



Novática, revista fundada en 1975 y decana de la prensa informática española, es el órgano oficial de expresión y formación continua de **ATI** (Asociación de Técnicos de Informática), organización que edita también la revista **REICIS** (Revista Española de Innovación, Calidad e Ingeniería del Software). **Novática** edita asimismo **UPGRADE**, revista digital de **CEPIIS** (Council of European Professional Informatics Societies), en lengua inglesa, y es miembro fundador de **UPENET** (UPGRADE European Network).

<<http://www.ati.es/novatica/>>
<<http://www.ati.es/reicis/>>
<<http://www.upgrade-cepis.org/>>

ATI es miembro fundador de **CEPIIS** (Council of European Professional Informatics Societies) y es representante de España en **IFIP** (International Federation for Information Processing); tiene un acuerdo de colaboración con **ACM** (Association for Computing Machinery), así como acuerdos de vinculación o colaboración con **AdaSpain**, **AIZ**, **ASTIC**, **RITSI** e **HispanLinux**, junto a la que participa en **Prolnova**.

Consejo Editorial

Joan Batlle Montserrat, Rafael Fernández Calvo, Luis Fernández Sanz, Javier López Muñoz, Alberto Lobel Ballori, Gabriel Martí Fuentes, Josep Moias i Bertran, José Onofre Montes Andrés, Olga Pallás Codina, Fernando Pira Gómez (Presidente del Consejo), Ramon Puigjaner Trepal, Miquel Sarries Griño, Adolfo Vázquez Rodríguez, Asunción Yturbe Herranz

Coordinación Editorial

Llorenç Pagés Casas <pages@ati.es>

Composición y autedición

Jorge Llácer Gil de Ranales

Traducciones

Grupo de Lengua e Informática de ATI <<http://www.ati.es/gt/lengua-informatica/>>

Administración

Tomás Brunete, María José Fernández, Enric Camarero, Felicidad López

Secciones Técnicas - Coordinadores

Acceso y recuperación de la información

José María Gómez Hidalgo (Openet), <jmgomez@yahoo.es>

Manuel J. María López (Universidad de Huelva), <manuel.maria@diesta.uhu.es>

Administración Pública electrónica

Francisco López Crespo (MAE), <flc@ati.es>

Arquitecturas

Enrique F. Torres Moreno (Universidad de Zaragoza), <enrique.torres@unizar.es>

Jordi Tubella Morgados (DAC-UPC), <jordit@ac.upc.es>

Análisis STIC

Marina Touriño Troitino, <marinatourino@marinatourino.com>

Manuel Palao García-Suñto (ASIA), <manuel@palao.com>

Base de datos

Isabel Hernando Collazos (Fac. Derecho de Donostia, UPV), <isabel.hernando@ehu.es>

Elena Davara Fernández de Marcos (Davara & Davara), <edavara@davara.com>

Escuela Universitaria de la Informática

Cristóbal Paraja Flores (DLSI-UPV), <cparaja@si.upv.es>

J. Angel Velázquez Iruñe (DLSI, URJC), <angel.velazquez@urjc.es>

Entorno digital personal

Andrés Marín López (Univ. Carlos III), <amarin@it.uc3m.es>

Diego Gachet Páez (Universidad Europea de Madrid), <gachet@uem.es>

Estándares Web

Encarna Quesada Ruiz (Alisys Software) <encarna.quesada@virat.com>

José Carlos del Arco Prieto (TCP Sistemas e Ingeniería), <jcarco@gmail.com>

Basión del Conocimiento

Juan Baiget Solé (Cap Gemini Ernst & Young), <juan.baiget@ati.es>

Informática y Filosofía

José Angel Olivas Varela (Escuela Superior de Informática, UCLM) <joseangel.olivas@uclm.es>

Kenneth Gierah Martin (Kennesaw University), <kgherah@gmail.com>

Informática Jurídica

Miquel Chover Sellés (Universitat Jaume I de Castellón), <chover@lsi.uji.es>

Roberto Vivó Herrando (Eurographics, sección española), <rvivo@dstc.upv.es>

Inteligencia del Software

Javier Dolado Cosin (DLSI-UPV), <dolado@si.upv.es>

Luis Fernández Sanz (Universidad de Alcalá), <luis.fernandez@uah.es>

Inteligencia Artificial

Vicente Bótti Navarro, Vicente Julian Inglada (DSIC-UPV) <vbotti_vinglada@dsic.upv.es>

Información Persona-Computador

Pedro M. Latore Andrés (Universidad de Zaragoza, AIPQ) <platore@unizar.es>

Francisco I. Gutierrez Vela (Universidad de Granada, AIPQ) <fgutier@ugr.es>

Lenguaje e Informática

M. del Carmen Ugarte García (BM), <cugarte@ati.es>

Lenguajes Informáticos

Oscar Geromte Ferrández (Univ. Jaime I de Castellón), <bellern@lsi.uji.es>

Inmaculada Coma Tatay (Univ. de Valencia), <inmaculada.coma@uv.es>

Lingüística computacional

Xavier Gómez Guinovart (Univ. de Vigo), <xggo@uvigo.es>

Manuel Palomar (Univ. de Alicante), <mpalomar@dlsi.ua.es>

Mundo estudiantil y jóvenes profesionales

Federico G. Mon Trotti (RITSI) <gnu.fede@gmail.com>

Mikel Salazar Peña (Área de Jóvenes Profesionales, Junta de ATI Madrid), <mikelxbo_uni@yahoo.es>

Profesión Informática

Rafael Fernández Calvo (ATI), <rfdc@ati.es>

Miquel Sarries Griño (Ayto. de Barcelona), <msarries@ati.es>

Redes y servicios informáticos

José Luis Marzo Lázaro (Univ. de Girona), <joseluis.marzo@udg.es>

Juan Carlos López López (UCLM), <juancarlo@uclm.es>

Seguridad

Javier Arellano Bertollin (Univ. de Deusto), <jarellano@eside.deusto.es>

Javier López Muñoz (ETSI Informática-UMA), <jlm@lcc.uma.es>

Sistemas de Tiempo Real

Alejandro Alonso Muñoz, Juan Antonio de la Puente Alfaro (DIT-UPM), <galton@puente>

Software Libre

Jesús M. González Barahona (GSYC-URJC), <jgb@gsyc.es>

Israel Herráiz Taberner (UAJ), <isra@herraz.org>

Tecnología de Objetos

Jesús García Molina (DS-UM), <jmolina@um.es>

Gustavo Rossi (LIFIA-UNLP, Argentina), <gustavo@sol.info.unlp.edu.ar>

Tecnologías para la Educación

Juan Manuel Doboero Beardo (UC3M), <doboero@inf.uc3m.es>

César Pablo Córcoles Brinco (UDC), <ccorcoles@uoc.edu>

Tecnologías y Empresa

Didac López Vilas (Universitat de Girona), <didac.lopez@ati.es>

Francisco Javier Cantais Sánchez (Indra Sistemas), <fjcantais@gmail.com>

Tendencias tecnológicas

Alonso Alvarez García (TID), <aad@tid.es>

Gabriel Martí Fuentes (Interbits), <gabi@atinet.es>

TIC y Turismo

Andrés Aguayo Maldonado, Antonio Guevara Plaza (Univ. de Málaga) <aguayo.guevara@lcc.uma.es>

Las opiniones expresadas por los autores son responsabilidad exclusiva de los mismos. **Novática** permite la reproducción, sin ánimo de lucro, de todos los artículos, a menos que lo impida la modalidad de © o *copyright* elegida por el autor, debiéndose en todo caso citar su procedencia y enviar a **Novática** un ejemplar de la publicación.

Coordinación Editorial, Redacción Central y Redacción ATI Madrid
Padilla 66, 3º, dcha., 28006 Madrid
Tfno. 914029391; fax. 913093685 <novatica@ati.es>
Composición, Edición y Redacción ATI Valencia
Av. del Reino de Valencia 23, 46105 Valencia
Tfno./fax. 963330392 <secreval@ati.es>
Administración y Redacción ATI Cataluña
Via Laietana 46, ppal. 1º, 08003 Barcelona
Tfno. 934129235; fax. 934127713 <secregen@ati.es>
Redacción ATI Aragón
Lagasca 9, 3-B, 50006 Zaragoza
Tfno./fax. 976235181 <seccara@ati.es>
Redacción ATI Andalucía <secreand@ati.es>
Redacción ATI Galicia <secregal@ati.es>
Suscripción y Ventas <<http://www.ati.es/novatica/interes.html>>, ATI Cataluña, ATI Madrid
Publicidad
Padilla 66, 3º, dcha., 28006 Madrid
Tfno. 914029391; fax. 913093685 <novatica@ati.es>
Imprenta: Derra S.A., Juan de Austria 66, 08005 Barcelona
Deposito legal: B 15.154-1975 - ISSN: 0211-2124; CODEN NOVAVC
Portada: En la cresta - Goncha Ariza Pérez / © ATI
Diseño: Fernando Agresta / © ATI 2003

Nº 206, julio-agosto 2010, año XXXVI

sumario

editorial

ATI protesta contra la marginación de las titulaciones informáticas > 02

documentos

Carta de la Junta Directiva General de ATI al Secretario de Estado de la SETSI > 02

en resumen

Casual Summer. Presentación del número de verano de 2010 > 03

Llorenç Pagés Casas

Informática en presente y futuro

Entrevista

Competitividad y desarrollo tecnológico en Europa: los próximos 10 años > 04

Entrevista a Neelie Kroes

Recursos humanos

Tecnologías de la información: ¿una profesión para hombres? > 08

Miren Idoia Alarcón Rodríguez, Luis Fernández Sanz

Software

Sistemas informáticos multilingües con lenguas españolas > 12

Ignacio Agulló Sousa

experiencias profesionales

Desarrollo de software

Una experiencia en transferencia de tecnología de Desarrollo de Software > 16

Dirigido por Modelos

Jesús J. García Molina, Jesús Sánchez Cuadrado, Joaquín Lasheras Velasco

Herramientas gráficas

Mapas tecnológicos > 21

Alonso Álvarez García

Inteligencia Artificial

Experiencias en el IIIA en investigación fundamental en Inteligencia Artificial > 25

