

# ¿Qué es en definitiva la inteligencia artificial?

Como las computadoras crecen en complejidad  
las máquinas están comenzando  
a parecer inteligentes,  
pero, ¿pueden pensar actualmente?

Raymond Kurzweil

A los «fans» de Norbert Wiener les resultará familiar esta canción infantil alemana:

«¿Sabes cuántas estrellitas hay en la tienda azul del cielo?  
¿Sabes cuántas nubes pasan a lo lejos sobre el mundo entero?

El Señor Dios las ha contado de modo que ninguna falte de su inmenso número»<sup>1</sup>.

Wiener era amigo de recitar esta canción, en parte porque combinaba dos de sus tópicos favoritos —la astronomía y la meteorología— pero también porque gozaba mostrando que los expertos eran capaces, a mitad del siglo veinte, de perfeccionar sus actividades. Desde los escritos de Wiener, uno siente orgullo porque los científicos podrían actuar al menos como uno de los coros de Dios. Podríamos también enorgullecernos de que ahora hemos trasvasado el trabajo de seguir el curso de nuestras estrellas y nubes a las computadoras, que probablemente están haciendo una labor más completa que los científicos de los días de Wiener, aunque tal vez todavía no tan elevada como la de Dios. También podemos hacer la pregunta que si las

---

<sup>1</sup> NORBERT WIENER. *Cybernetics; or, Control and Communication in the Animal and the Machine*, 2nd ed., p. 30. MITT Press.

## ¿Qué es en definitiva la inteligencia artificial?

computadoras pueden hoy asumir uno de los muchos coros de Dios, ¿qué hay de los coros humanos? Ciertamente, ¿qué hay de nuestro máspreciado coro, del pensamiento?

Las máquinas que piensan, o como un reciente libro tituló, *Las máquinas quienes piensan*<sup>\*</sup>, han sido un tópico de reflexión y debate desde que Charles Babbage diseñó su «Máquina Analítica» en el comienzo del siglo diecinueve y meditó sobre el sujeto con su amiga Lady Ada Lovelance, hija de Lord Byron<sup>2</sup>. En aquellos días había unos pocos partidarios para tales proyectos ambiciosos, y esos pocos tendían a ser escépticos. En los últimos años, sin embargo, las máquinas sujeto de pensamiento, o por decirlo de un modo más conservador, las máquinas que al menos parecen inteligentes, han terminado absorbiendo el interés de una amplia audiencia, incluyendo Wall Street y Madison Avenue.

El incremento de la atención ha traído consigo una controversia creciente. Los expertos todavía debaten sobre lo que es una máquina inteligente, y todavía siguen en desacuerdo sobre qué es lo que constituye el campo de estudio que ha sido llamado inteligencia artificial (o simplemente AI)<sup>3</sup> desde que John McCarty le dio ese nombre en una conferencia de Dartmouth en 1956. La incertidumbre ha estado siempre ahí pero ahora hay más en juego. La AI ha sido comparada a la ingeniería genética, y ha traído, además de una disciplina académica, una inversión de capital y una atención pública. Ahora bien, aunque la ingeniería genética tiene sus propias controversias, carece de incertidumbre respecto de su definición. Los artículos y las conferencias no empiezan con la cuestión de qué es en definitiva la ingeniería genética. Si tiene que ver con la manipulación de genes, no hay duda de que es ingeniería genética.

La controversia que rodea a la AI se evidencia por una cierta discordia dentro de este campo. Los estudiosos académicos han acusado al desarrollo industrial de la AI de ser superficial y dar una prioridad demasiado alta a las metas comerciales a corto plazo más que a la investigación a largo plazo. Los investigadores industriales han acusado a los académicos de ofrecer demostraciones superficiales, de no desarrollar robustos sistemas que trabajen de verdad. Las compañías han acusado a sus competidores de no usar técnicas reales de AI, y así sucesivamente.

Desafortunadamente, no nos vamos a encontrar publicado en un tiempo próximo el trabajo definitivo que despeje esta confusión para todos y para siempre. Puesto que ése no será el resultado de este artículo, también yo me

---

\* «Machines who think» por oposición a «machines that think» (N. del T.).

<sup>2</sup> CHARLES BABBAGE. 1973. *Babbage's Calculating Engines*. Charles Babbage Research Institute Reprint Series for the History of Computers, vol. 2. Los Angeles: Tomash Publishers.

<sup>3</sup> Artificial intelligence (N. del T.).

abstendré de hacer semejante afirmación. Pero un examen de las cuestiones relativas a la inteligencia artificial puede sin duda ayudarnos a determinar aquellas cuestiones que no se pueden responder, como aquellas otras que están respondiéndose con validez. En la última categoría incluiría una comprensión de qué es capaz de llevar a cabo hoy la disciplina de la inteligencia artificial, qué será capaz de llevar a cabo y cómo podemos llegar a esos objetivos del mejor modo posible.

Me aproximaré a estas cuestiones y a la cuestión central —¿qué es en definitiva la inteligencia artificial?— desde tres perspectivas: el pasado, el presente y el futuro.

### *La inteligencia artificial en el pasado*

Más que recitar la letanía usual de los primeros experimentos, comenzaré en cambio examinando algunas de las raíces intelectuales del movimiento de la AI, que creo será más revelador. Uno se ve tentado a volver a las máquinas de cálculo mecánico, que fueron propuestas tan tempranamente en el s. XVIII por Pascal y Leibniz. Quizás la más famosa fue la máquina analítica de Babagge de 1833, la primera máquina diseñada para usar un programa almacenado, para ser leído por tarjetas perforadas. Aunque la máquina nunca fue construida, los programas estaban hechos para ella, y las intensas discusiones de Babagge y Lovelace las han mantenido los que reflexionan sobre cuestiones tales como un código automodificable, programas que jugarían al ajedrez o compondrían música, y cómo (o si) la inteligencia automatizada podría ser relacionada con el pensamiento humano.

