

HAWKINS, J., (2004): *On intelligence*; New York, Times Books.
ISBN: 0-8050-7456-2

Ros Rodríguez, Silvia

Universidad de Cádiz
Departamento de Filología Francesa e Inglesa
Facultad de Filosofía y Letras
Avda. Gómez de Ulla, 1
11003 Cádiz
Tel: (+34) 678436917
e-mail: silvia.ros@uca.es

Recibido junio 2005; aceptado septiembre 2005

BIBLID[1133-682X (2005) 13; 179-185]

1. Sinopsis

Jeff Hawkins, uno de los arquitectos computacionales de más éxito y reconocido prestigio, aún en esta obra, *On intelligence*, los avances en neurociencia y computación, a través de una teoría que ofrece una explicación de cómo funciona realmente el cerebro humano y en qué consiste la inteligencia, con el fin de crear máquinas inteligentes.

Con un atractivo estilo, apto para un amplio abanico de lectores, y con una detallada explicación repleta de ejemplos ilustrativos, este trabajo constituye un punto de referencia para aquellos que intentan implementar computacionalmente el comportamiento o la inteligencia humana.

Aunque la presente obra disiente de los planteamientos de las primeras teorías computacionales, ya que la pretensión de Hawkins no es tanto centrarse en una inteligencia "artificial", sino en el funcionamiento de la inteligencia "real", su fin último es coincidente, pues el resultado al que se pretende llegar es la creación de máquinas que exhiban inteligencia. Por esta razón, y en un intento de comparar los procedimientos que, aunque por distintos caminos, nos lleven a un mismo resultado, este trabajo es de interés central en el marco general del cognitivismo y, más concretamente en el campo de la lingüística, para los que comparáramos la teoría mentalista o computacional de la lengua.

El primer capítulo ofrece una visión general del objetivo de esta obra, en particular, y de los intereses computacionales de Hawkins, en general, estableciendo una separación respecto a los planteamiento de la Inteligencia Artificial, cuya sede fundacional está representada por el MIT.

A diferencia de los científicos de este campo, Jeff Hawkins considera que el punto clave para la construcción de máquinas inteligentes en la comprensión del funcionamiento del cerebro humano, aspecto que dentro de la IA no es considerado necesario para conseguir simular computacionalmente el comportamiento inteligente humano.

El segundo capítulo se centra en la teoría conexionista o de redes neuronales, que se interesa en aprender qué tipo de comportamientos pueden exhibir las neuronas agrupadas en redes. El hecho de que el conocimiento derivado de las redes y la memoria estuviesen distribuidos a través de su conectividad se asemejaba bastante al funcionamiento de los cerebros reales y esta interacción de neuronas podría mostrar ciertas propiedades de la inteligencia hasta entonces desconocidas. No obstante, para Hawkins, este marco no es apropiado por no cumplir tres condiciones que él considera esenciales para entender el cerebro, a saber, la inclusión del factor tiempo en su funcionamiento, la importancia de la retroalimentación y la consideración de su arquitectura física. La falta de atención a estos factores se evidencia asimismo en la mayoría de IA, entre los científicos cognitivos y en la gran mayoría de los neurocientíficos.

En el fondo tanto la teoría de redes neuronales como la IA presentarían el mismo problema, y es que ambas se centran en el comportamiento, considerando que éste constituye una característica esencial de la inteligencia cuando sólo es una de sus manifestaciones. Sin embargo, Hawkins dice compartir el planteamiento funcionalista de estas corrientes, defendiendo que la inteligencia o la mente es una propiedad que deriva de la organización del cerebro, y no una característica inherente especial e inexplicable del mismo, por lo que cabe la posibilidad de construir máquinas inteligentes, aunque desde una perspectiva diferente, donde la investigación sobre la inteligencia parta del cerebro, y más concretamente del mecanismo con que la naturaleza nos ha dotado para ser seres inteligentes: el neocórtex.

Así, el neocórtex, considerado por Hawkins como la clave de la inteligencia, es el punto en el que se centra el siguiente y tercer capítulo.

Aunque la arquitectura del córtex está dictaminada por los genes, las distintas áreas funcionales que existen en él se caracterizan por su plasticidad, haciendo que la estructura de este sistema sea muy flexible.

En cuanto a su funcionamiento, partiendo de la pretensión de encontrar semejanzas y no diferencias entre estas áreas, Hawkins nos dice que la información que obtenemos a través de los sentidos crearía patrones, los cuales constituyen el input del córtex, siendo allí procesados por un mismo algoritmo cortical, una función común a todas las áreas, y dando lugar a un modelo del mundo, bastante parecido al mundo real, que es retenido en nuestra memoria a través de la formación y reforzamiento de conexiones sinápticas.

