencia y el razonamiento abstracto porque, en primer lugar, no está orientado a representar conocimientos fácticos precisos. Una vez que los hechos son difusos, es difícil acertar con el razonamiento.

Por un lado, esperamos que una IA presente representaciones internas ricas. En la actualidad, el aprendizaje profundo trata de eludir esto, con un montón de vectores que capturan algo de una manera aproximada, pero nunca el absoluto. Resulta que la IA no puede explicar como llegar a conclusiones. El sistema GPT-3 de composición de texto, del que tanto se habla, es un buen ejemplo de ello. Es incapaz de responder de forma fiable a preguntas sencillas como "si pones dos manzanas en un plato y añades una pera, ¿cuántas frutas hay?"

Mucho de lo que sabemos es bastante abstracto. Esperamos que una IA logre hacer abstracción y generalización. Por ejemplo, la relación "X es hermana de Y" es válida entre muchos pares de personas diferentes: Zeus es hermano de Poseidón, María Antonia Bolívar es hermana de Simón Bolívar, etc. Además, sabemos lo que son las hermanas y hermanos en general como concepto y podemos aplicar ese

conocimiento a otros individuos. Si dos personas tienen los mismos padres, deducimos que son hermanos. Si sabemos que María Antonia era hija de Juan Vicente y María de la Concepción, y descubrimos que Simón también era su hijo, podemos inferir que María Antonia y Simón son hermanos.

Las representaciones que subyacen a los modelos cognitivos (y al sentido común) se construyen a partir de relaciones abstractas, combinadas en estructuras complejas: trozos de tiempo ("9:25 PM"), trozos de espacio ("Catarata de El Salto Ángel"), sucesos concretos ("el atentado contra Juan Pablo II"), organizaciones sociopolíticas ("la ONU") y constructores teóricos ("semántica"), para utilizarlos en una explicación o relato reduciendo las situaciones complejas a los elementos esenciales, una enorme ventaja para razonar sobre el mundo.

Una limitación crucial de los sistemas de aprendizaje profundo es el así llamado «olvido catastrófico»: si una RNA ha sido entrenada para llevar a cabo una tarea (por ejemplo, jugar al ajedrez), si a continuación la entrenamos para llevar a cabo otra tarea distinta (por ejemplo, distinguir entre imágenes de perros y de gatos) olvidan completamente la tarea anteriormente aprendida (en este caso jugar al ajedrez). Es una prueba contundente de que en realidad las IAs no aprenden nada, en el sentido humano de

aprender.

Estos sistemas son difíciles de depurar y rara vez tienen la flexibilidad necesaria. Por ejemplo, está el caso de Deep Blue, el supercomputador de IBM que venció al gran maestro de ajedrez Garry Kasparov. Deep Blue ganó por pura fuerza bruta. Para cada movimiento legal que podía hacer, consideraba las respuestas de su oponente, sus propias respuestas a esas respuestas, y así sucesivamente durante seis o más pasos en línea. Con una función de evaluación rápida, calculaba una puntuación para cada posición posible y luego realizaba la jugada que le permitía obtener la mejor puntuación. Lo que permitió a Deep Blue vencer a los mejores humanos del mundo fue la potencia computacional aplastante. Podía evaluar hasta 330 millones de posiciones por segundo, mientras que Kasparov sólo podía evaluar unas pocas docenas antes de tener que tomar una decisión.

Una RNA como Deep Blue aprende y razona todo lo que hay que aprender sobre ajedrez porque tiene acceso a millones de partidas y puede explorar millones de posibilidades en un corto tiempo. No es así como los grandes maestros de ajedrez aprenden y juegan: a través de la experiencia van construyendo sistemas cognitivos altamente estructurados (estrategias dentro de estrategias) que les permiten operar con muchos menos datos y velocidad de

cómputo. Deep Blue no puede decirnos cómo concibe y entiende Kasparov un tablero de ajedrez. De hecho, ni siquiera puede decirnos nada de como jugamos al ajedrez.

Si queremos cambiar de contexto y que Deep Blue ahora sea campeón de Dominó, debemos entrenarlo otra vez con las nuevas experiencias, y a pesar de la memoria disponible olvidará como jugar al ajedrez. En la inteligencia tal como la conocemos, el nuevo conocimiento no sustituye lo aprendido. Sistemas híbridos como Siri, Alexa, el predictor de palabras de un teléfono inteligente y los motores de búsqueda de la web utilizan enfoques integrando muchos tipos de procesos diferentes, integrando el aprendizaje profundo y las técnicas simbólicas [6]. El sistema DALL-E, (ver Figura 1), puede dar buenos resultados en la respuesta a preguntas visuales y en la recuperación de imágenes y textos. Los sistemas construidos sobre esa base pueden producir textos gramaticales, pero muestran poca comprensión de lo que se desarrolla a lo largo del tiempo en el texto que producen, al faltarles composicionalidad.