Aunque fascinantes vistas en retrospectiva, estas primeras discusiones no produjeron inmediatamente el fermento desde el que las computadoras y la inteligencia artificial iban a aparecer. Aquellas emergieron, en cambio, a partir de algunos de los movimientos intelectuales más poderosos de principios del siglo veinte. Uno de ellos fue el movimiento filosófico llamado positivismo lógico, que se esforzó por examinar la epistemología con el mismo rigor que estaba entonces de moda en el mundo de las matemáticas. La epistemología es el estudio de los orígenes, métodos y límites del conocimiento —qué podemos conocer y cómo podemos conocerlo—. Aunque las computadoras como tales no estaban en las mentes de los primeros positivistas lógicos, no es sorprendente que una atención por definir la naturaleza del conocimiento con precisión matemática había de tener relevancia para el subsiguiente nacimiento de las computadoras.

Un trabajo influyente en el desarrollo del positivismo lógico fue el *Tractatus Logico-Philosophicus* de Ludwig Wittgenstein, publicado en 1921<sup>3</sup>. El li-

---

<sup>3</sup> LUDWIG WITTGENSTEIN. 1961. *Tractatus Logico-Philosophicus*. Newtrans. D. F. Pears and B. F. McGuinness. Routledge and K. Paul.

bro es doblemente interesante en lo que dice y en su propia estructura interna. El tratado contiene sólo siete declaraciones formales, numeradas del 1 al 7. Para ayudarnos a lo largo del él, Wittgenstein también incluye varios niveles de modificación de las declaraciones para clarificar las principales. Por ejemplo, las declaraciones 1.1. y 1.2. modifican la declaración 1. En cambio, la 1.1.1. se trae para explicar la 1.1. y así sucesivamente. El lector puede elegir leer el libro de izquierda a derecha, desde abajo, o incluso al revés. La estructura modular agradecería, estoy seguro, a los que propusieran un buen estilo de programación.

Wittgenstein comienza con la declaración 1, «El mundo es todo lo que acontece», y finaliza con la declaración 7, «Sobre lo que no podemos hablar, debemos guardar silencio». Como se puede deducir de estos puntos de comienzo y de final, se trata de un trabajo ambicioso. Las declaraciones que hay entre ellos determinan con un concepto formal de lenguaje qué puede ser dicho, y establecer la conexión entre lo que puede ser dicho y lo que puede ser pensado o conocido.

Declara:

4.0.0.3.1. Toda filosofía es una «crítica del lenguaje».

5.6. Los límites de mi lenguaje son los límites de mi mundo.

5.6.1. No podemos pensar lo que no podemos decir.

En el mundo de Wittgenstein hay ciertos hechos elementales, hay proposiciones sobre las relaciones entre los hechos elementales, y hay ciertas transformaciones permisibles sobre tales proposiciones que producen proposiciones compuestas. Su modelo de pensamiento humano consiste en que podemos recibir impresiones de sentido que comprenden hechos elementales. Podemos entonces transformar esos hechos elementales y derivar relaciones entre ellos de acuerdo con ciertos procesos lógicos admisibles. Cualquier pensamiento fuera de este esquema es, o bien falso, o sin sentido.

Si he simplificado demasiado la teoría, es para extraer dos puntos importantes. Ambos tienen un peso en las raíces intelectuales de la inteligencia artificial. El primero —y Wittgenstein establece este punto de diferentes maneras— es que lo que nosotros podemos pensar es lo que cuenta. Establece una vinculación directa entre el pensamiento humano y un proceso formal que puede ser descrito sólo como un cálculo. Para reorganizar las declaraciones de Wittgenstein: no podemos pensar lo que no podemos decir; no podemos decir o al menos no deberíamos decir lo que no tiene significado en el lenguaje que hablamos; y las manifestaciones en cualquier lengua son ciertamente no significativas a menos que puedan derivarse de una secuencia formal de cálculo como transformaciones sobre una base de datos de proposiciones elementales.

Esta descripción del pensamiento humano como una secuencia formal

de cálculo sería sostenida dos décadas más tarde en las tesis de Church-Turing, que discutiré brevemente. No es una tesis familiar a la que cualquiera se adhiere necesariamente, y permanece como discutible en la literatura filosófica. El mismo Wittgenstein terminó rechazándola, y en sus últimos trabajos tenía mucho que decir sobre los sujetos que, según había argüido en el *Tractatus*, quedarían relegados al silencio.

El segundo punto establecido en el *Tractatus*, que tendría más tarde un significado para las teorías computacionales, era que el pensamiento está incrustado en el lenguaje. Es también interesante notar que el lenguaje así concebido en el *Tractatus* tiene más la calidad del lenguaje de programación LIPS o incluso PROLOG que la que tiene el alemán nativo de Wittgenstein.

Aunque Wittgenstein terminó rechazando muchas de las ideas expresadas en su primer gran trabajo, sus pensamientos fueron elaborados dentro de la teoría formal del positivismo lógico por filósofos tales como Alfred Ayer y Bertrand Russell. Russell, además, extendió este nuevo formalismo en la epistemología al mundo de las matemáticas en el trabajo original que escribió con Alfred North Whitehead, *Principia Mathematica*, que representaba la fundamentación de la moderna teoría de series<sup>4</sup>.

