

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).

< <http://www.ati.es/novatica/>
< <http://www.ati.es/reicis/>

ATI es miembro fundador de **CEPIS** (Council of European Professional Informatics Societies), representa a España en **IFIP** (International Federation for Information Processing) y es miembro de **CLEI** (Centro Latinoamericano de Estudios de Informática) y de **CECUA** (Confederación of European Computer User Associations). Asimismo, tiene un acuerdo de colaboración con **ACM** (Association for Computing Machinery) y colabora con diversas asociaciones informáticas españolas.

Consejo Editorial

Guillem Alsina González, Rafael Fernández Calvo (presidente del Consejo), Jaime Fernández Martínez, Luis Fernández Sanz, José Antonio Gutiérrez de Mesa, Silvia Leal Martín, Didac López Viñas, Francesc Noguera Puig, Joan Antoni Pastor Collado, Viktu Pons i Colomer, Moisés Robles Gener, Cristina Vigli Díaz, Juan Carlos Vigo López

Coordinación Editorial

Llorenç Pagès Casas <pages@ati.es>

Composición y autedición

Impresión Offset Derra S. L.

Traducciones

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

Administración

Tomas Brunete, María José Fernández, Enric Camarero

Secciones Técnicas - Coordinadores

Acceso y recuperación de la Información

José María Gómez Hidalgo (Pragsis Technologies), <jmgomez@pragsis.com>

Manuel J. Maña López (Universidad de Huelva), <manuel.mana@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>

José Filich (Universidad Politécnica de Valencia), <jfilich@disca.upv.es>

Auditoría SITIC

Marina Touriño Troitino, <marinatourino@marinatourino.com>

Sergio Gómez-Landero Pérez (Endesa), <sergio.gomezlandero@endesa.es>

Derecho y tecnologías

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

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

Enseñanza Universitaria de la Informática

Cristóbal Pareja Flores (DSIF-UCM), <cpajera@sip.uom.es>

J. Ángel Velázquez Ilurbe (DLSI I, URJ), <angel.velazquez@urj.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 (Virati), <encarna.quesada@virati.com>

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

Gestión del Conocimiento

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

Gobierno corporativo de las TI

Manuel Palao García-Suñel (ATI), <manuel@opalao.com>

Miguel García-Menéndez (ITI), <mgarciamenendez@itirendsinstitute.org>

Informática y Filosofía

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

Roberto Feltoro Orta (UNED), <rfeltoro@gmail.com>

Informática Gráfica

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

Roberto Vivó Hernando (Eurographics, sección española), <rvivo@dsic.upv.es>

Ingeniería del Software

Luis Fernández Sanz, Daniel Rodríguez García (Universidad de Alcalá), <luis.fernandez.daniel.rodriguez@uah.es>

Inteligencia Artificial

Vicente Boti Navarro, Vicente Julián Inglada (DSIC-UPV), <{vbotti,vjnglada}@dsic.upv.es>

Interacción Persona-Ordenador

Pedro M. Latorre Andrés (Universidad de Zaragoza, AIFD), <platorre@unizar.es>

Francisco L. Gutiérrez Vela (Universidad de Granada, AIFD), <fgutierrez@ugr.es>

Lengua e Informática

M. del Carmen Ugarte García (ATI), <cuarte@ati.es>

Lenguajes Informáticos

Oscar Belmonte Fernández (Univ. Jaime I de Castellón), <obelfern@lsi.uji.es>

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

Lingüística computacional

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

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

Modelado de software

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

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

Mundo estudiantil y jóvenes profesionales

Federico G. Mon Troiti (RITS), <gml.fede@gmail.com>

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

Profesión Informática

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

Miguel Sarrías Grito (ATI), <mitsa@arris.net>

Redes y servicios telemáticos

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

Ana Pont Sanjuan (UPV), <apont@disca.upv.es>

Robotica

José Cortés Arenas (Sopra Group), <joscortea@gmail.com>

Juan González Gómez (Universidad Carlos III), <juan@iearobotics.com>

Seguridad

Javier Areitio Bertolin (Univ. de Deusto), <jareitio@deusto.es>

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

Sistemas de Tiempo Real

Alejandro Alonso Muñoz, Juan Antonio de la Puente Alfaro (DIT-UPM), <{aalonso,jpuente}@dit.upm.es>

Software Libre

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

Israel Herráiz Tabernero (Universidad Politécnica de Madrid), <isra@herraiz.org>

Tecnologías para la Educación

Juan Manuel Dodero Beardo (UC3M), <dodero@inf.uc3m.es>

César Pablo Córcoles Briongo (UOC), <ccorcoles@uoc.edu>

Tecnologías y Empresa

Didac López Viñas (Universidad de Girona), <didac.lopez@ati.es>

Alonso Álvarez García (TID), <aag@tid.es>

Tendencias tecnológicas

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

Juan Carlos Vigo (ATI), <juancarlosvigo@atinet.es>

TIC y Turismo

Andrés Aguiayo Maldonado, Antonio Guevara Plaza (Univ. de Málaga), <{aguayo,guevara}@lccuma.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