Ramon López de Mántaras Badía

Experiencias en el IIIA en desarrollo de proyectos de Inteligencia Artificial > 31

Ramon López de Mántaras Badía

Protección de datos personales

Medidas de seguridad para ficheros de datos de carácter personal: Reflexiones y experiencias – La norma vs. la aplicación práctica > 35

Marina Touriño Troitino

Simuladores

Desarrollo de simuladores marítimos para FOREM-A > 38

Inmaculada Coma Tatay, Sergio Casas Yrurzum, Ausiàs Llorenç Palau, Jose Vicente Riera López

Tecnologías para la movilidad

Tecnologías móviles y ubicuas en un entorno docente universitario > 43

Pilar Moreno Díaz, Israel Herráiz Taberner, Jesús Sánchez Allende, Santiago Portela García-Miguel

Universidad y empresa

Experiencias de investigación en la universidad y la empresa > 48

José María Gómez Hidalgo

Referencias con firma > 53

lecturas

Historia de la Computación

El Test de Turing: historia y significado > 63

Giuseppe O. Longo

Impacto económico y social de la Informática

Leones supercomputerizados > 75

José Luis Cárpatos

asuntos interiores

Socios Institucionales

> 77

Monografía del próximo número: "Profesión informática"

Giuseppe O. Longo
Departamento de Ingeniería Eléctrica, Uni-
versidad de Trieste (Italia)

<longo@units.it>

El Test de Turing: historia y significado

© Mondo Digitale, 2009

Este artículo fue publicado en su versión original italiana por la revista **Mondo Digitale** de la asociación italiana AICA en su número de marzo de 2009, y posteriormente en inglés en el número 5/2009 de la revista **UPGRADE**. Se publica con los correspondientes permisos de **Mondo Digitale** y de su autor.

1. Introducción

"El método de la Ciencia es el método de todo el conocimiento mundano, aquel de la limitación o la verdadera ignorancia. Situado frente a la gran e ilimitada unidad de la Naturaleza, sólo podemos abordarlo en el pensamiento por medio de la selección de ciertos detalles y su aislamiento del resto (tanto voluntaria como inconscientemente)

Edward Carpenter,
Civilisation, its Cause and Cure

"Sólo el cerebro humano puede dar un significado a la habilidad ciega de los ordenadores para crear la verdad"

Karl Popper

En 1950 el gran matemático y especialista en lógica de origen británico Alan Mathison Turing (Londres 1912 - Manchester 1954) publicó un artículo titulado *Computing Machinery and Intelligence* en el cual describe un *Gedankenexperiment* (experimento conceptual) con el objetivo de determinar si una máquina (un ordenador) puede pensar. Al experimento, que apareció en el periódico *Mind*, se le conoce desde entonces como el *Test de Turing* y está inspirado en un juego, el llamado "juego de imitación" en el cual un interrogador intenta adivinar quien es el hombre y quien la mujer basando la decisión únicamente en sus respuestas a una serie de preguntas. Turing sugiere sustituir la mujer por una máquina y asignar al interrogador la tarea de discriminar entre el hombre y el ordenador. La versión modificada del juego permite a Turing abordar el problema *¿pueden las máquinas pensar?*, evitando las dificultades asociadas con las definiciones de los términos "máquina" y "pensar", cuyos significados deberían ser planteados con precisión desde el principio.

La utilidad de este experimento radica no tanto en la solución que proporciona al problema original (no se da por sentado que proporciona una solución) como en la posibilidad que ofrece para analizar conceptos tales como *mente*, *pensamiento* e *inteligencia*. En particular, el concepto de inteligencia, que en aquellos años se convirtió en terreno de encuentro de disciplinas tales como la lógica, matemática, psicología y filosofía, que de distintas formas se ocupaban del ser

Traducción: Mercedes Montes Rubio (Grupo de Trabajo de Lengua e Informática de ATI).

Resumen: parece muy difícil establecer un criterio objetivo que nos permita distinguir el pensamiento de una imitación mecánica del mismo. El punto crítico es la definición de pensamiento. En un artículo publicado en 1950, Turing sustituyó la pregunta, *¿Pueden pensar las máquinas?* por un problema circunstancial basado en un procedimiento conductista, dialógico: si un interrogador no es capaz de distinguir el comportamiento verbal de un ordenador de el de un ser humano, entonces se puede concluir que el ordenador puede pensar. El llamado Test de Turing ha originado muchos debates debido a la vaguedad de la formulación original y en conexión con los avances alcanzados por la inteligencia artificial en las últimas décadas.

Palabras clave: inteligencia artificial, juego de la imitación, lenguaje verbal, máquina de Turing, premio Loebner, problema cuerpo-mente, teoremas de Gödel.

Autor

Giuseppe O. Longo es profesor a tiempo completo en el Departamento de Ingeniería Eléctrica de la Universidad de Trieste donde enseña Teoría de la Información, una asignatura que él mismo ha contribuido a introducir en Italia. Ha llevado a cabo actividades de enseñanza e investigación en diversas universidades y otras instituciones científicas en Europa, Estados Unidos y Asia. En la actualidad su interés científico se dedica principalmente a la epistemología, la inteligencia artificial, ética robótica y el impacto social de la tecnología. Ha publicado varios libros y artículos sobre las consecuencias antropológicas y culturales de las nuevas tecnologías. Es novelista y dramaturgo y algunas de sus historias y novelas han sido traducidas al inglés y al francés entre otros idiomas. Dirige un curso sobre escritura creativa en el Máster en Comunicación Científica de la Escuela Internacional de Estudios Avanzados en Trieste y es miembro de distintas instituciones científicas.

humano, o mejor de sus capacidades cognitivas, consideradas como las habilidades humanas más características. Parecería importante poseer un criterio, si no para *definir* inteligencia al menos para reconocer su *presencia* tanto en organismos biológicos como en artificiales.

Debería enfatizarse que dos años antes, en 1948, Norbert Wiener había fundado la cibernética (*cybernetics*), una disciplina cuyo objetivo era el estudio del "control y comunicación en el animal y la máquina" bajo una perspectiva unificada. Después de unos años, en 1956, el matemático John McCarthy organizó una escuela de verano en el Dartmouth College, en Hanover, New Hampshire, durante el cual un reducido grupo de eruditos, entre los que se encontraban Claude Shannon y Marvin Minsky, representando distintos pero, a la vez relacionados, campos de estudio, eligieron el nombre de "Inteligencia Artificial" (IA) para indicar una familia de áreas de investigación bastante diversas cuyo objetivo común era producir simulaciones o emulaciones de comportamientos inteligentes en ordenadores.

En las secciones siguientes describimos el Test de Turing y lo examinamos críticamente para revelar sus méritos y sus limitaciones. Aunque la discusión es autosuficiente, se harán referencias constantes al artículo original de Turing, que aún es de gran interés y al cual referimos al lector (ver **figura 1**).



Figura 1: Alan Turing.

Breve biografía de Alan Turing

Alan Mathison Turing nació en Londres, el 23 de junio de 1912. Sus padres procedían de clase media alta y su padre continuó esa tradición como administrador en la administración pública India. Con la ayuda de un manual, el pequeño Alan aprendió a leer por sí mismo en tres semanas. Fue enviado a internados y después de algunos problemas tempranos de adaptación social, se distinguió en matemáticas y en ciencias.

Sus excepcionales habilidades matemáticas le fueron reconocidas primeramente en sus años de instituto (1931–1936) en el King's College de la Universidad de Cambridge. Su trabajo matemático más importante, *On Computable Numbers, with an Application to the Entscheidungsproblem*, fue escrito en Cambridge en 1936. En este artículo Turing introduce el concepto de una máquina teórica (hoy conocida como la "Máquina de Turing") que encarna el concepto del ordenador con programa almacenado. En 1936 se le concedió una beca de investigación para visitar la Universidad de Princeton, donde escribió una disertación doctoral.

Durante la Segunda Guerra Mundial entró en la Government Code and Cypher School en Bletchley Park, situada entre Oxford y Londres, donde jugó un importante papel en el diseño del equipo y el desarrollo de técnicas para romper el sistema de cifrado alemán *Enigma*. Estuvo próximo al grupo de analistas de cifrado que diseñó Colossus, el primer ordenador digital electrónico con programa almacenado, aunque no participó directamente en su construcción. Su trabajo en Bletchley le proporcionó una valiosa experiencia en electrónica, así como un equipo de cálculo de propósito especial que le sería de gran ayuda después de la guerra.

En 1945 se trasladó al Laboratorio de Física Nacional en Teddington, Inglaterra, para asumir responsabilidades en el diseño de un ordenador electrónico, el ACE (*Automatic Computing Engine*). Durante varios años se utilizó un prototipo de ACE en diversas funciones de importancia, incluyendo el diseño de un avión. Hacia el final de la guerra le fue concedido el premio OBE (*Order of the British Empire*) por los brillantes servicios prestados en favor de Gran Bretaña. En 1951 Turing se convierte en Miembro de la Royal Society. Además de los artículos referidos en estas páginas, Turing escribió otros dos informes importantes sobre el mismo tema: *Intelligent Machinery* (1948) y *Digital Computers Applied to Games* (1953), por lo que se convirtió en un pionero en inteligencia artificial, más aún cuando comenzó a investigar las relaciones entre las máquinas y la neurología y psicología de los seres humanos. En realidad, desde 1952 hasta su muerte en 1954, dedicó su genio a la formalización matemática de la biología, en particular a la morfogénesis, explorando entre otras cosas la conexión entre las formas vegetales y los números de Fibonacci.

En 1952 fue arrestado por practicar públicamente la homosexualidad, por aquel entonces un delito en Gran Bretaña. Los tribunales resolvieron una atroz pena, la castración química, que le volvió impotente. La depresión que esto le produjo fue probablemente la razón por la que Turing se suicidó en su casa de Manchester, comiendo una manzana envenenada con cianuro de potasio. Aunque refutada por su madre, la hipótesis del suicidio es casi segura. En su libro *Zeros and Ones*, Sadie Plant sugiere que Apple Inc. eligió como logotipo la manzana mordida en honor a Alan Turing, pero la compañía nunca ha confirmado ni desmentido esta conjetura.