A partir de los fundamentos de la teoría de series surgió a su vez una sorprendente teoría nueva del matemático Alan Turing, en 1937, que conduciría a la moderna teoría computacional. En efecto, Turing recuperó la afirmación hecha originalmente por Wittgenstein, de que el verdadero pensamiento es cálculo, sólo que esta vez expresó explícitamente la idea como una definición formal del cálculo más que del pensamiento humano.

Un propósito del trabajo pionero de Turing era señalar un problema en cuya respuesta habían fracasado los *Principia*: el así llamado vigésimotercer problema del gran matemático alemán David Hilbert. Esta cuestión, brevemente expuesta, es la cuestión de si es posible trazar un método que pueda establecer la verdad o la falsedad de cualquier declaración en un cierto lenguaje de lógica llamado cálculo de predicados<sup>5</sup>. Examinando lo que Hilbert quiso decir con la palabra «método», Turing estableció una definición formal del método como algoritmo. Además trazó un concepto duradero de un algoritmo como un programa que podría correr sobre lo que ha sido conocido como máquina de Turing.

La máquina de Turing ha perdurado como nuestro principal modelo teórico de cálculo por su combinación de simplicidad y poder. Su simplicidad deriva de su cortísima lista de capacidades. Puede leer una cinta y deter-

---

<sup>4</sup> ALFRED NORTH WHITEHEAD and BERTRAND RUSSELL, 1957. *Princ. Mathematica*, vol. 1, Cambridge Univ. Press.

<sup>5</sup> DOUGLAS M. CAMPBELL and JOHN C. HIGGINS, eds. N. d. *Mathema People, Problems and Results*. Wadsworth International.

minar su siguiente operación basado en la lectura de un cero o un uno; puede mover la cinta a izquierda o a derecha; puede escribir un cero o un uno sobre la cinta; puede saltar a otro orden; y puede detenerse. Así, Turing fue capaz de mostrar que esta máquina extremadamente simple puede calcular algo que cualquier otra máquina pueda calcular, independientemente de su grado de complejidad.

Hay dos razones de por qué la máquina de Turing tuvo el éxito que tuvo. La primera es esta asombrosa combinación de simplicidad y poder. La segunda, que Turing descubrió algo inesperado: que hay problemas bien definidos para los que podemos probar que existe una respuesta, pero para los que podemos también probar que la respuesta no puede ser nunca encontrada. Esos son los llamados problemas insolubles.

El más famoso de éstos, el problema del «Castor Ocupado», fue descubierto por Tibor Rado<sup>6</sup>. Puede ser expresado como sigue. Dado un número entero positivo  $N$ , construimos todas las máquinas de Turing que tienen  $N$  declaraciones, lo cual es decir  $N$  distintas configuraciones internas (ésta siempre será un número finito); eliminamos aquellas que devienen en infinitos lazos; y entonces seleccionamos la máquina que escribe el mayor número de unos sobre su cinta. El número de unos que esta máquina Turing escribe es llamado el «Castor Ocupado» de  $N$ . Rado mostró que no hay algoritmo, esto es, no hay máquina de Turing, que pueda calcular esta función para todas las  $N$ . Lo crucial del problema es la clasificación de esas  $N$  máquinas de Turing que devienen, en infinitos lazos. Si programamos lo que se llama una máquina de Turing universal para simular todas las  $N$  máquinas de Turing que declaran, el simulador entra por sí mismo en un lazo infinito. La función del «Castor Ocupado» puede ser calculada por algunos  $N$ , y es además, curiosamente, un problema insoluble el separar esos  $N$  para los que podemos determinar que son «Castor Ocupado» de  $N$  de aquellos otros para los que no podemos hacerlo.

Aparte de su interés como ejemplo de problema insoluble, la función del «Castor Ocupado» es también interesante por sí misma como función inteligente. Dicho con más precisión, en una función que requiere aumentar la inteligencia al calcular para aumentar los argumentos. Como nosotros aumentamos  $N$ , la complejidad de los procesos necesitaba calcular el «Castor Ocupado» de  $N$  aumentos. Con  $N$  igual a 6, estamos trabajando con adición, el Castor Ocupado de 6 es igual a 35. En el 7, el Castor Ocupado aprende a multiplicar, y el Castor Ocupado de 7 es igual a 22,961. En el 8 aprende a exponenciar, y el número de unos que nuestro octavo Castor Ocupado escribe sobre su cinta es aproximadamente  $10^{43}$ . A la vez que llegamos a 10, estamos tratando con un proceso más complicado que la exponenciación, y para representar el Castor Ocupado de 10 necesitamos una exótica anota-

---

<sup>6</sup> JOHN M. HOPCROFT. 1984. Turing machines. *Sci. Am.* 250 (M. 86-98).

ción en la que tenemos un montón de exponentes cuya elevación está determinada por otro montón de exponentes cuya elevación está determinada por otro montón de exponentes, y así sucesivamente. Para el duodécimo Castor Ocupado necesitamos una anotación más exótica todavía. Es probable que la inteligencia humana sea ampliamente sobrepasada antes de que el Castor Ocupado llegue a 100.

Turing mostró que hay tantos problemas insolubles como solubles, el número de cada uno de los cuales es el más bajo orden de infinidad, la llamada infinidad contable.

Trabajando independientemente durante los años 30, tres matemáticos —Kurt Gödel, Alan Turing, y Alonzo Church— mostraron que la respuesta a la cuestión vigésimo tercera de Hilbert, originalmente propuesta en el año 1900, es no. No hay un método o algoritmo que pueda determinar la verdad o falsedad de ninguna declaración lógica en el cálculo de predicados, no podemos tampoco distinguir aquellas declaraciones que pueden ser probadas de aquellas que no lo pueden ser.