Plaza de España 6, 2ª planta, 28008 Madrid

Tlf: 91 4029391; fax: 91 3093685 <novatica@ati.es>

Administración y Redacción ATI Cataluña

Calle Àvila 50, 3a planta, local 9, 08005 Barcelona

Tlf: 93 4125235; fax: 93 4127113 <secretgen@ati.es>

Redacción ATI Andalucía

<secretand@ati.es>

Redacción ATI Galicia

<secretgal@ati.es>

Suscripción y Ventas

<novatica.subscripciones@atinet.es>

Publicidad

Plaza de España 6, 2ª planta, 28008 Madrid

Tlf: 91 4029391; fax: 91 3093685 <novatica@ati.es>

Imprenta:

Impresión Offset Derra S.L., Lluís 41, 08005 Barcelona

Depósito legal: B 15.154-1975 -- ISSN: 0211-2124; CODEN NOVAC

Portada: "El futuro en tus manos" - Concha Arias Pérez / © ATI

Diseño: Fernando Agresta / © ATI 2003

Nº 234, octubre-diciembre 2015, año XLI

editorial

40 años de Novática > 03

40º aniversario de Novática

Novática: un breve apunte histórico > 05

Rafael Fernández Calvo

Presentación. 2015-2025: En la encrucijada de los nuevos tiempos > 08

Llorenç Pagès Casas, Manolo Palao

Actividades internacionales de ATI

Grupo "Green ICT" de CEPIS > 13

Francisco Javier Esteve Zarazaga

CEPIS y su programa "Women in ICT" > 14

Maribel Sánchez-Segura

Legal and Security Issues Special Interest Network (LSI SIN) de CEPIS > 15

María Teresa Villalba de Benito

ATI en el grupo "ICT Professionalism" de CEPIS > 16

Ricardo Colomo-Palacios

Participación de ATI en la red de especial interés "Computing at Schools" > 17

Jesús Moreno León

Visión general de la dimensión internacional de ATI

y de su participación en proyectos europeos > 18

Luis Fernández Sanz

Año 2025: El futuro de la Informática

Política

Equidad Digital en países desarrollados y en desarrollo > 20

Ramon Puigjaner Trepal

Gobierno electrónico: Paradigma de relación entre administración y ciudadanía > 27

Sebastià Justicia Pérez, Francisco López Crespo

Sociedad

El papel de las TIC en la productividad laboral y de los recursos.

¿Estamos utilizando la innovación tecnológica de forma equivocada? > 32

Lorenz M. Hilty

Innovación responsable en TIC > 36

Norberto Patrignani

Un 2025 sin un mejor gobierno corporativo de las tecnologías de la información > 40

Manolo Palao

La "economía colaborativa" desde un punto de vista legal:

los casos de Uber, Blablacar y Airbnb > 45

Alexander Benlal

¿Y el futuro de la educación? > 48

Juana M. Sancho Gil

El futuro de la educación: Como la tecnología puede impactar en la educación > 53

Francesc Santanach Delisau, Josep Prieto Blázquez

TIC, Smart Cities, Big Data e Internet de las Cosas > 58

José Javier Medina Muñoz

Apuesta por un e-liderazgo consciente para la nueva ola digital > 62

Silvia Leal Martín, Jorge Urrea Filgueira

Tecnología

Reflexiones sobre el futuro de la Interacción Persona-Ordenador > 65

Gerrit van der Veer, Julio Abascal

Horizonte 2025: Fricciones y sinergias entre seguridad y privacidad > 70

Alberto Cita, Sara Degli-Esposti, José María del Álamo

La tecnología tras las divisas virtuales: implicaciones a futuro > 74

César Pérez-Chirinos

Las "TIC verdes" en el Horizonte 2025 > 80

Francisco Javier Esteve Zarazaga

Una visión del desarrollo de software para 2027 > 85

Jesús J. García Molina

Los siguientes 10 años de Ingeniería de Software: Retos y oportunidades > 90

Antonio Vallecillo Moreno

El software libre en 2025 > 95

Jesús M. González Barahona

Algunas reflexiones sobre el presente y futuro de la Inteligencia Artificial > 97

Ramon López de Mántaras Badía

!Vaya tela! Atrapados por la Web > 102

José Antonio Gil, Ana Pont

Las TIC y el futuro de la accesibilidad para personas con discapacidad > 105

Loïc Martínez Normand

Ficción

Nuevo Getsemaní > 108

Miquel Barceló García

asuntos interiores

Coordinación editorial / Programación de Novática / Socios Institucionales > 111

Ramon López de Mántaras Badia

Institut d'Investigació en Intel·ligència Artificial (IIIA), Consejo Superior de Investigaciones Científicas (CSIC)

<mantaras@iia.csic.es>

Algunas reflexiones sobre el presente y futuro de la Inteligencia Artificial

1. Introducción

El objetivo de la Inteligencia Artificial (IA), lograr que una máquina tenga una inteligencia de tipo general similar a la humana, es uno de los objetivos más ambiciosos que se ha planteado la ciencia. Por su dificultad, es comparable a otros grandes objetivos científicos como explicar el origen de la vida, el origen del universo o conocer la estructura de la materia.