Los psicólogos cognitivos distinguen entre información ascendente (directamente

de nuestros sentidos), y conocimiento descendente (conocimiento previo sobre el mundo): letras y dígitos son categorías distintas, las

palabras y los números se componen de elementos extraídos de esas categorías. Podemos sumar números y concatenar palabras en frases, etc. Todo lo que percibimos lo integramos en un modelo cognitivo de la situación y en nuestra comprensión del mundo en su conjunto.

Los conceptos integrados

en las teorías son vitales para un aprendizaje eficaz. Supongamos que una niña de preescolar ve por primera vez la fotografía o video de una iguana. De allí en

adelante, la niña será capaz de reconocer iguanas en la vida real, distinguiéndolas fácilmente de los venados. Asimismo, la niña podrá deducir, a partir de sus conocimientos generales sobre los animales, que las iguanas comen y respiran y que nacen pequeñas, crecen, se reproducen y mueren. Y

que se parecen lo suficiente a otros reptiles como lagartijas y caimanes como para ponerlos a todos en una misma familia (reptiles, una abstracción). Para tener éxito, una inteligencia general tendrá que integrar los hechos que adquiere en teorías generales más ricas que ayuden a organizar esos hechos [7].

A lo largo de nuestra vida cotidiana, llevamos la cuenta en varios hilos de todo tipo de objetos individuales, sus propiedades y sus relatos. “Antes solía trabajar como mensajero, mi vehículo tiene una rueda con un reventón y el año pasado reparé la nevera”. Nuestra experiencia se compone de entidades que persisten y mutan a lo largo del tiempo, y gran parte de lo que sabemos se organiza en torno a esas particulares historias.

Aprender a partir de una pizarra absolutamente en blanco, como pretenden hacer la mayoría de los investigadores de aprendizaje automático, hace que el juego sea mucho más difícil de lo que debería ser. Se trata de la crianza sin la naturaleza, cuando la solución más eficaz es, obviamente, combinar ambas. Ya se sabe que nacemos cableados de fábrica para el lenguaje, con una innata psicología empírica e intuitiva, y es probable que también

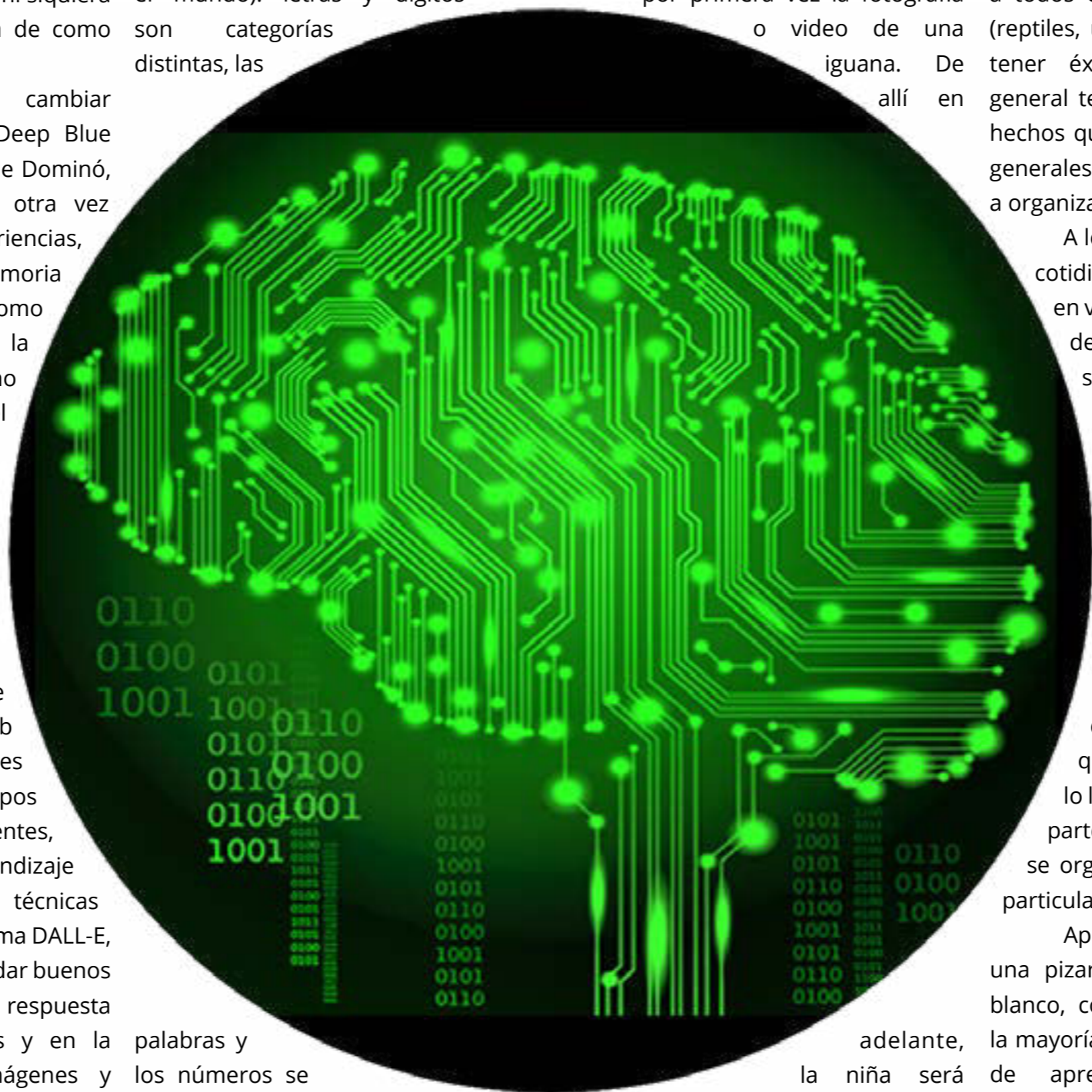
arribemos con el concepto de que el mundo está formado por objetos perdurables que se desplazan por trayectorias conectadas en el espacio y el tiempo, con un sentido de la geometría y la cantidad. La dualidad “naturaleza y crianza” es una falsa dicotomía. Las pruebas de la biología, de la psicología y la neurociencia del desarrollo son abrumadoras: ambas trabajan juntas.