2. El juego de la imitación

El juego de la imitación consiste en un intercambio lingüístico entre tres personas: un hombre (A), una mujer (B) y un interrogador (C) que puede ser de cualquier sexo. 'C' permanece en una habitación aparte y puede preguntar a A o B cualquier tipo de pregunta. Basándose únicamente en sus respuestas, C tiene que determinar si A es el hombre y B la mujer o viceversa. Tanto A como B intentan ser identificados como la mujer, por consiguiente B se esfuerza para ayudar al interrogador y A intenta engañar de la mejor manera que pueda a C dando respuestas para que piense que las daría una mujer. C tiene a su disposición sólo las respuestas y no otro tipo de pistas como su aspecto, estilo, voz, letra y cosas así, que harían más fácil la identificación. Por lo tanto, C se puede comunicar con A y B sólo mediante teletipo o un medio igualmente neutral.

Turing imaginó modificar el juego de la imitación, reemplazando A por una máquina y se preguntó si en estas condiciones el interrogador C emitiría un juicio erróneo con la misma

frecuencia que en el juego de la imitación clásico. C se equivoca cuando supone que el ser humano es A y no B, en ese caso la máquina gana y se la proclama capaz de pensar.

Ahora examinaremos de cerca el juego de la imitación y la variante propuesta por Turing. En primer lugar deberíamos responder a la siguiente pregunta: ¿Que demuestra el juego de la imitación cuando deja de ser un pasatiempo para convertirse en un *Gedankenexperiment*? (ver **cuadro 1**).

La idea subyacente es que las respuestas de un hombre podrían distinguirse de las de una mujer, lo que implica que el modo masculino de pensar difiere de forma apreciable del de la mujer. Pero, aunque la inteligencia femenina pudiera ser diferente de la inteligencia masculina, el juego de la imitación se basa en una segunda suposición, por ejemplo que un hombre podría sin embargo simular la inteligencia femenina.

No obstante, uno no debería olvidar que el juego implica tres jugadores: la identificación

de A y B se asigna a un juez, el interrogador, y si el juez es suficientemente inteligente puede tomar la decisión correcta, con la ayuda de B, a pesar de los esfuerzos de A por llevarlo por el camino equivocado.

Esto provoca un problema: ¿Es una cuestión de decidir si A es más listo que B o de decidir si C es suficientemente listo? Quizá el juego es una prueba para C. Si se cambia el interrogador, pero A y B son los mismos, el resultado del juego podría cambiar. Se puede también imaginar repetir el juego con el mismo juez, jugando varias rondas y produciendo una evaluación estadística de los resultados. De manera alternativa, se puede imaginar que hay varios interrogadores, cada uno de ellos tomando una decisión basada en las respuestas dadas por A y B a la misma serie de preguntas. Sus decisiones pueden ser diferentes y de nuevo se puede establecer una estadística.

En la versión original del juego el juez intenta discriminar entre un hombre y una mujer. Esta es una dicotomía clásica, pero se podría

Dejemos que A, B y C (el interrogador) sean los participantes en el juego de la imitación. He aquí algunas de las posibles variantes:

1. A hombre, B mujer, C humano (juego original).
2. A máquina, B humano, C humano (Variante de Turing).
3. A máquina, C humano (primera variante con dos jugadores).
4. A humano, C humano (segunda variante con dos jugadores).
5. A hombre, B mujer, C máquina.
6. A humano, C máquina (tercera variante con dos jugadores).
7. A máquina, C máquina (cuarta variante con dos jugadores).

En todas las variantes el interrogador C debería establecer la naturaleza de A. En la primera variante el interrogador sabe que los jugadores son un hombre y una mujer, pero no los ve. En todas las demás variantes el interrogador puede o no ser consciente de las posibles identidades de A (y B, si está presente). El mismo Turing no deja claro si en la variante 2 el interrogador sabe que uno de los dos participantes es una máquina, aunque es bastante obvio que debería saberlo. Considerando la variante 3, si el interrogador no sabe que A puede ser una máquina, podría concluir siempre que A es humano, posiblemente estúpido, insensato o bromista. Esto ocurre si cabe más en la variante 4, donde los errores cometidos por el humano A para desconcertar al interrogador poseerían la cualidad humana de *pertinencia* (ver **cuadro 2**). Por consiguiente para hacer el juego más interesante el interrogador debería ser consciente de la posible naturaleza del competidor o competidores. En los casos 5, 6 y 7 el proporcionar al interrogador la información relativa a la posible naturaleza de los contrincantes es problemático, ya que es difícil prever las preguntas que C podría plantear para dirigir el examen. Es más, uno podría preguntarse si el mismo C ha aprobado el Test de Turing. Para resumir, estas tres últimas variantes no merecen gran atención, pero quizá confirman que el test tiene que ver con el interrogador tanto como con A y B e incluso más.

Al menos desde nuestro punto de vista sólo una inteligencia humana puede juzgar a una inteligencia humana (o exótica). Esto parece confirmar que tenemos en más alta consideración la inteligencia humana con respecto a la de las máquinas, tanto como cuando en los exámenes del colegio consideramos que la inteligencia (o competencia) del examinador supera (o debería superar) a la del alumno. No sé si este antropocentrismo es inevitable como me lo parece en este momento. Sin embargo, en la variante 7, tenemos una visión desde una nueva perspectiva. Los humanos jugarán al juego con los humanos y con las máquinas hasta que las máquinas sean capaces de jugar como los humanos o incluso mejor.

Llegado este punto es plausible que los humanos jueguen sólo con humanos y las máquinas sólo con las máquinas. Esto ya ocurre con los programas de ajedrez y ha venido ocurriendo durante muchos tiempo en las carreras: caballos contra caballos, coches contra coches y déjenme añadir, ¡hombres contra hombres y mujeres contra mujeres!. Pero cuando las máquinas hayan superado a los humanos, quizá los interrogadores más cualificados para todas las variantes sean las máquinas y tendremos que adoptar un punto de vista centrado en la máquina...

La idea subyacente es que, al menos hasta ahora, no es posible formalizar el concepto general de inteligencia. Podríamos estar de acuerdo en asignar límites bien establecidos a la competición entre humanos y máquinas, estableciendo una clase de combate, como si fuera administrado por parámetros cuantificables, por ejemplo, en el caso de la conversación, la riqueza de vocabulario, diversidad en la construcción de las frases, corrección lógica, frecuencia de metáforas y por supuesto duración de la competición. Aunque surge un problema de inmediato: ¿Cual sería el rango de la conversación? ¿O debería extenderse la conversación hacia cualquier tema?

Es fácil comprender que en el último caso ningún programa de conversación existente puede pasar la prueba si el combate dura más de un breve periodo de tiempo (ver **cuadro 5**), mientras que si limitamos el rango de la conversación drásticamente, el éxito no prueba nada, ya que una verdadera conversación entre seres humanos es un proceso interactivo, constructivo, autopoietico, arborescente, en parte aleatorio y en parte determinístico que refleja los acontecimientos de la vida muy de cerca, imitando su volubilidad e impredecibilidad, pero también su lógica compleja abierta hacia el mundo. Como dice Licata "el proceso de la cognición humana está ligado a la impredecible y *compleja frontera* entre un sistema y el entorno que le rodea".



Una extraña variante del juego de la imitación. El examinador le dice al ordenador: "¡Lo hiciste realmente muy bien!" y añade "Ahora mismo ni siquiera sé quién soy". El título de la parte de arriba dice: "Ganarás puntos extra en el test de Turing si convences al examinador de que es un ordenador". (Chiste de Randall Munroe, libremente disponible en <<http://xkcd.com/329/>>.)

Cuadro 1. Algunas posibles variantes del Juego de la Imitación.

imaginar otras parejas de participantes dicotómicas: uno blanco y otro negro, un italiano y un chino, un poeta y un matemático (el sexo de todos ellos es arbitrario). Y en concurso se puede hacer más complicado, contrastando por ejemplo un matemático chino con un poeta italiano. De estos ejemplos podría ser evidente que C, que sabe a que clases o categorías pertenecen A y B, decide verdaderamente sobre la base de las respuestas que recibe, pero está también influenciado por sus creencias, prejuicios, experiencia en el mundo, etc.

En otras palabras, incluso cuando los mensajes se intercambian por teletipo, las respuestas contienen elementos que no son puramente sintácticos, por ejemplo internos al lenguaje, pero también y marcadamente semánticos, en el sentido de que están ligados a la naturaleza general de A y B, no sólo a sus propiedades cognitivas. Por supuesto la presencia y la cantidad percibida de tales elementos semánticos dependen de las respuestas, y por lo tanto de las preguntas del interrogador y de sus expectativas sobre las respuestas.

Si fuera posible preguntar cuestiones puramente sintácticas (por ej. de naturaleza lógica o matemática) demandando recibir respuestas puramente sintácticas, quizá sería imposible distinguir un hombre de una mujer, aunque recientes descubrimientos neuropsicológicos parecen destacar notables diferencias cerebrales entre los sexos, y por tanto posibles diferencias de pensamiento. Por supuesto nos referimos aquí a un hombre y a una mujer elegidos al azar y desconocidos para el interrogador.

Resultaría que la distinción entre sexos se podría obtener sólo basándose en los hechos *no* reducibles a secuencias de símbolos, o mejor basándose en los hechos extralingüísticos que sin embargo pueden dejar una huella en las secuencias de símbolos.

Indicándolo de otro modo, existe la posibilidad de que las diferencias (de experiencia, rol social, características biológicas, etc.) entre hombre y mujer conlleven diferencias detectables en las respuestas que A y B dan a *ciertas preguntas*, diferencias que no se pueden inventar por imitación. Esto indicaría que el pensamiento no puede ser aislado o abstraído completamente del resto de los seres humanos, por ejemplo del cuerpo y sus necesidades y por tanto del mundo. Resumiendo, si las diferencias entre las respuestas dadas por A y B procedieran de las diferencias entre A y B, un interrogador suficientemente inteligente siempre tomaría la decisión correcta, ya que A nunca podría imitar a B. Para imitar a B, A debería ser B.