A lo largo de los últimos siglos, este afán por construir máquinas inteligentes nos ha conducido a inventar modelos o metáforas del cerebro humano. Por ejemplo, en el siglo XVII, Descartes se preguntó si un complejo sistema mecánico compuesto de engranajes, poleas y tubos podría, en principio, emular el pensamiento. Dos siglos después, la metáfora fueron los sistemas telefónicos ya que parecía sus conexiones se podían asimilar a una red neuronal.

Actualmente, el modelo dominante en IA es el modelo computacional basado en el ordenador digital. Este artículo contiene algunas reflexiones sobre la IA actual así como sobre su futuro.

2. La hipótesis del Sistema de Símbolos Físicos y la IA: IA débil versus IA fuerte

Centrándonos en la IA actual, es importante distinguir entre IA débil e IA fuerte, dos visiones que se corresponden, respectivamente, con los dos siguientes intentos de definición:

Definición 1 (IA débil)- La IA es la ciencia e ingeniería que permiten diseñar y programar ordenadores de forma que realicen tareas que requieren inteligencia.

Definición 2 (IA fuerte)- La IA es la ciencia e ingeniería que permitirá replicar la inteligencia humana mediante máquinas

En una ponencia, con motivo de la recepción del prestigioso "Premio Turing" en 1975, Allen Newell y Herbert Simon formularon la hipótesis del "Sistema de Símbolos Físicos" según la cual *todo sistema de símbolos físicos posee los medios necesarios y suficientes para llevar a cabo acciones inteligentes.*

Por otra parte dado que los seres humanos somos capaces de mostrar conductas inte-

Resumen: En este artículo se presentan algunas reflexiones sobre el pasado, presente y futuro de la Inteligencia Artificial (IA). Partiendo de la hipótesis del sistema de símbolos físicos se distingue entre la IA fuerte y la IA débil y se analizan los principales modelos existentes: simbólico, conexionista, evolutivo y corpóreo. A continuación se reflexiona sobre el problema más importante de la IA: la adquisición de conocimientos de sentido común el cual constituye el requerimiento fundamental para que las máquinas actuales dejen de tener inteligencias artificiales especializadas y empiecen a tener inteligencias artificiales de tipo general. Finalmente se incluyen algunas reflexiones sobre el futuro de la IA.

Palabras clave: Conocimientos de sentido común, Inteligencia Artificial, IA débil, IA fuerte, sistemas integrados.

Autor

Ramon López de Mántaras Badia es Profesor de Investigación del Consejo Superior de Investigaciones Científicas (CSIC) y Director del Instituto de Investigación en Inteligencia Artificial del CSIC. Doctor en Física por la Universidad de Toulouse III (Francia), *Master of Science* en Informática por la Universidad de California-Berkeley (EEUU) y Doctor en Informática por la Universidad Politécnica de Barcelona. Fue Profesor Titular de la Facultad de Informática de Barcelona y Catedrático de la Universidad Pierre et Marie Curie de París. Es uno de los pioneros de la inteligencia artificial en España y miembro del comité de redacción de varias revistas internacionales. Ha recibido, entre otros, el "Premio Ciudad de Barcelona" de investigación, el "European Artificial Intelligence Research Award", el "2011 American Association of Artificial Intelligence (AAAI) Robert S. Engelmore Award" y el "Premio Nacional de Informática" de la Sociedad Científica Informática de España en 2012. También es miembro de honor ("Fellow") de la Asociación Europea de Inteligencia Artificial) desde 2000. Fue presidente de ECAI'04 (2004 European Conference on Artificial Intelligence) y Presidente de IJCAI'07 (2007 International Joint Conference on Artificial Intelligence). Es autor, o co-autor, de unas 260 publicaciones científicas. Actualmente investiga en razonamiento y aprendizaje basado en casos, en robots autónomos capaces de aprender interactuando con su entorno físico, y en inteligencia artificial aplicada a la música. <<http://www.iia.csic.es/~mantaras>>.

ligentes en el sentido general, entonces, de acuerdo con la hipótesis, nosotros somos también sistemas de símbolos físicos.

Conviene aclarar a qué se refieren Newell y Simon cuando hablan de "sistema de símbolos físicos" (SSF). Un SSF consiste en un conjunto de entidades denominadas símbolos que, mediante relaciones, pueden ser combinados formando estructuras más grandes (como los átomos que se combinan formando moléculas) y que pueden ser transformados aplicando un conjunto de procesos.

Estos procesos pueden crear nuevos símbolos, crear y modificar relaciones entre símbolos, almacenar símbolos, comparar si dos símbolos son iguales o distintos, etc. Estos símbolos son "físicos" en tanto que tienen un sustrato físico-electrónico (en el caso de los ordenadores) o físico-químico-biológico (en el caso de los seres humanos).