El trabajo de estos matemáticos tuvo unas resonancias que todavía duran hoy. El Teorema de la Incompletud de Gödel, por ejemplo, que mostró que todos los sistemas formales de suficiente poder son capaces de generar proposiciones que no pueden ser decididas en absoluto, se ha considerado el más importante en el ámbito de las matemáticas<sup>7</sup>. Los trabajos de Gödel, Turing y Church representaron las primeras pruebas formales de que hay límites que definen lo que la lógica, las matemáticas y el cómputo pueden hacer. Esos descubrimientos contradicen fuertemente la declaración 7.5. de Wittgenstein: «Si una cuestión puede ser formulada, puede ser respondida».

Por añadidura, para encontrar algunos profundos límites a los poderes de cálculo, Church y Turing también avanzaron, independientemente, una afirmación que ha venido a conocerse como la tesis Church-Turing: que si un problema que pudiera ser presentado a la máquina de Turing no es soluble por dicha máquina, entonces tampoco lo es por el pensamiento humano. Otros han retomado la tesis de Church-Turing para proponer una equivalencia esencial entre lo que un humano puede pensar o conocer y lo que es calculable. La tesis de Church-Turing se puede ver como una restitución, de alguna manera más precisa en los términos, de una de las principales tesis de Wittgenstein en el *Tractatus*.

Sería de señalar que aunque la existencia de problemas insolubles es una certeza matemática, la tesis de Church-Turing no es en absoluto una proposición matemática. Es una manifestación que con varios disfraces está en el corazón de algunos de nuestros más profundos debates filosóficos. Ello

---

<sup>7</sup> DOUGLAS R. HOFSTADTER. 1979. *Gödel, Escher, Bach: An Eternal Gold Braid*. Basic Books.

## ¿Qué es en definitiva la inteligencia artificial?

tiene un lado negativo y otro positivo. El lado negativo es que los problemas que no pueden ser resueltos por ningún medio teórico de cómputo tampoco pueden ser resueltos por el pensamiento humano. Aceptar esta tesis significa que hay cuestiones cuyas respuestas puede mostrarse que existen pero que tal vez no se encontrarán nunca. El lado positivo es que si los seres humanos pueden resolver un problema u ocuparse en alguna actividad inteligente, las máquinas pueden últimamente construirse para actuar en el mismo sentido. Esta es una tesis central del movimiento de la inteligencia artificial, que las máquinas pueden ser hechas para realizar funciones inteligentes, que la inteligencia no es la parcela exclusiva del pensamiento humano. Por consiguiente podemos llegar a una posible definición de la inteligencia artificial: esta AI representa los intentos de suministrar demostraciones prácticas de la tesis de Church-Turing.

En su más fuerte formulación la tesis de Church-Turing señala resultados del determinismo y de la libre voluntad, que ha sido descrita como una actividad llena de intenciones que no es ni determinada ni casual, parecería contradecir la tesis de Church-Turing. No obstante, la verdad de la tesis es últimamente una cuestión de creencia personal, y los ejemplos de comportamiento inteligente de las máquinas están, probablemente, para influir en la creencia de uno al menos en el lado positivo del asunto. El escepticismo de Lady Lovelace al considerar la posibilidad de máquinas inteligentes no era una duda relativa a los límites de la computadora mecánica de palancas y ruedas dando vueltas que a ella le proponían. Hoy es posible imaginar la construcción de máquinas cuya «ferretería» compite en complejidad con el cerebro humano. Así nuestros algoritmos crecen en complejidad y las máquinas parecen ser más inteligentes y más acabadas, las discusiones sobre la tesis de Church-Turing serán más prácticas que el debate altamente teórico del tiempo de Church y Turing.

A finales de los 40, la unión entre pensamiento y cálculo era necesariamente teórica. Con el desarrollo de las computadoras electrónicas durante la segunda guerra mundial, la discusión se orientó rápidamente a la realidad de lo que podía hacerse con la maquinaria disponible. Los primeros esfuerzos en la AI fueron entusiastas, productivos y la causa de que se echaran un poco las campanas al vuelo. Herbert Simon y Allen Newell, en un artículo publicado en 1958, dicen lo siguiente: «Ahora hay máquinas en el mundo que piensan, que aprenden y que crean. Además, su habilidad para hacer todo esto va a crecer rápidamente —en un futuro próximo— hasta que la categoría de problemas que puedan manejar sea correlativa a la categoría de aquellos en los que la mente humana se ha aplicado»<sup>8</sup>.

---

<sup>8</sup> HERBERT A. SIMON and ALLEN NEWELL. 1958. Heuristic problem solving: The next advance in operations research. *Oper. Res.* 7-8.

El artículo continúa prediciendo que en 10 años (esto es, en 1968) una computadora digital será campeona mundial de ajedrez. Ocho años más tarde, en 1965, Simon escribió en otro artículo que «las máquinas serán capaces, en veinte años, de hacer cualquier trabajo de los que puede hacer el hombre»<sup>9</sup>.

No quiero ahora criticar a Simon. El ha contribuido tanto o más que otros al progreso sustancial que de hecho ha tenido lugar, y está lejos de ser el único en hacer semejantes promesas incumplidas. Mi opinión es sólo que el campo de la AI comenzó con una energía romántica que lo autorizaba a llevar a cabo algunas imponentes realizaciones intelectuales pero al mismo tiempo originaba un problema de credibilidad del que, en algún grado, todavía se resiente.