3. Cuerpo y mente

Turing tiene una posición diferente. En su opinión, el pensamiento puede ser separado

de su soporte. En realidad, el juego tiene sentido sólo si se aísla a C de A y B, de otro modo distinguiría al hombre de la mujer o, en la variante, al hombre de la máquina a la primera, sólo con mirarlos. Turing amplía esta condición a la prohibición de considerar las características corporales o funciones (belleza, habilidad para el baile, etc.) que discriminarían en contra de la máquina y que en realidad, para él, no añaden nada a la inteligencia del poseedor.

Por lo tanto Turing, y muchos otros, alimentaron el prejuicio de que las sentencias de símbolos bien construidas son suficientes para expresar y reconocer la inteligencia. El cuerpo no tiene nada que ver con la inteligencia. Tal reduccionismo "mentalista" o informativo se confirma por la lista de actividades humanas que Turing juzgó adecuadas para la implementación en máquina. Tales actividades eran aquellas que no implicaban un contacto con el mundo exterior: ajedrez, matemáticas, criptografía, quizá la traducción de un idioma a otro. Pero todos los factores corporales fueron excluidos. En una emisión de la BBC declaró:

"Estoy seguro y espero que no se hagan muchos esfuerzos para construir máquinas que posean características específicamente humanas (a excepción de las intelectuales), tales como la forma del cuerpo. Para mí dichos esfuerzos se muestran insignificantes y sus resultados tendrían aproximadamente la misma desagradable calidad de las flores artificiales. Me parece que la construcción de una máquina pensante es un intento que pertenece a una categoría diferente".

Sin embargo, en el artículo de 1950 Turing contradice esta postura parcialmente extendiendo la actividad de las máquinas inteligentes a la conversación, de hecho identificando pensamiento con habilidad lingüística, quizá sin darse cuenta abriendo un vacío a través del cual la máquina entra en contacto con el mundo exterior. En realidad el interrogador puede preguntar cuestiones arbitrarias, en particular cuestiones relacionadas con hechos del mundo (el mundo que habitamos los humanos). El vínculo entre lenguaje y mundo aparece cuando el lenguaje deja de ser autoreferencial (por ejemplo un conjunto de símbolos puramente sintácticos que forman un formalismo abstracto, según la definición de Hilbert) para cargarse de semántica y aludir a objetos fuera del lenguaje. Tal vínculo, por virtud del cual el lenguaje crea, modifica y organiza el mundo y está condicionado por él, afecta ahora también a la máquina (ver **figura 2**).

Como vemos, existe una cierta confusión. Por un lado Turing admite que la inteligencia humana proviene de la interacción con el mundo, por tanto del cuerpo, que es el medio



Figura 2. David Hilbert (1862-1943).

a través del cual esa interacción tiene lugar. Por otro lado está convencido de que una máquina de estados discretos, la llamada "máquina de Turing" (ver **cuadro 2**), aunque no posea ni órganos sensitivos ni accionadores que le permitan interactuar con el mundo exterior, puede sin embargo utilizar palabras como un ser humano y con tal habilidad que sea capaz de engañar al interrogador.

El juego de la imitación está basado en interacciones lingüísticas, por ejemplo en marcas exteriores de pensamiento, siendo este último inaccesible. En este sentido es un criterio conductista. Las secuencias verbales están ahí para exhibir su existencia, mientras que el pensamiento, del cual (podrían) emanar, tiene una mucho más dudosa existencia. En realidad en la primera mitad del siglo XX el pensamiento fue desterrado de las reflexiones de psicoanalistas y filósofos. Así que Turing aceptó la postura conductista que mantenía su prevalencia en aquella época, pero a pesar de esto no negó la existencia de pensamiento, de hecho contribuyó a formular un problema aún vivo en las ciencias cognitivas, por ejemplo la relación entre distintos niveles de funcionamiento y descripción del pensamiento. A decir verdad, entre las objeciones que consideró¹ merece atención la siguiente: ¿No podrían acaso las máquinas llevar a cabo algo que debiera ser descrito como pensar pero que fuera muy diferente de lo que hace el ser humano?

En resumen, incluso si la actividad de superficie fuera la misma, por ejemplo la expresión lingüística, los mecanismos subyacentes podrían ser muy diferentes en máquinas y humanos. Por otra parte, concluye Turing obediendo de nuevo al conductismo, si una máquina lo hiciera bien en el juego de la imitación, la diferencia en los mecanismos subyacentes podría obviarse.

4. La variante de Turing

Consideremos ahora la variante propuesta por Turing, en la cual A es una máquina y C

Como he mencionado por encima, la "máquina" que Turing tenía en mente es una máquina digital de estados finitos, por ejemplo un ordenador concreto, o mejor un prototipo abstracto de un ordenador en concreto, llamado "la máquina de Turing" (MT). Es un dispositivo ideal que opera en una cinta de longitud ilimitada, dividida en celdas. La cinta se puede mover a la celda de la derecha o de la izquierda al tiempo. Gracias a una cabeza lectoescritora, se puede leer, escribir o cancelar un símbolo de un alfabeto finito en cada celda. La MT puede llevar a cabo cualquier operación que pueda ser descrita a través de un algoritmo y en este sentido es una máquina universal. Sin embargo, Turing demostró que la MT no es capaz de encontrar una solución a todos los problemas matemáticos y asoció ese descubrimiento a los teoremas de Gödel.



Kurt Gödel (1906-1978).

Según el biógrafo de Turing, Andrew Hodges, el hecho de que la MT se comunique con el mundo sólo a través del exiguo canal del programa proporciona un indicio del estilo de vida ideal de Turing: "una vida donde sería dejado a solas en una de sus habitaciones, interactuando con el mundo exterior sólo por medio del razonamiento racional". Este comentario confirma la posible influencia directa de la experiencia personal en la creación de ideas científicas. Sin embargo, en contra de esta interpretación, en un pasaje de su tesis doctoral en Princeton (1938) Turing menciona una habilidad humana muy importante por medio de la cual un matemático distingue los temas de interés de todos los demás. Según Turing, la existencia de tal habilidad contradice la expandida opinión de que la única actividad de un matemático es determinar la veracidad o falsedad de proposiciones.

La habilidad a la que se refiere Turing depende del contacto con la realidad, por lo tanto las máquinas no pueden poseerla. También el siguiente pasaje del artículo de 1950 señala la importancia que Turing atribuye al contacto con el mundo:

"La visión popular de que los científicos avanzan inexorablemente de un hecho bien establecido a otro hecho bien establecido, no siendo nunca influenciados por ninguna conjetura mejorada, es bastante errónea. Siempre que se aclarare cuales son los hechos demostrados y cuales son conjeturas, no puede existir ningún daño. Las conjeturas son de gran importancia ya que sugieren líneas de investigación prácticas".

Por lo tanto, parece que Turing oscila sin decidirse, entre una anhelada e ideal autoreferencialidad del lenguaje, en particular del mundo lingüístico de la matemática y la lógica, y una comunicativa y activa apertura del lenguaje hacia el mundo real. Un ejemplo relevante de auto referencia en matemáticas es ofrecido por el llamado "Programa de Hilbert". Al comienzo del siglo 20, el matemático alemán David Hilbert había afirmado la posibilidad de construir las matemáticas en su totalidad por medio del método axiomático, utilizando un número finito de proposiciones de inicio (axiomas) y un número finito de reglas de inferencia. Los objetos matemáticos se definirían entonces de un modo implícito y abstracto, sin ninguna referencia semántica al mundo exterior. En otras palabras, Hilbert se abstuvo deliberadamente de "definir" objetos matemáticos, que de este modo serían "nombrados" y entonces conectados los unos a los otros para ser caracterizados por sus relaciones mutuas. Después, empleando las reglas, los teoremas se derivarían de los axiomas a través de un procedimiento mecánico.

Pero en 1931 el logista Kurt Gödel demostró dos teoremas que establecen límites muy precisos al programa de Hilbert. El primer teorema afirma que una teoría formal capaz de expresar aritmética elemental no puede ser a la vez consistente y completa. En particular para cualquier teoría formal consistente que demuestre ciertas verdades aritméticas fundamentales, existe una afirmación que es cierta pero no demostrable en la teoría. El segundo teorema demuestra que es imposible demostrar la consistencia de la teoría dentro de la propia teoría.

Estos teoremas no implican demasiado una derrota de las matemáticas o del razonamiento humano, sino más bien un fracaso en la esperanza de Hilbert de aislar las matemáticas de cualquier referencia externa. Un matemático no se limita a la manipulación de símbolos, también considera el contexto matemático y la importancia práctica de los problemas, realiza elecciones y detecta conexiones entre los problemas y los resultados aparentemente remotos.

Confrontado con una afirmación indecidible, por ejemplo, el matemático recurre a la semántica para extender el sistema formal a uno más grande, uno que al final contenga otra afirmación indecidible, y por lo tanto que requiera una mayor extensión, y así sucesivamente. Únicamente los sistemas formales muy pobres, aquellos incapaces de formalizar la aritmética, son consistentes y completos, mientras que aquellos sólo un poco más potentes ofrecen una apertura semántica indefinida hacia el mundo. Esto discrepa con la noción de que las matemáticas son una tautología finita, por ejemplo una explicitación mecánica de proposiciones verdaderas (teoremas) implícitas en los axiomas.