Efectivamente, en el caso de los ordenadores los símbolos se realizan mediante circuitos electrónicos digitales y en el caso de los seres humanos mediante redes de neuronas.

En definitiva, de acuerdo con la hipótesis SSF, la naturaleza del sustrato (circuitos electrónicos o redes neuronales) carece de importancia siempre y cuando dicho sustrato permita procesar símbolos. No olvidemos que se trata de una hipótesis y por lo tanto no debe de ser ni aceptada ni rechazada a priori. En cualquier caso su validez o refutación se deberá verificar, de acuerdo con el método científico, en ensayos experimentales.

La IA es precisamente el campo científico dedicado a intentar verificar esta hipótesis en el contexto de los ordenadores digitales, es decir verificar si un ordenador convenientemente programado es capaz o no de tener conducta inteligente de tipo general.

“ La IA fuerte implicaría que un ordenador convenientemente programado no simula una mente sino que es una mente y por consiguiente debería ser capaz de pensar igual que un ser humano ”

Es importante el matiz de que debería tratarse de inteligencia de tipo general y no especializada ya que la inteligencia de los seres humanos es de tipo general.

Exhibir inteligencia específica es otra cosa bien distinta. Por ejemplo, los programas que juegan al ajedrez a nivel de Gran Maestro son incapaces de jugar a las damas a pesar de ser un juego mucho más sencillo, se requiere diseñar y ejecutar un programa distinto e independiente del que le permite jugar al ajedrez para que el mismo ordenador juegue también a las damas.

En el caso de los seres humanos no es así ya que cualquier jugador de ajedrez puede aprovechar sus conocimientos sobre este juego para, en cuestión de segundos, jugar a las damas perfectamente.

El diseño y realización de inteligencias artificiales que únicamente muestran comportamiento inteligente en un ámbito muy especializado, constituye lo que se conoce por “IA débil” en contraposición con la “IA fuerte” a la que se referían Newell y Simon y otros padres fundadores de la IA [1].

Aunque estrictamente la hipótesis SSF se formuló en 1975, ya estaba implícita en las ideas de los pioneros de la IA en los años 50 e incluso en las ideas de Alan Turing en sus escritos [2][3] sobre máquinas inteligentes.

De hecho, quien introdujo esta distinción entre IA débil y IA fuerte fue el filósofo John Searle en un artículo crítico con la IA publicado en 1980 [4] que provocó, y sigue provocando, mucha polémica.

La IA fuerte implicaría que un ordenador convenientemente programado no simula una mente sino que es una mente y por consiguiente debería ser capaz de pensar igual que un ser humano. Searle en su artículo intenta demostrar que la IA fuerte es imposible.

La IA débil, por otro lado, consistiría según Searle en construir programas que ayudan al ser humano en sus actividades mentales en lugar de duplicarlas.

La capacidad de los ordenadores para realizar tareas específicas mejor que las personas

ya se ha demostrado. En ciertos dominios los avances de la IA débil superan en mucho la pericia humana, como por ejemplo en buscar soluciones a formulas lógicas con muchas variables.

También se asocia con la IA débil el hecho de formular y probar hipótesis acerca de aspectos relacionados con la mente (por ejemplo la capacidad de razonar deductivamente, de aprender inductivamente, etc.) mediante la construcción de programas que llevan a cabo dichas funciones. Todos los avances logrados hasta ahora en el campo de la IA son manifestaciones de la IA débil.

3. Los principales modelos en Inteligencia Artificial: Simbólico, conexionista, evolutivo y corpóreo

El modelo dominante en IA ha sido el *simbólico*, que tiene sus raíces en la hipótesis SSF. De hecho sigue siendo muy importante y actualmente se considera el modelo clásico en IA (también denominado por el acrónimo GOFAI de *Good Old Fashioned AI*).

Es un modelo *top-down* que se basa en el razonamiento lógico y la búsqueda heurística como pilares para la resolución de problemas, sin que el sistema inteligente necesite formar parte de un cuerpo ni estar situado en un entorno real.

Es decir, la IA simbólica opera con representaciones abstractas del mundo real que se modelizan mediante lenguajes de representación basados principalmente en la lógica matemática y sus extensiones. Por este motivo, los primeros sistemas inteligentes resolvían principalmente problemas que no requerían interactuar directamente con el entorno como por ejemplo demostrar teoremas o jugar al ajedrez (los sistemas de IA que juegan al ajedrez no necesitan la percepción visual para ver las piezas en el tablero ni manos artificiales para mover las piezas).

Ello no significa que la IA simbólica no pueda ser usada para, por ejemplo, programar el módulo de razonamiento de un robot físico situado en un entorno real, pero en los primeros años los pioneros de la IA no disponían de lenguajes de representación del conocimiento ni de programación que permitieran hacerlo de forma eficiente y por este

motivo los primeros sistemas inteligentes se limitaron a resolver problemas que no requerían interacción directa con el mundo real.