El romanticismo del temprano trabajo en la AI (algo corriente por otra parte), se refleja también en una fuerte tendencia a usar términos antropomórficos para describir sus técnicas. Cualquier discusión de inteligencia artificial es probable que incluya referencias a expertos, directores expertos, demonios, comunicación mediante pizarras, aprendizaje, procesos de inferencia lógica y fuentes de conocimiento. Personalmente tengo sentimientos encontrados sobre tal terminología. Por una parte, me divierte su uso tanto como a cualquiera. También puede decirse que tales términos son razonablemente descriptivos de los métodos mencionados. Ciertamente expresan mejor lo que describen que términos como verdad, encanto, extrañeza, con respecto de algunos fenómenos recientemente descubiertos en la física atómica. El lado negativo de esta tendencia a la antropomorfización es la acusación, a veces justificada, de que tal terminología es vaga y que su principal propósito es contribuir a que la técnica parezca más compleja y misteriosa de lo que realmente es. Puede apuntarse, por ejemplo, que un proceso de inferencia lógica, como los usados en recientes sistemas especiales, no es una operación comparativamente más complicada que el rápido transformador *dun-sounding* de Fourier.

Ha habido intentos durante todo este tiempo por acallar este entusiasmo exagerado (ver, por ejemplo, la cita 10). Quizá en aquellos primeros años un desbocado entusiasmo era más importante que la moderación. Sin embargo, creo que hoy hay un consenso según el cual es esencial establecer modelos y expectativas realistas si se quiere realizar el enorme potencial de este campo. Las realizaciones de la tecnología AI son ahora suficientemente sustanciales para que este campo no pueda ser únicamente considerado como algo del futuro.

---

<sup>9</sup> HERBERT A. SIMON. 1965. *The Shape of Automation for Men Management*, p. 96. Harper & Row.

<sup>10</sup> MARVIN MINSKY and SEYMOUR PAPERT, 1969. *Perceptrons: An Introduction to Computational Geometry*. MITT Press.

*La inteligencia artificial en el presente*

Así trasladamos nuestra atención al presente. Preferiría evitar de nuevo una confirmación de las sospechas usuales: no es que la exposición de los triunfos y las frustraciones del mundo de la AI de hoy no sea de interés, pero es un material asequible en otro lugar. En cambio, intentaré compartir con ustedes una visión personal del estado de este arte. Para hacer esto, citaré seis afirmaciones que se han hecho en meses recientes y declararé si estoy de acuerdo o no con cada una. Como a los investigadores de la AI, como yo mismo, se les acusa a menudo de ser imprecisos, contestaré a cada afirmación con una respuesta corta y precisa. Una vez hecho esto, sucumbiré a las tentaciones de mi educación y también daré una larga y vaga respuesta.

La primera afirmación es que «la AI es LISP». Mi *respuesta* breve es no.

LISP es un lenguaje de programación informática que se ocupa de la evaluación de expresiones simbólicas. Esas expresiones están compuestas principalmente de listas y funciones. Las listas por su parte son definidas como secuencias ordenadas de unidades, que pueden ser números, símbolos, expresiones, u otros. Por jerarquías definidas de listas pueden crearse estructuras de datos muy complejas; esas estructuras pueden estar interrelacionadas.

Antes de explicar mi breve respuesta negativa, permítanme decir una palabra a favor de LISP. Ante todo, el LISP es un elegante y satisfactorio lenguaje que, reflejando sus raíces matemáticas, obtiene un gran nivel de poder mediante la repetición. La repetición, como muchos recordarán por la teoría de los números elementales, nos permite hacer un número infinito de afirmaciones probando una proposición verdadera en un caso, para  $n=0$ , y luego derivando la proposición para  $n+1$  de la proposición para  $n$ . La mayoría de los matemáticos confían en la repetición, y esta técnica permite programas simples de LISP para realizar transformaciones relativamente poderosas. En segundo lugar, el LISP es uno de los pocos lenguajes de alto nivel que permite automodificar el código, lo cual es otro poderoso concepto en la teoría del cálculo. En tercer lugar, el LISP permite estructuras de información altamente flexibles que están sujetas a la representación de conceptos más complicados que los números. El LISP, además, ha cambiado durante los años más que la mayoría de los lenguajes. Muchos de esos cambios se han añadido significativamente a su modernización y capacidad.

Hay dos razones de por qué respondo no. La primera es que creo que hay una excesiva confianza y énfasis en el LISP. Un reciente informe del mercado de la inteligencia, proponiendo perspectivas a los inversores para arriesgar capital, declaraba que si el artículo en el que estaban considerando invertir no estaba escrito en LISP, entonces era casi seguro que no era verdadera inteligencia artificial. Aunque la mayoría de los investigadores serios

de la AI, incluso apasionados defensores del LISP, criticarían esa afirmación, no es un sentimiento poco frecuente.

Mi segunda razón es probablemente más sobresaliente y tiene que ver con algunos inconvenientes del LISP. El mismo informe del mercado decía que el LISP usado es ineficaz y caro, pero que esto no es toda la verdad. Es cierto que el advenimiento de micro-chips de bajo costo ha disminuido el precio de las máquinas de LISP. Han bajado decenas de miles de dólares. Sin embargo, como anotaré brevemente, el futuro de la AI dependerá notablemente de un proceso en paralelo —el ajuste de muchos procesadores para realizar muchas operaciones simultáneamente— más que del cálculo paso a paso. Los productos de un centenar de pequeñas máquinas unidas en paralelo son prácticos y pueden ser manufacturados a un precio razonable. Nadie propondría sin embargo manufacturar el producto que incluyen cien eficientes máquinas de LISP. El proceso en paralelo no era considerado cuando el LISP fue diseñado, y el LISP no es particularmente adecuado para él. En mi opinión, los procesos en paralelo serán más importantes para el futuro de la AI que los procesos de lista.