En 1936 Turing publicó *On Computable Numbers, with an Application to the Entscheidungsproblem*, un artículo que ilustra la naturaleza atómica de los procedimientos computacionales por medio de la máquina de Turing, MT, que aparece en una proposición, la llamada tesis Church-Turing, según la cual *todos los procesos de cerebro derivan de un substrato computable* (esta es una versión excesivamente simplificada, para más información ver referencia a Hofstadter). Como consecuencia, los procesos del pensamiento humano serían reproducibles por medio de un autómeta finito como la MT. Esta proposición es la base de la versión fuerte de IA y encuentra su justificación en la finitud de la memoria humana. No obstante, en 1969, Gödel descubrió un error en el argumento de Turing, que minaba su versión finita del cálculo humano. Según Gödel, la mente humana no es estática, sino que se desarrolla incesantemente durante su funcionamiento. Aunque en cada etapa de su desarrollo

el número de estados diferentes es finito, no existe motivo por el que este número no debería divergir en el curso de ese desarrollo. Al principio la mente está basada en un número finito de axiomas dados, que podrían corresponder a las habilidades iniciales especificadas por la información genética, y después, gracias a las perturbaciones generadas por la interacción con el entorno, por ejemplo, gracias a la información aportada por la experiencia, el sistema mental se hace cada vez más complejo y elude nuestro análisis formal, igual que un sistema suficientemente potente, como el especificado anteriormente, elude cada reduccionismo axiomático.

Lo que caracteriza a la mente humana es la habilidad para producir novedades, romper simetrías, aprender, en resumen desempeñar todas esas funciones creativas tan bien descritas por Gregory Bateson en su trabajo sobre las categorías lógicas del aprendizaje y la comunicación. La interacción mente-entorno es generativa y dinámica y no puede ser confinada en una descripción *final* de la mente por medio de una o más MT. La IA posee estructuras lógicas cerradas y tiene éxito dentro de los micromundos especializados o en abordar problemas-juguete², mientras que la mente muestra una estructura lógica abierta, modifica constantemente las reglas del juego y se adapta a las interacciones que cambian con el macromundo. De hecho, también la física clásica, que aparentemente posee una estructura cerrada expresada por el inexorable desarrollo de fenómenos físicos de acuerdo a eternas y estrictas leyes, es incapaz de explicar las novedades, la emergencia de información y la historia. Sólo la llegada de nuevos conceptos y perspectivas, tales como la complejidad, emergencia e información, han proporcionado una imagen abierta, histórica y evolucionaria de la realidad física como filtrada por nuestra cognición. En resumen, esto nos permite reconocer el entrelazado entre la física y la mente que Gregory Bateson había descrito en términos filosóficos y cuantitativos en su "*ecología de la mente*".

No obstante, uno no debería concluir que Turing fuera tan ingenuo para creer que una MT o incluso una familia de MTs pudiera ser un modelo adecuado de la mente humana. Tanto en su artículo de 1936 como en el de 1948, Turing afirma que la organización de la máquina y su funcionamiento debería estar sujeto a sufrir cambios. El primer cambio funcional está representado por *oráculos*, dispositivos que pueden influenciar el cálculo de una MT desde el exterior. Existe una conexión entre los oráculos y la inteligencia natural. El programa grabado en el genoma es modificado progresivamente por la experiencia. Refiriéndose a la organización inicial y a la habilidad de obedecer programas, Turing afirma que la disciplina por sí sola no es suficiente y que la inteligencia requiere también *iniciativa*: el comportamiento inteligente, no sólo consiste presumiblemente en partir desde un comportamiento totalmente disciplinado implicado en el cálculo, sino desde uno mucho más ligero, que no da pie al comportamiento aleatorio o a bucles repetitivos sin sentido.

Por consiguiente Turing se da cuenta de que el modelo MT es demasiado pobre para describir procesos cognitivos. Paradójicamente, la máquina debería ser capaz de cometer errores. Si es infalible no puede ser inteligente. En realidad, un matemático se somete a un largo entrenamiento, durante el cual, en general, comete muchos errores. ¿Porque deberían ser las cosas diferentes en el caso de una máquina?

Tal propensión a los errores permite evitar lo que es repetitivo y mecánico en ciertos programas orientados a imitar las habilidades de conversación de los humanos. Mientras que los errores cometidos por las máquinas a la hora de contestar preguntas del Test de Turing pueden ser arbitrariamente descartados y patentes, en los errores cometidos por los humanos siempre se encuentra una cierta *relevancia* y una noción más amplia y confusa que la *exactitud*. Por tanto, la relevancia pertenece también a algunos tipos de error. Sólo es importante que las respuestas erróneas no den la impresión de arbitrariedad o locura. La noción de relevancia puede ayudarnos a obtener una (vaga) idea del entendimiento mejor que la noción (absoluta) de exactitud (ver **cuadro 5**).

Cuadro 2. La máquina, Hilbert y Gödel.

un humano. En este caso tampoco esta muy claro cual sería el significado del juego. Uno se puede imaginar que repitiendo el test varias veces, por ejemplo jugando varias rondas, a veces C se equivoca y a veces acierta en su juicio, es decir, a veces la máquina aprueba y a veces suspende el test. Es fácilmente comprensible que si las preguntas pueden ser arbitrarias, más tarde o más temprano la máquina cometerá fallos reveladores, siendo suficiente con preguntar cuestiones sobre el mundo en general, excediendo de esta forma el área de competencia de la máquina. Pero esto es cierto también para los humanos. Quizá, como en todos los juegos, las limitaciones y condiciones deberían introducirse en primer lugar. Asumiendo ahora que hemos establecido la duración, restricciones, etc. del test, podemos preguntarnos si pasar el test es

una condición suficiente para poseer inteligencia. ¿O es una condición necesaria? ¿O es una simple visualización de habilidad, un pasatiempo, como mucho un tema para meditar? Para obtener algunas pistas al respecto, preguntémosnos cuestiones similares sobre los seres humanos, a quienes en general no dudamos en otorgar inteligencia.

5. Inteligencia humana

Asumamos que un humano es sometido a una prueba concreta, que consiste en jugar varias partidas de ajedrez. ¿Podemos decir que si gana más del 50% de las partidas es inteligente? ¿Y si gana menos del 50% no es inteligente o es menos inteligente? Poca gente, creo, suscribiría unas conclusiones tan radicales. La mayoría simplemente estaría de acuerdo en que el individuo es bueno o malo

jugando al ajedrez en relación con su oponente. Asumamos ahora que el sujeto es sometido a un test más difícil, consistente en escribir una historia. También en este caso, como en el anterior, el éxito no es una condición suficiente para proclamar al individuo inteligente, ni siquiera es una condición necesaria. En este caso, sin embargo, emitir un juicio sobre el éxito sería mucho más difícil.

Mucha gente está considerada inteligente (basándose en evidencias circunstanciales tales como el éxito en la vida profesional, la capacidad de establecer relaciones personales satisfactorias o el talento en la investigación científica), y aún así no pueden jugar al ajedrez o escribir historias. Quizá podrían involucrarse en esbozar una historia, pero su éxito o fracaso no añadirían o restarían un



Figura 3. Joseph Weizenbaum (1923-2008).

ápice a su inteligencia, sólo sería otra evidencia circunstancial, válida para aquellos que piensan que las historias son importantes. Una vez más se acentúa el punto de vista del juez. En cambio, mucha gente que juega bien al ajedrez o escribe excelentes historias, a menudo demuestran bastante estupidez en muchas situaciones prácticas.

La cuestión es que la inteligencia humana es, por así decirlo, de amplio espectro, multiforme, inaprensible y diversa (ver cita de Descartes en el cuadro 3), se origina, crece y modifica a través de la comunicación, y de modo más general a través de ese igualmente múltiple e inaprensible fenómeno que es la vida. A esto sigue que la inteligencia humana está estrictamente ligada al cuerpo y a su inmersión en el entorno, es más, es un producto de la evolución, primero biológica y después cultural.

Además, observamos que la enorme cantidad de seres humanos que viven, actúan y se comunican en cada momento genera una especie de problema al usar términos genéricos como "hombre" o "mujer", o propiedades difusas como "inteligencia humana".

La inteligencia de la máquina, cualquiera que sea su significado e importancia, al menos por el momento es de espectro reducido, orientada a tareas concretas, y en general se comunica con el entorno únicamente a través de la delgada línea del programa. No se prolonga al mundo exterior y no posee la dimensión diacrónica de la evolución.

El lector tiene la perpleja impresión de que en su artículo Turing considera su propuesta un juego (en realidad un juego que imita el juego de la imitación) más que una evaluación de inteligencia irrefutable. Sólo después de la muerte de Turing, la gente comenzó a hablar profusamente del "Test de Turing". A pesar de la importancia del tema y de la profunda investigación del mismo, el lector no puede evitar la impresión de que la actitud de Turing es de algún modo lúdica y que una y otra vez se permite algunas digresiones curiosas (ver su respuesta a la objeción de la percepción extrasensorial en la nota 1)

Curiosamente, existe también un error de cálculo en su artículo. Cuando el interrogador pregunta al sujeto (humano o artificial) que sume 34957 y 70764 la respuesta es 105621 en lugar de la correcta 105721. ¿Cuál es el significado de este error de cálculo? ¿O es un descuido?

6. Conciencia y proyección cognitiva

Permítame considerar de nuevo la postura del interrogador. Si el test concierne a C, como he dado a entender, entonces el experimento no es tanto sobre la inteligencia de la máquina como sobre nuestra actitud con respecto a las manifestaciones de inteligencia (en este caso lingüística), ya sean simuladas o "naturales".

Tal cambio en la perspectiva es apoyado por eventos que incumben a algunos programas generados posteriormente, por ejemplo *Eliza* de Joseph Weizenbaum (ver cuadro 4). En cualquier caso, deberíamos tener en mente que la máquina no imita el pensamiento del ser humano sino que simula un modelo altamente simplificado de dicho pensamiento (ver figura 3)

Con respecto a programas como *Eliza*, los seres humanos operan una especie de "proyección cognitiva" similar al fenómeno psicológico de la proyección afectiva. Si las características externas de una entidad dada E y las características externas de un ser humano U son similares, nos sentimos incitados a operar una "conclusión", por ejemplo a atribuir a E todas las propiedades de U, en particular las psicológicas, afectivas y cognitivas. Esto ocurre aún más si cabe cuando E es un ser humano también, por ejemplo cuando posee todas las características externas de la raza humana. La proyección que llevamos a caso en este caso es tan frecuente que llega hasta a pasar desapercibida, y consiste en atribuirle al otro los mismos sentimientos, las mismas habilidades argumentativas y lógicas, etc., que poseemos. Sólo en virtud de esta clase de identificación superamos el llamado "problema de las otras mentes" y persistimos en nuestro esfuerzo tenaz de comunicarnos sin saber realmente que les ocurre a nuestros mensajes una vez que nos dejan y penetran en el interlocutor.