Actualmente, la IA simbólica se sigue usando para demostrar teoremas o jugar al ajedrez, pero, gracias a los progresos en representación de conocimientos y visión artificial, también se puede usar para aplicaciones que requieren percibir el entorno y actuar sobre él como por ejemplo el aprendizaje y la toma de decisiones en robots autónomos.

Simultáneamente, con la IA simbólica también empezó a desarrollarse una IA bio-inspirada llamada *conexionista*. Los sistemas conexionistas no son incompatibles con la hipótesis SSF pero, contrariamente a la IA simbólica, se trata de una modelización *bottom-up*, ya que también se basan en la hipótesis de que la inteligencia emerge a partir de la actividad distribuida de un gran número de unidades interconectadas que procesan información paralelamente.

En la IA conexionista estas unidades son modelos aproximados de la actividad eléctrica de las neuronas biológicas. Ya en 1943 se propuso un modelo simplificado de neurona en base a la idea de que una neurona es esencialmente una unidad lógica [5].

Este modelo es una abstracción matemática con entradas (“*dendritas*”) y salidas (“*axones*”). El valor de la salida se calcula en función del resultado de una suma ponderada de las entradas de forma que si dicha suma supera un umbral pre-establecido entonces la salida es un “1”, en caso contrario la salida es “0”.

Conectando la salida de cada neurona con las entradas de otras neuronas se forma una red neuronal artificial. Inspirándose en un modelo de reforzamiento de las sinapsis entre neuronas biológicas, propuesto por Donald Hebb [6], se vio que estas redes neuronales artificiales se podían entrenar para aprender funciones que relacionaran las entradas con las salidas mediante el ajuste de los pesos que sirven para ponderar las conexiones entre neuronas, por este motivo se pensó que serían mejores modelos para el aprendizaje, la cognición y la memoria, que los modelos basados en la IA simbólica.

“ Una de las críticas más fuertes a todos estos modelos no corpóreos se basa en que un agente inteligente necesita un cuerpo para poder tener experiencias directas con su entorno ”

Sin embargo, los sistemas inteligentes basados en la IA conexionista tampoco necesitan formar parte de un cuerpo ni estar situados en un entorno real y, desde este punto de vista, tienen las mismas limitaciones que los sistemas de IA simbólica.

Por otra parte, las neuronas reales poseen complejas arborizaciones dendríticas con propiedades, no solo eléctricas sino también químicas, nada triviales. Pueden contener conductancias iónicas que producen efectos no-lineales. Pueden recibir decenas de millares de sinapsis variando en posición, polaridad, magnitud.

Además, hoy en día sabemos que en el cerebro hay unas células llamadas *gliales* que regulan el funcionamiento de las neuronas siendo incluso más numerosas que éstas. No existe ningún modelo conexionista que incluya a dichas células por lo que en el mejor de los casos estos modelos son muy incompletos y en el peor de los casos erróneos.

En definitiva, toda la enorme complejidad del cerebro queda muy lejos de los modelos actuales y plantea serias dudas sobre la utilidad de grandes proyectos como el proyecto “*Blue Brain*” de la UE.

Esta inmensa complejidad del cerebro también me hace a pensar que la llamada “*singularidad*” (futuras super-inteligencias artificiales que superarán con mucho la inteligencia humana en un plazo de unos treinta años) es una predicción con muy poco fundamento.

Otra modelización bio-inspirada, también compatible con la hipótesis SSF, es la computación evolutiva. Los éxitos de la biología evolucionando organismos complejos, hicieron que desde los años 60 algunos investigadores se plantearan la posibilidad de imitar la evolución con el fin de que los programas de ordenador mejorasen automáticamente las soluciones a los problemas para los que habían sido programados, mediante un proceso evolutivo. La idea es que estos programas, gracias a operadores de mutación y cruce de “*cromosomas*” que modelizan a los programas, generan nuevas generaciones de programas modificados cuyas soluciones son mejores que las de los programas de las generaciones anteriores.

Dado que podemos considerar que el objetivo de la IA es la búsqueda de programas capaces de producir conductas inteligentes, se pensó que se podría usar la programación evolutiva para encontrar dichos programas dentro del espacio de programas posibles. La realidad es mucho más compleja y esta aproximación tiene muchas limitaciones aunque ha producido excelentes resultados en problemas de optimización y también en la invención de nuevos dispositivos, como por ejemplo un nuevo tipo de antena.

Una de las críticas más fuertes a todos estos modelos no corpóreos se basa en que un agente inteligente necesita un cuerpo para poder tener experiencias directas con su entorno (decimos que el agente está *situado* en su entorno) en lugar de que un programador proporcione descripciones abstractas de dicho entorno codificadas mediante un lenguaje de representación de conocimientos.

Sin un cuerpo, estas representaciones abstractas no tienen contenido semántico para la máquina. Sin embargo, mediante la interacción directa con el entorno el agente puede relacionar las señales que percibe mediante sus sensores con representaciones simbólicas generadas a partir de lo percibido.