No se cuestiona que el LISP continuará siendo un lenguaje importante en AI y que continuará evolucionando. De todas formas, no creo que continúe jugando el papel dominante que ha tenido hasta la fecha.

La segunda afirmación que me gustaría examinar es «La inteligencia artificial son las técnicas de AI». De nuevo, mi breve respuesta es no.

La afirmación parece ser una tautología, el tipo de declaración sin sentido sobre las que Wittgenstein nos recomienda guardar silencio. Si definimos algunos términos, no obstante, empezará a tener significado. Por AI quiero significar el arte de crear máquinas que realizan tareas que se considera requieren inteligencia cuando son realizadas por humanos. Esta es la definición de AI que Minsky dio a mitad de los años sesenta. La he visto al menos veinte veces desde entonces, generalmente sin estar atribuida a nadie. Por técnicas de AI, entiendo a lo que la gente generalmente se refiere al decir «técnicas de AI», esto es, aquellas técnicas sobre las que las revistas aceptarían artículos. Es el mismo juego de técnicas sobre los que han estado trabajando los centros de investigación de AI en los últimos cinco años, pero no los que se investigaron antes de ese tiempo.

Un problema respecto a la segunda afirmación, como seguramente clarifica mi definición, es que el concepto de técnicas de AI es algo vago y arbitrario. Sugiero que, en cambio, hablemos de técnicas de extracción de características de alto nivel o técnicas de manejo experto, más que usar la expresión imprecisa de «técnicas de AI». Al menos tendríamos una mejor idea de aquello a lo que nos estamos refiriendo.

Pero existe un problema más importante respecto de esta afirmación. Indiferentemente de cualquier confusión sobre el significado de los términos, mi experiencia ha sido la de que el volumen de la tecnología sucesiva-

## ¿Qué es en definitiva la inteligencia artificial?

mente aplicada a los problemas de la AI, no tienen relación en absoluto con tales técnicas de AI, sino más bien con lo que llamamos técnicas de dominio relativo. Por ejemplo, mis colegas y yo estamos trabajando sobre una máquina que reconocerá el lenguaje humano sin restricciones de vocabulario. Mientras que hay técnicas de AI complicadas, tales como el manejo experto, y otros, el volumen de nuestra tecnología pertenece específicamente al dominio de la averiguación. Nosotros nos aprovechamos de los lingüistas, de la ciencia del lenguaje, de la psicoacústica, del procesamiento de signos, de la teoría de la información, de los factores humanos, de los montajes de computadores, del diseño de circuito integrado a gran escala y de otros campos.

Otra de las primeras definiciones de Minsky de inteligencia artificial era la de «un saco repleto de trucos». No podría estar más de acuerdo. Cada tarea específica que asociamos con una máquina inteligente requerirá un número diferente de técnicas, con métodos derivados principalmente de nuestra comprensión de cada problema. Oímos sobre los algoritmos de percepción generalizada que pueden reconocer cualquier tipo de patrón si está expresado en un lenguaje hablado, caracteres impresos, mapas de la superficie terrestre o huellas dactilares. Resulta que tales concesiones son absolutamente correctas; semejantes algoritmos reconocen, de hecho, cada tipo de patrón. Sólo que hacen todas esas tareas muy pobremente. Realizar bien cualquiera de ellas, con proporciones satisfactorias de exactitud, requiere un gran nivel de conocimiento profundamente imbuido en los algoritmos y específico para el dominio de la averiguación, tanto como que este aspecto de la tecnología está lejos de las genéricas técnicas de AI.

La tercera afirmación es: «La AI es un proceso en paralelo». Mi respuesta breve es sí.

Respondo que sí no porque la actividad de la AI deba ser paralela por definición, nadie más que ella debe usar el LISP, sino porque creo que los montajes en paralelo son el futuro, particularmente para los sistemas más complejos que tendemos a considerar como inteligencia artificial. El cerebro, como sabemos, más que ésta suple la lenta velocidad inherente a las células nerviosas en comparación con la silicona, con un paralelismo casi total y sin duda con un algoritmo superior.

Los procesos en paralelo pueden representar un uso más efectivo de los recursos que los procesos en serie. Nuestra completa confianza en los procesos en serie puede compararse ahora a una sociedad en la que un trabajo sólo puede ser hecho por una persona al tiempo.

Otra razón para usar procesos en paralelo es que aplicando suficiente fuerza bruta a un problema es con frecuencia necesario llevar a término el resultado deseado. Por ejemplo, la organización típica de un sistema especial tiene tres componentes: una masiva base de datos adecuada al área de especialización; un juego de roles sobre cómo ordenar, manipular y transformar

la base de datos, y un procesador de inferencia lógica que puede aplicar esos roles a esta base de conocimiento organizado. El cuello de botella en tales sistemas no es el costo de la memoria que acumula la información y los roles sino el tiempo real que requieren las operaciones de inferencia. Esto puede superarse a un precio razonable por montajes exhaustivos en paralelo, que la próxima generación de sistemas especiales apoyará indudablemente.

Un ejemplo de los comienzos de la AI concierne a los juegos de mesa. El algoritmo esencial para juegos de mesa como el ajedrez es tremendamente simple y bien conocido: generar una rama de árbol, tirar para abajo y podar. Curiosamente, durante los treinta años que las máquinas de juego de mesa han existido, hemos encontrado que lo más efectivo para mejorar su situación ha sido la aproximación a la fuerza bruta. Aplicando poderes más grandes de cálculo ha sobrevenido una mejora mayor que la que tiene la complejidad algorítmica. No es una coincidencia que el mejor juego de ajedrez por computadora es ahora incorporado por las supercomputadoras. La computadora que acabe ganando el campeonato mundial de ajedrez será indudablemente una computadora en paralelo.