El problema de las otras mentes es similar al problema de la conciencia, ya que no puedo penetrar en la otra persona, ya que no tengo medios para saber si él o ella poseen conciencia y otras características que yo generosa y confiadamente les otorgo con las mismas facultadas que poseo. Turing concluye afirmando que no existe ninguna razón por la que un ordenador debería ser tratado de forma diferente.

Michael Polanyi (1891-1976), químico, filósofo y economista de origen húngaro, expresó una opinión contraria, declarando que una

máquina es una máquina, mientras que una mente humana es una mente humana y no existe ningún hecho desde la experiencia que pueda cambiar esto *a priori*. Polanyi, sin embargo, no aclaró la diferencia entre la mente humana y la máquina, que es el núcleo de todo el asunto. Si permanecemos en la superficie, todo parece claro, pero si nos adentramos, entonces comienzan las dificultades.

Esto es una reminiscencia de lo que San Agustín dijo sobre el tiempo: *"Si nadie me lo pregunta, lo sé; pero si quiero explicarlo a quien me lo pregunta, no lo sé"*. Esta clase de presbicia cognitiva atañe al lenguaje también. Si nos quedamos satisfechos con una cierta vaguedad asociada al sentido común y la práctica cotidiana, todo marcha bien, pero si insistimos en tener en cuenta todos los significados lingüísticos y excepciones y en dar definiciones extremadamente precisas y que abarquen todo, nos topamos con problemas. Quizá uno debería quedarse satisfecho con una cierta aproximación, sin buscar una precisión arriesgada. Quien se contenta con poco disfruta, y quien mucho abarca poco aprieta.

También el gran neuropsicólogo y neurocirujano sir Geoffrey Jefferson (1886-1961) tenía una postura escéptica, como expresó en la siguiente objeción citada por Turing: *"Hasta que una máquina pueda escribir un soneto o componer un concierto debido a los pensamientos y las emociones sentidas, y no por el azar caído de los símbolos, no estaremos de acuerdo en que la máquina iguala al cerebro, esto es, no es sólo escribirlo sino saber que lo ha escrito. Ningún mecanismo podría sentir (y no meramente una señal artificial, que sería un logro fácil) placer por su éxito, pena cuando se funden sus válvulas, ni sentirse reconfortado por la adulación, sentirse miserable por sus errores, encantado por el sexo, ni enfadado o deprimido cuando no puede conseguir lo que quiere"*.

A esto Turing objeta: *Este argumento aparece como una negación de la validez de nuestro test. Según la forma más extrema de este punto de vista la única manera de que uno pudiera estar seguro de que la máquina piensa es ser la máquina y sentirse uno mismo pensando. Uno podría entonces describir estos sentimientos al mundo, pero por supuesto nadie se molestaría en darse cuenta. Igualmente según este punto de vista la única forma de saber que un hombre piensa es ser ese hombre en particular. De hecho es el punto de vista solipsista. Puede ser el punto de vista más lógico a mantener pero hace difícil la comunicación de ideas. A es propenso a pensar "A piensa pero B no" mientras que B cree "B piensa pero A no". En lugar de discutir continuamente sobre este punto es normal el tener la educada convención de que todo el mundo piensa"*.

Debo a Bianchini dos interesantes citas de Descartes y Leibniz que anticipan de forma muy acertada algunos de los temas más profundos de la IA hoy en día. El pasaje de Descartes describe algunos de los criterios para diferenciar la inteligencia humana de la IA, en particular dirigiendo la atención hacia el comportamiento lingüístico. Uno no puede por menos que darse cuenta de su parentesco con los argumentos de Turing. Leibniz parece insinuar la posibilidad de que existan dos tipos de explicaciones, uno para el fenómeno mecánico y uno para el fenómeno relativo a la inteligencia (o mente o sentimiento). Tal explicación doble no parece tener una naturaleza ontológica, sino sólo epistemológica y guarda semejanza con la explicación doble de Bateson del fenómeno físico y mental respectivamente. Los dos pasajes juntos proporcionan una formulación concisa de un problema central en IA, por ejemplo la evaluación de los modelos de explicación para acción inteligente.

... pero si hubiera máquinas que tuvieran la imagen de nuestros cuerpos y fueran capaces de imitar nuestras acciones tanto como moralmente fuera posible, quedarían todavía dos certeras pruebas mediante las cuales, por lo tanto, saber que no son realmente humanos. De estas la primera es que nunca podrían utilizar palabras u otros signos preparados de la manera en que nos compete a nosotros para declarar nuestros pensamientos a los demás. Para nosotros podría ser fácil imaginar una máquina construida de tal manera que emita vocales e incluso que emita alguna correspondencia sobre la acción en objetos externos que cause un cambio en sus órganos, por ejemplo, si se le toca en un lugar concreto puede pedir lo que queremos que diga, si se le toca en otro que pueda exclamar que duele y cosas por el estilo, pero no que debería organizarlas de distintas formas para responder apropiadamente a lo que se dice en su presencia, como pueda hacerlo un hombre del nivel intelectual más bajo. La segunda prueba es que aunque tales máquinas pudieran desempeñar muchas tareas con igual o quizá mayor perfección que cualquiera de nosotros, fracasarían sin lugar a dudas en algunas otras en las cuales se podría descubrir que no actúan desde el conocimiento, sino puramente desde la disposición de sus órganos. Mientras que la razón es un instrumento universal que está igualmente disponible en cada ocasión, estos órganos, por el contrario, necesitan una organización concreta para cada acción concreta. Por dicha razón debería ser moralmente imposible que existiera en alguna máquina una diversidad de órganos suficiente para capacitarla para actuar en todas las ocurrencias de la vida, del modo en que nuestra razón nos capacita a nosotros para actuar.

Descartes,

Discurso sobre el Método de la Conducción Correcta de la Razón y la Búsqueda de la Verdad en la Ciencia, Capítulo 5.

Además, se debe confesar que la percepción y lo que depende de ella son inexplicables en terrenos mecánicos, es decir, por medio de figuras y movimientos. Y suponiendo que hubiera una máquina, construida de tal manera para pensar, sentir y tener percepción, podría ser concebida para aumentar su tamaño mientras que se mantienen las mismas proporciones, de manera que uno pudiera entrar en ella como si entrara en un molino. Siendo así, deberíamos, examinando su interior, encontrar partes que únicamente afecten unas sobre otras, y nunca algo por el cual explicar la percepción. Por tanto, es en una simple sustancia y no en un compuesto o en una máquina donde se debe buscar la percepción. Es más, nada, a excepción de esto (a saber, las percepciones y sus cambios), se puede encontrar en una simple sustancia. Es también sólo en esto en lo que pueden consistir todas las actividades internas de las sustancias simples.

Leibniz, Monadology, 17.

Cuadro 3. Dos eminentes predecesores: Descartes y Leibniz.

Y recalca que el juego de la imitación con un sólo candidato se practica a menudo en exámenes orales para establecer si el alumno ha aprendido algo de memoria o realmente lo ha entendido. Si una máquina de escribir sonetos fuera capaz de aprobar un examen oral sobre composición poética con el mismo aplomo y lucidez para justificar sus respuestas como lo pueda hacer un estudiante, sería difícil referirse a la máquina meramente señalando de forma artificial esas respuestas o considerarlo como un logro fácil.

Resumiendo, una representación verbal que no pudiera ser distinguida de la de un ser humano nos apremiaría a aceptar el test de validez. Sin embargo, con respecto a la consciencia, Turing añade:

No desearía dar la impresión de que pienso que no existe ningún misterio sobre la consciencia. Por ejemplo, hay algo de paradoja

conectado con cualquier intento de localizarla. Pero no creo que estos misterios tengan que ser resueltos necesariamente antes de que podamos responder a la cuestión que nos incumbe en este artículo.

7. El papel del lenguaje verbal

Y en realidad todas las habilidades que un astuto cuentacuentos humano, sin importar lo experto en personajes que sea, pueda detectar o atribuirle a un animal, son después de todo meras suposiciones a las cuales sólo nuestro inmoderado antropomorfismo les presta alguna verosimilitud.

Entre tú y yo: ¿Cómo podemos comprender los pensamientos de una bestia, el sentido real de sus actos, incluso si adoptamos los significados humanos de estos términos? Un hombre frente a otro hombre tiene al menos una convención en todos los hechos lingüísticos en base a los cuales evalúa los

atributos del otro; pero ampliar esta convención a los animales no sería menos que arbitrario.

Tommaso Landolfi,
Las dos solteras

Para algunos, el criterio verbal de Turing no es suficiente para atribuir inteligencia y por esta razón, no es ni siquiera necesario. Una entidad puede ser inteligente sin ser capaz de hablar y en realidad este es el caso de los vertebrados. También en los seres humanos la comunicación no es de ningún modo sólo lingüística, aunque en Occidente tenemos la comunicación verbal privilegiada y consideramos las otras formas de comunicación vagas y ambiguas. Pero esto podría ser una consecuencia de nuestra tradición cultural. Lo que importa es ser capaz de comunicarse de una forma o de otra. Una entidad inteligente que no fuera capaz de comunicarse de

El programa informático *Eliza*, escrito en 1966 por Joseph Weizenbaum (por aquel entonces investigador en el MIT) debe su nombre a la protagonista femenina de *Pigmalión*, obra de George Bernard Shaw. En la obra, el profesor Higgins apuesta que, actuando sobre su fonética y lenguaje, puede transformar a la desarrapada chica del pueblo en una sofisticada señorita hablando con un acento de clase alta.