Algunos expertos en IA [7] incluso llegaron a afirmar que no era ni siquiera necesario generar dichas representaciones internas, es decir que no es necesario que un agente tenga que tener una representación interna del mundo que le rodea ya que, en palabras de Brooks, el propio mundo es el mejor modelo posible de sí mismo y la mayor parte de las conductas inteligentes no requieren razonamiento sino que emergen a partir de la interacción entre el agente y su entorno.

Esta idea generó mucha polémica y el mismo Brooks unos años más tarde admitió que hay muchas situaciones en las que una representación interna del mundo es necesaria para que el agente tome decisiones racionales.

La aproximación *corpórea* con representación interna ha ido ganando terreno en la IA y actualmente muchos la consideramos imprescindible para avanzar hacia inteligencias de tipo general. De hecho basamos una gran parte de nuestra inteligencia en nuestra capacidad sensorial y motora.

En otras palabras, el cuerpo da forma a la inteligencia [8] y, por lo tanto, sin cuerpo no puede haber inteligencia de tipo general. Esto es así porque el *hardware* del cuerpo, en particular los mecanismos del sistema sensor y del sistema motor, determinan el tipo de interacciones que un agente puede realizar.

A su vez, estas interacciones conforman las habilidades cognitivas de los agentes dando lugar a lo que se conoce como *cognición situada*. Es decir, se sitúa a la máquina en entornos reales, como ocurre con los seres humanos, con el fin de que tengan experiencias interactivas que, eventualmente, les permitan llevar a cabo algo similar a la teoría del desarrollo cognitivo de Piaget [9], según el cual un ser humano sigue un proceso de maduración mental por etapas y quizá los distintos pasos de este proceso podrían servir de guía para diseñar máquinas inteligentes.

Estas ideas han dado lugar a una nueva subárea de la IA llamada robótica basada en el desarrollo cognitivo (*Developmental Robotics*) [10] que en mi opinión ofrece interesantes oportunidades para enfrentarse al problema más importante de la IA: la adquisición de conocimientos de sentido común. Dotarlas de sentido común es el requerimiento fundamental para que las máquinas actuales dejen de tener inteligencias artificiales especializadas del tipo “*débil*” y empiecen a tener inteligencias artificiales de tipo general.

4. El problema de dotar de sentido común a las máquinas

Posiblemente, la lección más importante que hemos aprendido a lo largo de los casi 60 años de existencia de la IA es que lo que parecía más difícil (diagnosticar enfermedades, jugar al ajedrez a nivel de gran maestro, etc.) ha resultado ser relativamente fácil y lo que parecía más fácil ha resultado ser lo más difícil. La explicación a esta aparente contradicción hay que buscarla precisamente en la dificultad de dotar a las máquinas de conocimientos de sentido común.

Los conocimientos de sentido común no se suelen recoger en libros o enciclopedias y sin embargo todos poseemos una gran cantidad de estos conocimientos fruto de nuestras vivencias y experiencias interactuando con

“ Posiblemente, la lección más importante que hemos aprendido a lo largo de los casi 60 años de existencia de la IA es que lo que parecía más difícil (diagnosticar enfermedades, jugar al ajedrez a nivel de gran maestro, etc.) ha resultado ser relativamente fácil y lo que parecía más fácil ha resultado ser lo más difícil ”

nuestro entorno ya que la cognición humana es una cognición situada y corpórea. Por ejemplo, sabemos que para mover un objeto atado a una cuerda hay que tirar de la cuerda en vez de empujarla.

El cuerpo, en particular los detalles del sistema sensor y del sistema motor, determina el tipo de situaciones que un agente puede percibir y abordar. A su vez, estas situaciones conforman las habilidades cognitivas de los agentes. Consecuentemente, para especificar concretamente dichas habilidades cognitivas es necesario tener en cuenta las interacciones del agente con su entorno.

Las aproximaciones no corpóreas no permiten interacciones directas con el entorno por lo que, inevitablemente, dan lugar a falsos problemas y, por lo tanto, a falsas soluciones. Tienden a definir los problemas en términos de tareas en un entorno especificadas desde una perspectiva abstracta de objetos y relaciones. Las capacidades cognitivas no se deberían estudiar haciendo abstracción del sistema sensor y el sistema motor.

5. Los sistemas integrados como paso previo hacia la IA de tipo general

Las capacidades más complicadas de alcanzar son aquellas que requieren interactuar con entornos no restringidos: percepción visual, comprensión del lenguaje, razonar con sentido común y tomar decisiones con información incompleta.

Diseñar sistemas que tengan estas capacidades requiere integrar desarrollos en muchas áreas de la IA. En particular, necesitamos lenguajes de representación de conocimientos que codifiquen información acerca de muchos tipos distintos de objetos, situaciones, acciones, etc., así como de sus propiedades y de las relaciones entre ellos. También necesitamos nuevos algoritmos que, en base a estas representaciones, puedan responder de forma robusta y eficiente preguntas sobre prácticamente cualquier tema.

Finalmente, dado que necesitarán conocer un número prácticamente ilimitado de co-

sas, estos sistemas deberán ser capaces de aprender nuevos conocimientos de forma continua a lo largo de toda su existencia.