Del mismo modo, los sistemas de IA no deberían intentar aprenderlo todo a partir de las correlaciones entre píxeles y acciones, sino partir de una comprensión básica del mundo como base para desarrollar modelos más ricos.

INTELIGENCIA ARTIFICIAL COMO INTELIGENCIA ALIENÍGENA

Los descubrimientos de las ciencias cognitivas pueden decirnos mucho en nuestro intento de construir una inteligencia artificial con la flexibilidad y la generalidad de la mente humana. No es necesario que las máquinas reproduzcan la mente humana, pero un conocimiento profundo de ésta puede conducir a importantes avances en la IA.

Según Douglas Hofstadter [8], la etiqueta de “inteligencia” es inapropiada para describir las ideas extraídas por una potencia bruta de cálculo en conjuntos masivos de datos, porque, desde su punto de vista, el hecho de que los resultados parezcan inteligentes es irrelevante si el



proceso subyacente no se parece al pensamiento inteligente.

El hecho en la década de 1980 ya existía un programa tan sencillo como Candide que usando Machine Learning podía hacer traducción automática eficaz sin requerir comprensión: todo lo que necesitaba era un montón de data de textos bilingües. Candide se convirtió en una prueba de concepto para el enfoque que conquistó la IA y es dominante ahora con la ayuda de RNAs, y que como resultado ha convertido a Google en la máquina de inteligencia artificial más grande del planeta, simplemente porque tiene acceso a ingentes cantidades de datos. De nuevo, no es así como

los humanos piensan.

LA IA REFLEJA LOS PREJUICIOS DE LA DATA O DE QUIENES LA ALIMENTAN

Bajo la promesa de hacernos la vida más fácil, la IA está cambiando lo que significa ser humano, quizás para mal. En la transición del siglo XVIII al siglo XIX surgió en Gran Bretaña el Movimiento Luddita, los primeros "tecnófobos", como reacción de los artesanos tejedores contra los primeros ingenios textiles, que amenazaban su existencia debido a la tecnología precursora que permitía a máquinas tejedoras producir textiles de alta calidad y bajo costo, y

condenando a esos artesanos al desempleo y hambre. Los artesanos se constituyeron en un movimiento subversivo que atacaba y destruía las máquinas tejedoras (sabotaje industrial). Fueron masacrados por el Gobierno Británico, y desde entonces son el referente preferido de resistencia al cambio tecnológico. Si bien no debemos preocuparnos de una hipotética emergencia de consciencia maligna de una IA que amenace a la humanidad, si debemos cuidarnos más bien de un humano o grupo de humanos, empresas o gobiernos usen esa IA para propósitos nefastos.