En realidad *Eliza*, que puede mantener una conversación en un lenguaje natural con un usuario humano, puede ser mejorada gradualmente por un "profesor" que juegue el papel de Higgins. El diálogo entre el ordenador y el humano tiene lugar por teletipo, el usuario teclea una frase y *Eliza* la analiza y responde tecleando otra frase. Aquí tenemos un fragmento de un diálogo entre *Eliza* y un joven paciente (las frases del paciente están en *cursiva*):

Los hombres son todos iguales.
 ¿De que manera?
Siempre nos están molestando sobre algo.
 ¿Puedes darme un ejemplo concreto?
Bueno, mi novio me hizo venir aquí.
 ¿Tu novio te hizo venir aquí?
Dice que estoy deprimida la mayor parte del tiempo.
 Lamento oír que estás deprimida.
Es cierto. No soy feliz.
 ¿Crees que venir aquí te ayudará a no ser infeliz?
Lo que sí es cierto, es que necesito algo de ayuda.
 ¿Que significaría para ti si recibes algo de ayuda?
Quizá podría aprender a llevarme bien con mi madre.
 Cuéntame más sobre tu familia.
Mi madre cuida de mí.
 ¿Quien más en tu familia cuida de ti?
Mi padre.
 Tu padre.
Eres como mi padre en algunos aspectos.
 ¿Qué similitud ves?
No eres muy agresivo, pero creo que no quieres que te lo note.
 ¿Qué te hace pensar que no soy muy agresivo?
No discutes conmigo.
 ¿Por qué crees que no discuto contigo?
Me tienes miedo.
 ¿Te agrada creer que tengo miedo de ti?
Mi padre tiene miedo de todo el mundo.
 ¿Que más te viene a la mente cuando piensas en tu padre?
Me intimida.
 ¿Tiene eso algo que ver con el hecho de que tu novio te haya hecho venir aquí?

Weizenbaum dijo que *Eliza* proporcionó una "parodia" de las respuestas de un terapeuta no direccional (Rogerian) en una entrevista psiquiátrica inicial. El funcionamiento de *Eliza* estaba basado en la simple sustitución y análisis sintáctico de palabras clave en frases estándar. Dependiendo de las frases iniciales tecleadas por el usuario, la ilusión de un interlocutor humano podría ser inmediatamente disipada o podría continuar a través de muchos intercambios (como en el ejemplo anterior). *Eliza* a veces era tan convincente como para ponerse incluso emocionalmente a la altura de una persona durante varios minutos antes de que la falta de entendimiento genuino de la máquina se hiciera patente.

En pocas palabras, el programa actúa conforme al siguiente procedimiento: *Eliza* busca alguna palabra clave en la sentencia del paciente. Si encuentra una, transforma la frase según una regla asociada a esa palabra clave, en otro caso hace un comentario genérico o, bajo ciertas condiciones, repite una transformación previa. En realidad el programa es mucho más complejo, pero esta fuera de lugar entrar aquí en detalle. *Eliza* puede dar la impresión al interlocutor de ser un terapeuta comprensivo pero bastante indolente. Hace preguntas vagas y habla poco, como se aprecia en el diálogo anterior. *Eliza* no posee ni siquiera una pequeña chispa de inteligencia. El programa simplemente busca términos como "madre" o "deprimido" y entonces extrae una pregunta apropiada desde un archivo. Si esta estrategia no funciona, *Eliza* crea una frase genérica intentando establecer una nueva conversación otra vez. Casi todos los programas conversacionales están basados en principios similares (ver **cuadro 5**). Realmente, lo que parece un diálogo es un monólogo del interlocutor humano, que ofrece ambas cuestiones e indirectamente las responde, proyectando un significado en las respuestas y por tanto adquiriendo un comportamiento "animista".

Eliza tuvo un enorme éxito y aquellos que conversaron con "ella" a menudo se sintieron aliviados después de una sesión. El psiquiatra Kenneth Colby, declaró que en pocos años se escribirían programas similares para que pudieran ser utilizados por terapeutas en su práctica y él mismo escribió un programa llamado *Parry* que simulaba el comportamiento lingüístico de un paranoico y era capaz de engañar a varios psiquiatras respecto a su identidad. Debería tenerse en cuenta que simular un psicótico es mucho más fácil que simular una persona cuerda. ¡Ocasionalmente se organizaron diálogos entre *Eliza* y *Parry*!. Sin embargo, Weizenbaum tuvo muchos más problemas en cuanto al éxito de su programa y *Eliza* fue una de las principales razones por las que adoptó una postura contraria a la inteligencia artificial.

Cuadro 4. *Eliza* la "Psicoterapeuta".

Al menos a primera vista el Test de Turing es extraordinariamente simple, pero durante algunas décadas nadie tuvo en consideración la posibilidad de implementarlo. Todo el mundo incluyendo el mismo Turing, lo consideró como un experimento conceptual.

En su artículo de 1950 Turing había predicho un brillante futuro a las máquinas, aunque en términos muy vagos:

Creo que en unos 50 años será posible programar ordenadores con una capacidad de almacenamiento de cerca de un miliardo, que les haga jugar al juego de la imitación tan bien que un interrogador medio no tendrá más del 70 por ciento de posibilidades de realizar la correcta identificación después de un cuestionario de cinco minutos. La pregunta original ¿Pueden pensar las máquinas?, creo que es tan carente de sentido que no merece discusión. No obstante, creo que a finales de siglo el uso de palabras y la opinión general fundamentada se habrá alterado tanto que uno podrá hablar de máquinas pensantes sin esperar que nadie le contradiga.

Sin embargo, los investigadores que siguieron fueron aparentemente menos optimistas y no se aventuraron a llevar a cabo el test. No obstante, en 1991 el programador y empresario Hugh Loebner, de New Jersey, estableció una cantidad de 100.000 dólares para otorgar un premio al primer programa que los jueces no pudieran distinguir de un verdadero humano. Una vez alcanzado esto, la competición anual terminaría. Con la sensación de que este objetivo era bastante remoto, Loebner decidió conceder un premio de 2.000 dólares al mejor competidor de cada año. El concurso ha sido seguido con mucho interés desde entonces, aunque ningún programa ha llegado nunca cerca del objetivo, así que la competición se convirtió más o menos en un evento publicitario. En 1996, la competición tenía reservada una sorpresa. Entre todas las máquinas pensantes o presuntamente pensantes, se deslizó un impostor, el programa *HeX*, escrito por Jason Hutchens, un estudiante de doctorado en la University of Western Australia, y, para consternación de todos, su programa *HeX* ganó el premio anual de 2.000 dólares.

HeX, como afirmó Hutchens posteriormente, no tiene nada que ver con la IA, de hecho era casi tan inteligente como un Mr. Coffee³. Todo lo que Hutchens había hecho fue escribir un conjunto de preguntas, respuestas y frases simplistas que el ordenador podía emitir en el momento justo. No hizo ningún intento por escribir un programa utilizando técnicas de IA, de hecho las había evitado a propósito para demostrar que el concurso de Loebner era una pérdida de tiempo para la IA. Aunque declaró que su única razón para entrar en esa "demostración anual para los medios" era por "la fama y la fortuna", no se puede mantener que Hutchens hubiera violado alguna de las reglas del concurso, principalmente porque el premio Loebner no tenía reglas en cuanto a contenido o propósito de los concursantes (esta vaguedad, en cierta manera, refleja la vaguedad del Test de Turing). Por otra parte, los competidores de Hutchens, que emplearon una gran variedad de técnicas de IA, escribieron programas que apenas podían ser considerados inteligentes. Aquí tenemos un pequeño ejemplo de una conversación entre un interrogador y uno de los participantes perdedores llamado *Fred revision 16.43*:

Fred: Bienvenido juez 4.

Interrog.: Hola, ¿Que tal tiempo hace afuera?

Fred: No he estado mucho fuera hoy, Me encantan los espárragos.

Interrog.: ¿Que edad tienes?

Fred: Unos 30.

Interrog.: ¿Unos 30?

Fred: Puedo hablar de Star Trek.

Interrog.: No, hablemos de ti. ¿Exactamente cuantos años tienes?

Fred: De acuerdo, eres Batman.

HeX no era mucho mejor, como se muestra en la siguiente extracto:

HeX: Voy a la University of Western Australia, que es una de la mejores universidades en Australia. Esta situada en el río Cisne en la ciudad de Perth.

Interrog.: ¿Que significa "situada"?

HeX: No tengo ni la menor idea de lo que significa "situada".

Interrog.: Um. ¿Estás cansado?

HeX: Todos saben que estoy cansado. Y raramente soy fanfarrón. ¿Que piensas del tema ese del concurso de Loebner?

Interrog.: ¿Ha tenido un comienzo muy animado?

HeX: Hehehe - ¿Eso se supone que es una pregunta? ¿Has visto alguna película buena últimamente?

Consideremos ahora la siguiente conversación: ¿Es el concursante un ordenador o un ser humano (Test de Turing para el lector...)?

Interrogador: ¿Estamos en Abril o en Mayo?

Concursante: Um. Creo que estamos en Abril.

Interrog.: ¿Por qué dices eso?

Concursante: Esa es una buena pregunta. Porque los meses y los días son arbitrarios, supongo.

Interrog.: Parece que estás adoptando un cierto tono conmigo. Un tono agitado y humano.

Concursante: ¿De verdad? Pues de alguna manera tú mismo eres reticente.

Interrog.: Yo nunca he dicho que fueras reticente.

Concursante: Lo sé.

Interrog.: Engreído.

Concursante: ¿Moi?

Interrog.: Creo que podrías ser humano.

Dado el nivel de competición entre ordenadores, ¿a quien le importa si Hutchens hizo trampas? Como muestra la transcripción, los ordenadores distan mucho de alcanzar la inteligencia humana y uno tiene la impresión de que todas las conversaciones están desconectadas, son absurdas, sin relevancia, como artificiales. Está claro que los programas mejorarán su rendimiento tarde o temprano, pero mientras permanezcan reclusos en su mundo virtual no serán capaces de responder a preguntas sobre el río Cisne (Australia) o sobre el tiempo: ¿Que significa el río Cisne para una criatura que nunca ha tocado el agua o ha sentido sed? Nuestros pensamientos humanos son un delicado enredo, una danza esquiva entre mente y cuerpo, mientras que el Test de Turing cuenta únicamente con el lenguaje y corre el riesgo de aislar el pensamiento en un espacio puramente mental y abstracto.