En definitiva, además de progresos individuales en cada una de estas áreas, debemos también diseñar sistemas que integren percepción, representación, razonamiento, acción y aprendizaje.

Éste es un problema muy importante en IA ya que todavía no sabemos cómo integrar todos estos componentes de la inteligencia, necesitamos arquitecturas cognitivas que integren a estos componentes de forma adecuada. Estos sistemas integrados son un paso previo fundamental para conseguir algún día inteligencias artificiales de tipo general.

6. El futuro de la IA

Entre las actividades futuras, creo que los temas de investigación más importantes seguirán estando basados en lo que podemos llamar “*massive data-driven AI*”, es decir en explotar la posibilidad de acceder a cantidades masivas de datos y poder procesarlos con hardware cada vez más rápido con el fin de descubrir relaciones entre ellos, detectar patrones y realizar inferencias y aprendizaje mediante modelos probabilísticos. Ejemplos claros de ello son el sistema Watson [11] de “*deep question-answering*” y los sistemas de “*deep learning*” [12].

Sin embargo estos sistemas basados en el análisis de enormes cantidades de datos deberán, en el futuro, incorporar módulos que permitan explicar cómo se ha llegado a los resultados y conclusiones propuestas ya que la explicabilidad es una característica irrenunciable para los usuarios. Actualmente, la principal limitación de los sistemas basados en *deep learning* es que son “cajas negras” sin capacidad explicativa.

Otras técnicas más clásicas de IA que seguirán siendo objeto de investigación son los sistemas multiagente, el razonamiento espacial, la planificación de acciones, el razonamiento basado en la experiencia, la visión artificial, la comunicación multimodal persona-máquina, la robótica humanoide y

animaloide y en particular las nuevas tendencias en robótica basada en el desarrollo (*developmental robotics*) [10].

En el caso de la robótica existe una nueva propuesta basada en el concepto de computación en la nube (*cloud computing*) que se conoce como *Cloud Robotics* [13]. La ventaja reside en que cada robot podrá aprender de las experiencias del resto de robots a los que esté interconectado a través de la nube, lo cual debería permitir un progreso rápido de aprendizaje y mejora de las capacidades de cada uno de los robots. Cuantos más robots formen parte del “*cloud*” más rápido deberían aprender y mejorar. La técnica de aprendizaje más apropiada para este enfoque sería el “*deep learning*” ya que los robots deberán aprender en base a grandes cantidades de datos.

También veremos progresos significativos gracias a las aproximaciones biomiméticas para reproducir en máquinas el comportamiento de animales tales como insectos. No se trata únicamente de reproducir el comportamiento de un animal sino de comprender cómo funciona el cerebro que produce dicho comportamiento. Se trata de construir y programar circuitos electrónicos que reproduzcan las secuencias de órdenes que el cerebro genera para, por ejemplo, producir movimientos (de las alas, las patas, etc.).

Algunos biólogos están interesados en los intentos de fabricar un cerebro artificial lo más complejo posible porque es una manera de comprender mejor el órgano y los ingenieros buscan información biológica para hacer diseños más eficaces. Mediante la biología molecular es posible identificar qué genes y qué neuronas juegan un papel en estos movimientos.

Otra importante fuente de inspiración para la IA es la ciencia de materiales, y en particular los nanomateriales. Por ejemplo, para el desarrollo de músculos artificiales una posible tecnología consiste en intercalar capas de caucho de silicio con capas de polímero electro-activo de tal forma que el conjunto flexiona al aplicar un campo eléctrico [14].

“ Veremos progresos significativos gracias a las aproximaciones biomiméticas para reproducir en máquinas el comportamiento de animales tales como insectos ”

También hay resultados interesantes en cartílagos artificiales construidos mediante filamentos de polímeros con moléculas atractoras de agua para mimetizar las propiedades de los cartílagos naturales [15].

Finalmente, se están usando compuestos de caucho de silicón cargado con nanopartículas de níquel para pieles artificiales ya que la resistencia del compuesto disminuye con la presión a que es sometido, o también sensores capacitivos, es decir cuya capacidad cambia con la presión [16].

Esta aproximación pluridisciplinar a la IA mimetizando a la biología y el uso de resultados en ciencia de materiales puede producir un efecto sinérgico que cambie profundamente la naturaleza de la IA e incluso quizá nuestra comprensión de qué es la inteligencia.

En cuanto a las aplicaciones, algunas de las más importantes seguirán siendo aquellas relacionadas con la web, los video-juegos, y los robots autónomos (en particular los automóviles con pilotaje automático, sillas de ruedas autónomas, robots domésticos, robots de exploración de planetas, etc.).

La economía y la sociología también usarán cada vez más modelos de IA, en particular *agent-based models* para simular interacciones entre grandes cantidades de agentes y llevar a cabo predicciones.

También en biología molecular y farmacología veremos un uso cada vez más intensivo de la IA. Por ejemplo, muchos fármacos tienen efectos secundarios inesperados, por ejemplo la Viagra que inicialmente se desarrolló para hipertensión resultó ser muy efectiva para tratar la disfunción eréctil; o la lovastatina, un efectivo tratamiento de la hipercolesterolemia, ha resultado ser un potente antibiótico.