Equipar a las empresas y gobiernos con sistemas

avanzados de IA para que la gestión y la producción sean más eficientes requerirá menos empleados humanos y provocará más desempleo. En el área de vigilancia extrema lleva a la pérdida de privacidad y derechos civiles. Estos dilemas éticos hacen que muchos expertos en IA señalen la necesidad de regular su desarrollo. En algunos casos, incluso debería prohibirse el uso de la IA. El ejemplo claro son las armas autónomas. Los tres principios básicos que rigen los conflictos armados: la discriminación (necesidad de distinguir entre combatientes y civiles o entre un combatiente en actitud de rendición y otro dispuesto a atacar), la proporcionalidad (evitar el uso excesivo de la fuerza) y la precaución (minimizar el número de bajas y daños materiales) son extraordinariamente difíciles de evaluar y, por tanto, casi imposibles de cumplir por los sistemas de IA que controlan las armas autónomas. Para su prohibición, es necesario asignar responsabilidad moral y penal a quienes codifican, fabrican, financian y despliegan robots asesinos.

Pero incluso si, a muy largo plazo, las máquinas tuvieran esta capacidad de participar activamente en acciones bélicas, sería indigno delegar la decisión de matar en una máquina. El escritor Frank Herbert, en su novela Dune, toca este tema al hablar de

una rebelión llamada la Yihad Butleriana, que acaba con todos los computadores y la inteligencia artificial en general, y cuyo principal grito de batalla es el mandamiento "No falsificarás una mente humana"

Será muy triste y un veredicto lapidario sobre nosotros, sus creadores, si al final la razón por la que podemos calificar una IA como verdaderamente inteligente es por lograr imitar una de las peores facetas del comportamiento del ser humano.

REFERENCIAS

[1] Patrick Henry Winston. Artificial Intelligence. Addison-Wesley. (1984)

[2] Judea Pearl. Causality: Models, Reasoning, and Inference. Cambridge University Press (2000).

[3] Stuart Hameroff and Roger Penrose. Orchestrated reduction of quantum coherence in brain microtubules: A model for consciousness. Mathematics and Computers in Simulation, Volume 40, Issues 3-4, April 1996, Pages 453-480 [https://doi.org/10.1016/0378-4754\(96\)80476-9](https://doi.org/10.1016/0378-4754(96)80476-9).

[4] Roger Penrose, Henry Stapp, Don Page, Michael Nauenberg, Fred Kuttner, Walter J. Christensen, Bruce Rosenblum, Stuart Hameroff and Ellen Langer. Consciousness

and the Universe: Quantum Physics, Evolution, Brain & Mind. Cosmology Science Publishers (2022). ISBN 13: 9781938024306

[5] Michelle Baddeley, Joanna Bryson, Nancy Chau, Barry Eichengreen, Francisco González, Amos N. Guiora, Peter Kalmus, Ravi Kanbur, Ramón López de Mántaras, María Martín-Torres, José M. Mato, Diana Owen, Alex Pentland, Carlo Ratti, Martin Rees, Victoria Robinson, Daniela Rus, José Manuel Sánchez Ron, Vivien A. Schmidt, Samuel H. Sternberg, Sandip Tiwari, Yang Xu, Ernesto Zedillo Ponce de León. ¿Hacia una nueva ilustración? Una década trascendente. BBVA OpenMind series, Turner Libros (2019). ISBN: 978-84-17141-20-2

[6] Bill Yuchen Lin, Wangchunshu Zhou, Ming Shen, Pei Zhou, Chandra Bhagavatula, Yejin Choi and Xiang Ren, CommonGen: A Constrained Text Generation Challenge for Generative Commonsense Reasoning.. arXiv (2020). <https://arxiv.org/abs/1911.03705>.

[7] Gary Marcus, Ernest Davis. Insights for AI from the Human Mind. Communications of the ACM, January (2021), Vol. 64 No. 1, Pages 38-41, doi 10.1145/3392663.

[8] Douglas Hofstadter. Fluid Concepts and Creative Analogies. Basic Books (1996).



Figura 1. Imágenes generadas por el autor usando la IA llamada DALL-E como respuesta a la frase "Dibuja a Hugo Chávez en el estilo pictórico de X". La IA busca en internet las imágenes apropiadas y las reinterpreta produciendo nuevos dibujos en el estilo apropiado, demostrando una asombrosa y preocupante capacidad de reconocer lenguaje, intención, y contexto en la producción de Deep Fakes. Los pintores usados como plantilla aparecen en orden: Velásquez, Rembrandt, Rafael, Miguel Ángel, Van Gogh, Manet, Sorolla, Seurat, Gauguin, Kalo, Picasso(a), Picasso(b), Reverón, Morisot, Warhol.