Un reproche a los procesos en paralelo es que instalan un montón de microcomputadoras juntas y proporcionan vías para su comunicación y la coordinación de sus actividades. Otro reproche mayor aún es que lleva a una formación de lo que podríamos llamar subprocesadores sobre un único chip que está dedicado a alguna tarea particular. Un subprocesador no es una computadora programable sino un delicado circuito diseñado para desempeñar un algoritmo deseado. Los ejemplos de las funciones que tales chips podrían realizar incluirían procesos de señal, procesos de lógica deductiva, curvas de tiempo, modelos análogos y procesos de imagen, a la vez que realizaciones de algoritmos especiales para algún sistema particular. Empleando la costumbre de integración a gran escala, es posible poner una docena más de subprocesadores sobre un sólo chip. Un producto con 100 chips así, suministra de este modo el equivalente poder de cálculo de unos 1.000 microprocesadores con un costo base de menos de 1.000 dólares.

Este poder de cálculo desde luego, se dedica a una serie particular de algoritmos, que es por lo que considero que el montaje del sistema es tan importante como los mismos algoritmos. Un típico producto usado para ser una sola computadora con terminales apropiadas y soporte lógico. Más aún, veremos algoritmos distribuidos por un complejo y diversificado montaje en paralelo.

La cuarta afirmación es: «La AI es interdisciplinar». Mi respuesta breve es sí.

Esta afirmación parece obvia a primera vista. Creo, sin embargo, que la expresión de cualquier trabajo en la AI es necesariamente interdisciplinar, y el desafío que esto representa no es del todo apreciado. Como mencioné antes, en mi experiencia las técnicas de dominio relativo constituyen una par-

## ¿Qué es en definitiva la inteligencia artificial?

tipación mayor de la tecnología usada sucesivamente en sistemas de AI que la de las técnicas genéricas de AI. Esto tiene importantes aplicaciones para el tipo de esfuerzo que se requiere.

De aquí que una gran parte del reto que supone resolver un problema de AI implica una organización y comunicación. Investigar por equipos juntando media docena de disciplinas distintas no se da fácilmente, y una vez reunidos, existe el problema de que los miembros de cada equipo hablan diferentes lenguajes técnicos.

Una de las primeras personas en reconocer esto como un grave obstáculo para el desarrollo de las máquinas inteligentes fue Norbert Wiener. Su libro «Cibernéticas», publicado en 1948, fue una mirada extraordinariamente comprensiva al futuro del cálculo. En él señala que nadie desde Leibniz ha tenido «un completo dominio de todas las actividades intelectuales de su época». Declara el inevitable resultado: «Hay campos de trabajo científico... que han sido explorados desde diferentes ángulos de las matemáticas puras, la estadística, la ingeniería eléctrica y la neurofisiología, en los que cada una de las nociones recibe un nombre distinto en cada grupo, y en el que un trabajo importante se ha hecho por triplicado o cuadruplicado, mientras que se retrasa otro trabajo importante por la indisponibilidad en un campo de resultados que puede haberse convertido ya en clásico en el siguiente campo»<sup>11</sup>.

La quinta afirmación es: «La AI sirve de modelo para la inteligencia humana». La respuesta breve es no.

Sería beneficioso si pudiera pero en el presente no conocemos lo suficiente como para adelantar un gran trecho en la modelización del conocimiento humano. Hay excepciones. En nuestro trabajo sobre el reconocimiento del lenguaje hemos adelantado en lo que se conoce sobre los procesos de audición del caracol y el nervio auditivo. Encontramos que nuestros algoritmos de reconocimiento del lenguaje actúan mejor si atendemos al proceso de principio-a-fin del modelo de audición humana. Sigo, sin embargo, respondiendo que no, porque este proceso de principio-a-fin sólo representa una pequeña parte de todo el proceso implicado en el reconocimiento de lenguajes, bien por máquina, bien por una persona.

Una de las razones por las que conocemos algo sobre el proceso de frente-fin auditivo consiste en que los caracoles y nervios auditivos de animales de ensayo son relativamente accesibles. La mayor parte de los procesos cognitivos tienen lugar, sin embargo, fuera del alcance de nuestras pruebas. No solamente hay muchas células profundamente metidas en el cerebro, sino que ahora nos damos cuenta de que la mayoría de los procesos tienen lugar dentro de las células mismas por una serie de complicadas transformaciones químicas. El progreso en la comprensión del cerebro con-

---

<sup>11</sup> Ref. 1, p. 2.

tinuará, pero esto no nos promete más fuentes de ideas para nuevos algoritmos de AI.

Mi última afirmación es la contraria de la anterior: «Podemos aprender sobre el conocimiento humano a partir de los algoritmos de la AI». Mi respuesta breve es sí.

Antes de que podamos ensayar experimentos con humanos para probar teorías alternativas, necesitamos tener teorías para probar. Una de las mejores fuentes para esas teorías son las técnicas que hemos encontrado al trabajar con máquinas. El hecho de que un algoritmo trabaje en una máquina no prueba que la misma técnica sea usada en el cerebro, pero prueba que éste es un camino que el cerebro podría recorrer, y suministra una potente teoría que podría estar sujeta a la experimentación neurofisiológica.