Pues bien, en lugar de esperar que estos beneficiosos efectos secundarios se descubran por casualidad, investigadores en farmacología aplican técnicas de IA para predecir qué fármacos ya existentes pueden tener otros usos terapéuticos, y en particular predecir las propiedades antibióticas de dichos fármacos con la consiguiente ganancia de tiempo que ello supone frente al desarrollo y aprobación de un nuevo fármaco.

Por último, las aplicaciones de la IA Inteligencia Artificial al arte (artes visuales, música, danza, narrativa) están cambiando de forma importante la naturaleza del proceso creativo. Los ordenadores ya no son solamente herramientas de ayuda a la creación, los ordenadores empiezan a ser agentes creativos.

Este punto de vista ha dado lugar a una nueva y muy prometedora área en Inteligencia Artificial denominada *Creatividad Computacional* [17].

7. A modo de colofón

En cualquier caso, por muy inteligentes que lleguen a ser las futuras inteligencias artificiales, de hecho siempre serán distintas a las inteligencias humanas ya que, tal como hemos argumentado, el desarrollo mental que requiere toda inteligencia compleja depende de las interacciones con el entorno y estas interacciones dependen a su vez del cuerpo, en particular del sistema perceptivo y del sistema motor.

Ello junto al hecho de que las máquinas muy probablemente seguirán procesos de socialización y culturización distintos a los nuestros incide todavía más en el hecho de que, por muy sofisticadas que lleguen a ser, serán inteligencias distintas a las nuestras.

El hecho de ser inteligencias ajenas a la humana y, por lo tanto, ajenas a los valores y necesidades humanas nos debería hacer reflexionar sobre posibles limitaciones éticas al desarrollo de la Inteligencia Artificial.

Referencias

[1] A. Newell, H. Simon. Computer Science as Empirical Inquiry: Symbols and Search, *Communications of the ACM* 19(3): pp. 113-126, 1976.

[2] A.M. Turing. Intelligent Machinery, National Physical Laboratory Report, 1948. Reimpreso en B. Meltzer y D. Michie (eds), *Machine Intelligence* 5, Edinburgh University Press, 1969.

[3] A.M. Turing. Computing Machinery and Intelligence, *Mind*, Vol. LIX(236): 433-460, 1950.

[4] J.R. Searle. Minds, Brains, and Programs, *Behavioral and Brain Sciences* 3(3): 417-457, 1980.

[5] W.S. McCulloch, W. Pitts. A Logical Calculus of Ideas Immanent in Nervous Activity", *Bulletin of Mathematical Biophysics* 5: 115-133, 1943.

[6] D.O. Hebb. *The Organization of Behavior: A Neuropsychological Theory*, New York: John Wiley, 1949.

[7] R.A. Brooks. Intelligence Without Reason. IJCAI'91 *Proceedings of the 12th International joint conference on Artificial intelligence - Volume 1*: 569-595, 1991.

[8] R. Pfeifer, J. Bongard. *How the body shapes the way we think*. The MIT Press, 2007.

[9] B. Y. Inhelder, J. Piaget. *The Growth of Logical Thinking from Childhood to Adolescence*, New York: Basic Books, 1958.

[10] J. Weng, J. McClelland, A. Pentland, O. Sporns, I. Stockman, M. Sur, E. Thelen. Autonomous Mental Development by Robots and Animals. *Science* 291: 599-600, 2001.

[11] D. Ferrucci, E. Brown, J. Chu-Carroll, J. Fan, D. Gondek, A.A. Kalyanpur, A. Lally, J.W. McDock, E. Nyberg, J. Prager, N. Schlaefer, C. Welty. Building Watson: An Overview of the DeepQA Project. *AI Magazine* 31(3): 59-79, 2010.

[12] Y. Bengio. Learning Deep Architectures for AI. *Foundations and Trends in Machine Learning* 2(1): 1-127, 2009.

[13] J.J. Kuffner. Cloud-Enabled Humanoid Robots. *IEEE-RAS International Conference on Humanoid Robots*, 2010.

[14] I. Anderson, S. Rosset, T. McKay, H. Shea. Stack design for portable artificial muscle generators: is it dangerous to be short and fat?. *Electroactive Polymer Actuators and Devices (EAPAD)*, San Diego, 2014.

[15] M. Chen, W.H. Briscoe, S.P. Armes, J. Klein. Lubrication at Physiological Pressures by Polyzwitterionic Brushes, *Science* 323(5922): 1698-1701, 2009.

[16] A. Schmitz, P. Maiolino, M. Maggiali, L. Natale, G. Cannata, G. Metta. Methods and technologies for the implementation of large-scale robot tactile sensors, *IEEE Transactions on Robotics* 27(3), 389-400, 2011.

[17] S. Colton, R. Lopez de Mantaras, O. Stock. Computational Creativity: Coming of Age. *AI Magazine* 30(3): 11-14, 2009.