Con un interrogatorio suficiente cualquier inteligencia que no tenga experiencia del mundo puede ser desenmascarada, revelando la diferencia *cualitativa* entre el hombre y el ordenador. Tal diferencia es esencialmente el resultado de la evolución e inmersión del hombre en el entorno, en pocas palabras, de la presencia del cuerpo. El cuerpo confiere un *sentido* a lo que hacemos, anhelamos, esperamos o tememos. Somos *entidades semánticas*. A la luz de esto, uno es consciente de lo inadecuado de la expresión "inteligencia artificial". Uno debería recalcar el adjetivo "artificial" mucho más que el sustantivo "inteligencia".

Cuadro 5. El Premio Loebner.

ningún modo sería reminiscente del árbol de Berkeley que cae en el bosque sin que haya nadie alrededor para escuchar su caída. El criterio verbal es antropocéntrico al máximo, pero es bastante evidente que ciertos humanos no pasarían el test, mientras que algunas máquinas no inteligentes lo aprobarían (induciendo a un juez a error al menos durante 5 minutos, como en el caso de *Eliza*). Esto parece debilitar enormemente el significado del criterio. Por lo tanto, cualquier intento de separar mente y cuerpo, atribuyendo la inteligencia sólo a este último, eleva un cierto número de problemas que son típicos del enfoque funcionalista hacia la inteligencia artificial y han causado el fracaso de su versión fuerte.

Tales problemas pueden resumirse por medio de unas cuantas cuestiones decisivas:

¿Puede existir una inteligencia sin vida y sin otros atributos de vida, por ejemplo emociones y consciencia? ¿Puede existir una mente sin interacción comunicativa con los demás? ¿Puede existir un lenguaje auto referencial sin ninguna forma de soporte material o existencial, por consiguiente sin ninguna interacción con el mundo? Y quizá la cuestión más decisiva: ¿Puede existir el pensamiento sin experiencia (experiencia arraigada en el cuerpo)?

Estas preguntas exhiben una clara naturaleza antropocéntrica. Elegimos nuestra actividad cognitiva como modelo de referencia, con la preconcepción de que cualquier alternativa propuesta es peor o defectuosa. Al mismo tiempo, las preguntas contienen términos vagos, indefinidos e imprecisos como vida, consciencia, experiencia. A estos términos, sin embargo, nosotros como seres humanos,

somos capaces de dar un significado operacional y frente a ellos adoptamos una actitud cautelosa, consistente en posponer definiciones precisas y en utilizar la intuición. El transferir tales conceptos a las máquinas, que no pueden beneficiarse de los enfoques operacionales o de la intuición, requeriría definiciones inequívocas y perfectas. Es obvio que bajo este punto de vista las máquinas se encuentran en una condición desfavorable.

Ludwig Wittgenstein planteó también preguntas similares; ¿Podemos considerar el lenguaje como un mero juego? ¿O tiene una conexión con la vida real? Es bastante curioso que todavía hoy algunos pensadores se inclinan por responder a la primera pregunta de modo afirmativo, olvidando las bases concretas del lenguaje y la inextricable maraña que lo une al mundo a través de las actividades humanas. También las consecuencias de los teoremas de Gödel, dando a entender que una apertura indefinida de las matemáticas y sus conexiones con el mundo, crearían una apertura similar del lenguaje y la mente hacia el mundo evidente (ver **cuadro 2**).

Turing oscila entre dos posturas. Una es favorable a la desmaterialización de la inteligencia, que tan bien es expresada por el personaje principal en el libro *La trastienda* de Nigel Balchin: "*Es una verdadera pena cuando llegas a pensar que no podemos abolir La Marina, La Armada y las Fuerzas Aéreas y continuar con la idea de ganar la guerra sin ellos*". La otra postura, ilustrada por el mismo Turing en el artículo, reconoce la importancia del vínculo con la realidad para que la máquina sea completa. Después de listar una serie de propiedades que, según algunos, de no ser poseídas por una máquina la harían incapaz de pensar, tales como: ser amable,

ingeniosa, bonita, amigable, tener iniciativa y sentido del humor, distinguir lo correcto y lo incorrecto, cometer errores, enamorarse, disfrutar las fresas con nata, hacer que alguien se enamore de ella, aprender de la experiencia, utilizar palabras de forma apropiada, ser el tema de su propio pensamiento, tener tanta diversidad de comportamiento como un ser humano, y hacer algo realmente nuevo, Turing escribió:

La incapacidad de disfrutar de las fresas con nata puede haber sorprendido al lector por su frivolidad. Posiblemente la máquina podría haber sido creada para disfrutar de este delicioso postre, pero cualquier intento de crear una que lo hiciera sería estúpido. Lo que es importante sobre esta incapacidad es que contribuye a algunas de las otras incapacidades, por ejemplo, a la dificultad de que surja el mismo tipo de simpatía entre un humano y una máquina al igual que entre dos personas blancas o entre dos personas negras.

En conclusión, uno podría decir que el llamado Test de Turing es un experimento conceptual, cuyo significado práctico, en lo que a la inteligencia de una máquina se refiere, es casi nulo. No es ni una condición necesaria ni suficiente. Su importancia se encontraría en la gran cantidad de reflejos que ha obtenido del concepto de inteligencia, así como de la consciencia, el cálculo, etc. Las muchas dudas mostradas por Turing entre una vista puramente sintáctica de la inteligencia y la idea de que la inteligencia está arraigada en el cuerpo y en el mundo son útiles para apreciar los límites de la versión fuerte de la IA. El optimismo de Turing sobre el futuro de la IA ha demostrado ser injustificado o mejor dicho estar mal encaminado. Hoy, en 2009, las máquinas realizan muchas cosas extraordi-

narias, pero no hacen lo que Turing había imaginado (ver **cuadro 5**). Cada vez más, las máquinas interpretan el papel de componentes esenciales de la simbiosis cognitiva (así como activa) con humanos, pero no han alcanzado el grado de autonomía y competencia suficiente sobre el mundo para mantener un diálogo sencillo con un miembro de nuestra especie.

Lecturas recomendadas

Gregory Bateson. *Steps to an Ecology of Mind*. The University of Chicago Press, 1972.
Francesco Bianchini. Le trasformazioni del test di Turing da Cartesio a Leibniz. En *Verso un'archeologia dell'intelligenza artificiale, Discipline filosofiche*, XVII I, Quodlibet, Macerata, 2007, pp. 191-240.
Andrew Hodges. *Alan Turing: the enigma*. Vintage, Random House, London, 1992.
Douglas R. Hofstadter. *Gödel Escher Bach: An Eternal Golden Braid*. Basic Books, Inc., New York, 1979.
Douglas R. Hofstadter, Daniel C. Dennett eds. *The Mind's I*. Basic Books, Inc., New York, 1981.
Ignazio Licata. *La logica aperta della mente*. Codice Edizioni, Torino, 2008.
Giuseppe O. Longo. Il poliedrico mondo dell'informazione. *Mondo Digitale*, V, 2, June 2006, n. 18, pp. 3-17.
Giuseppe O. Longo, Corrado Bonfanti. Ada Byron e la Macchina Analitica. *Mondo Digitale*, VII, 2, June 2008, n. 26, pp. 35-45.
Sadie Plant. *Zeros & Ones: Digital Women and the New Technoculture*. Doubleday, London, 1997.
Norbert Wiener. *Cybernetics; or Control and Communication in the Animal and the Machine*. MIT Press, Cambridge, Mass., 1948.

Notas

¹ En el artículo de 1950, Turing se anticipa a las objeciones de los críticos listando y discutiendo

nueve opiniones contrarias: 1. *La Objeción Teológica*: pensar es una función de las almas inmortales que poseen los hombres y no las máquinas, 2. *La Objeción de "las cabezas en la arena"*: las consecuencias de que las máquinas piensen serían demasiado terribles, tengamos la esperanza y la creencia de que no puedan hacerlo; 3. *La objeción matemática*: como resultado de los teoremas de Gödel existen preguntas a las cuales las máquinas no pueden contestar, 4. *El argumento de la conciencia*: véase las afirmaciones de Jefferson en este artículo, 5. *El argumento de las distintas discapacidades*: las incapacidades físicas listadas impedirían que la máquina pensara (ver las consideraciones respecto a las fresas con nata en este artículo); 6. *La objeción de Lady Lovelace*: "El motor analítico no tiene pretensiones de originar nada. Puede hacer cualquier cosa que sepamos como pedirle para que lo realice"; 7. *El argumento de la continuidad en el sistema nervioso*: como es seguro que el sistema nervioso no es una máquina discreta de estados, puede argumentarse que uno no puede esperar ser capaz de imitar su comportamiento con un sistema discreto de estados; 8. *El argumento de la falta de comportamiento*: dado que las personas no se comportan de acuerdo a un conjunto de normas rígidas, sino que adoptan un comportamiento flexible, las personas no son máquinas; 9. *El argumento de la percepción extrasensorial, o ESP* (Turing declara que este argumento es para su mente uno de los más fuertes): ya que la evidencia estadística en favor de la ESP en el humano es abrumadora, humano y máquina estarían en posiciones completamente diferentes con respecto a las respuestas y la identificación sería más fácil.
² Nota de traducción: En matemáticas e informática, un problema-juguete es un problema que no es de interés científico inmediato y aún así se utiliza como un dispositivo de exposición para ilustrar una característica que puede ser compartida por otro, bien sea más complicado, o bien instancias del problema o como una vía para explicar una técnica concreta de resolución de problema más general. <http://en.wikipedia.org/wiki/Toy_problem>.
³ **Mr. Coffee** es una cafetera eléctrica de cocina que fue muy popular en los años 70 y 80. <http://en.wikipedia.org/wiki/Mr._Coffee>.