### *La inteligencia artificial en el futuro*

Si dirigimos nuestra atención al futuro, debemos comprender que ya estamos a sus mismas puertas. Quizá lo mejor sería empezar diciendo lo que no espero ver en un futuro próximo. Estamos lejos todavía de la segunda predicción de Simon, de ser capaces de imitar en máquina la vasta categoría de la capacidad intelectual humana. Es dudoso, por ejemplo, que hacia el fin de este siglo las computadoras sean capaces de ver una película y escribir una crítica coherente.

El umbral en el que estamos es sin embargo muy significativo. Estamos ganando en habilidad para aplicar agudamente inteligencia selectiva a la máquina, o quizá diría inteligencia selectiva especializada, a un amplio campo de problemas. Me gustaría recalcar la palabra «especializada», porque las computadoras que se crearán en la próxima década no serán de amplio alcance intelectual. En cambio, veremos una proliferación de sistemas de áreas bien definidas de especialización; sistemas que tendrán un conocimiento magistral de alguna clase de enfermedades, o que tendrán habilidad para desarrollar ciertos tipos de estrategias de inversiones financieras, o que podrán ayudar a conducir una negociación complicada.

A pesar de mi énfasis en la palabra «especializada», es importante no subestimar el significado de lo automático en este tipo de especialización. Vemos el impacto que las computadoras están teniendo hoy sobre casi todas las áreas de esfuerzo. Añadiendo alguna inteligencia bien enfocada a nuestra computadora ya está bien demostrada la superioridad para adquirir diversas cantidades de información y conducir operaciones de repetición a altas velocidades sin cansancio, lo cual será una poderosa combinación. Nuestra meta, en realidad, no sería copiar la inteligencia humana en la próxima generación de computadoras, sino más bien concentrarnos solamente en las fuerzas propias de la inteligencia humana en el futuro más próximo.

## ¿Qué es en definitiva la inteligencia artificial?

Tan importantes al menos como la aparición de sistemas especiales son los esfuerzos por desarrollar computadoras inteligentes metidas en áreas tales como el reconocimiento del lenguaje, la comprensión de lenguajes naturales y la visión de cálculo. Los principales usuarios de máquinas inteligentes de la próxima generación no se han propuesto ser expertos en cálculo. Hay grandes esfuerzos, por ejemplo, por añadir alguna medida de inteligencia a los sistemas de instrucción de cálculo asistido. La próxima generación de máquinas de enseñanza intentará evaluar estudiantes para determinar sus fuerzas y debilidades, así como sus intereses, y proporcionará una formación al mismo tiempo entretenida y pedagógicamente adecuada a cada individuo.

Para comprometerse en alguna predicción por mi parte, creo que hacia el final de este siglo, la AI estará tan consolidada como lo están hoy los ordenadores personales. La mayor parte del soporte lógico será inteligente, al menos según los actuales modelos; se dará una interacción entre el principio inteligente —por parte de la máquina— y la determinación de los fines últimos por parte de los usuarios; se incrustarán bases de datos organizados, en el soporte lógico, y se acondicionarán a las telecomunicaciones; y máquinas de cálculo en paralelo a alta velocidad serán capaces de manejar cuanta información se necesite.

Si consideramos qué es lo que en verdad hace que un comportamiento parezca inteligente, una observación que podríamos hacer es que combina dos atributos. El primero es que el comportamiento parece razonablemente apropiado a la situación. El segundo que no es totalmente previsible. Las computadoras están ahora a punto de satisfacer esas dos condiciones. Por estar todavía sujetas en una vasta base de datos rápidamente cambiables y por aplicar a gran nivel procesos en paralelo para generar las interferencias y deducciones necesitadas, las respuestas generadas por tales sistemas combinarán sin duda la adecuación con la medida de lo imprevisible, dando al menos la impresión de comportamiento inteligente.

Añadir que consideraremos tales sistemas como realmente inteligentes es difícil de decir. Nuestros criterios sobre lo que constituye la inteligencia artificial están constantemente cambiando. En los cincuenta había un enorme entusiasmo porque las computadoras comenzaban a jugar al ajedrez y a dar jaques, aunque en un nivel elemental. Esto fue considerado AI en ese momento, si bien ya no consideramos como elementos de inteligencia artificial a las máquinas jugadoras de ajedrez tamaño bolsillo que se venden por 20 dólares y juegan mucho mejor. Hubo entusiasmo a finales de los cincuenta y principios de los sesenta cuando los programas de AI fueron capaces de suministrar teoremas y resolver problemas de cálculo. Hoy hay paquetes mucho más potentes que manejan ecuaciones consideradas bastante usuales, pero generalmente no se consideran ejemplos de AI. Aparentemente, tan pronto como entendemos un proceso suficientemente bien, co-

## Raymond Kurzweil

menzamos a considerarlo como un reto técnico y no como una ampliación de la inteligencia. Como Minsky apuntó recientemente, si un ser superior analizara el comportamiento humano para comprender al detalle cómo operamos, tampoco podría considerarnos muy inteligentes<sup>12</sup>.

Algunos observadores han sugerido en efecto que la inteligencia artificial está al filo de las posibilidades de la técnica, que sería definida como la esencia de los problemas de cálculo que no podemos resolver todavía.

Título original: What is Artificial Intelligence Anyway?

As the techniques of computing grow more sophisticated, machines are beginning to appear intelligent — but can they actually think?

Tomado de: *American Scientist*, Vol 73, May-June, 1985, pp. 258-264.

Traductor: José J. CABALLERO BONO.

---

<sup>12</sup> MARVIN MINSKY. 1984. Artificial intelligence. *A. I. Reporter*, Ap. p. 6.