Tras el acercamiento al procesamiento de la información, el capítulo cuarto aborda cómo se almacena este conocimiento, es decir, en qué consiste nuestra memoria.

Según el autor del libro, es precisamente el hecho de que la información que entra en el córtex sea almacenada y recuperada para resolver problemas, lo que distingue al cerebro de un ordenador, ya que este último sólo realiza cálculos.

Aunque los ordenadores también posean memoria, no cumpliría cuatro características que según Hawkins serían distintivas de la memoria neocortical: el almacenamiento de secuencias de patrones, el recordar los patrones autoasociativos, su almacenamiento en forma invariante y su distribución de manera jerárquica.

permitiendo al ser vivo aprender del medio y adaptar su comportamiento (creación del neocórtex, común a todos los mamíferos), y iii) expansión del neocórtex, que nos permite aprender mucho más de nuestro alrededor, e invención del lenguaje, por el que podemos transmitirlo a nuestros sucesores, lo cual nos convierte en la especie mejor adaptada del planeta.

Un segundo interrogante sería en qué consiste la creatividad, definida aquí como una simple operación de predicciones por analogía, que se produce en un continuo y en la que la memoria y la predicción juegan un importante papel. Aunque todos tenemos la capacidad de ser creativos, existen diferencias entre nosotros debidas tanto a cuestiones de naturaleza, como es la variación física que exhiben los cerebros, como de crianza, ya que no todos tenemos las mismas experiencias del mundo ni, consecuentemente, derivamos de ellas las mismas predicciones.

En cuanto al tema de la conciencia, la divide en dos categorías: lo que se corresponde a la noción cotidiana de conciencia de uno mismo (darse cuenta de algo), y lo que denomina *qualia*, término con el que alude a la idea de que los sentimientos asociados a las sensaciones sean independientes del input sensorial.

En lo que respecta al concepto de realidad, éste dependería de la adecuación con la que nuestro modelo cortical refleja la verdadera naturaleza del mundo, teniendo en cuenta que la mayoría de lo que percibimos no proviene, al menos directamente, de los sentidos, sino de nuestro modelo de memoria interno.

Por último, la existencia de la imaginación se explica por el hecho de que nuestras predicciones se conviertan en input, derivando las consecuencias de las mismas sin necesidad de que ocurran realmente.

Ya en el último capítulo, se intenta mostrar de qué forma podría implementarse este algoritmo cortical en una máquina inteligente, teniendo en cuenta, no obstante, que ésta no exhibiría comportamiento humano, debido a que la mente humana no está creada sólo por el neocórtex, sino que influyen también los sistemas emocionales del cerebro primitivo y la complejidad del cuerpo humano.

De esta manera, la inteligencia no se evidenciaría en un comportamiento humano, sino en la habilidad predictiva de una memoria jerárquica que funcionase como la nuestra, sin necesidad de producir los mismos resultados, ya que su modelo del mundo podría diferir del nuestro. Los elementos clave serían la capacidad de almacenamiento y la conectividad o procesamiento en paralelo.

Cerrando este trabajo, y con la intención de dar validez al modelo expuesto, encontramos un apéndice con una serie de predicciones que derivan del mismo, para que sean probadas experimentalmente.

2. Comentario

Como señalamos al principio, aunque la intención de Hawkins con este trabajo es establecer las bases para la creación de máquinas inteligentes, y en esto coincide con el resto de científicos que trabajan en IA, los medios para alcanzar dicho objetivo son muy diferentes a los utilizados hasta ahora.

La idea clave en torno a la cual gira su trabajo es exponer una teoría que explique en qué consiste la inteligencia real y cómo está producida por el cerebro; una vez dado este paso pre-

vio, que se considera esencial, podremos conseguir una inteligencia artificial con fundamentos sólidos. Los anteriores intentos de simular la inteligencia humana habrían fracasado precisamente por no seguir este proceder.

Ante esta opinión cabe señalar, en primer lugar, que, desde sus inicios en los años cincuenta, la inteligencia artificial ha producido logros que, si bien no llegan ni lejanamente a la producción total de una mente completa, sin embargo replican con bastante propiedad muchas de las capacidades de la mente humana, por lo que los avances realizados son sustanciales; en segundo lugar, no podemos culpar de esta falta de éxito total al hecho de que no se haya intentado entender primeramente el funcionamiento real del cerebro, ya que este proceder no nos proporcionaría ninguna ventaja sino que, además de suponernos una gran traba, constituiría, asimismo, un paso innecesario en nuestra creación de máquinas inteligentes.

El inconveniente que se nos plantea desde el punto de vista adoptado en esta obra, sería la necesidad de adentrarnos en el terreno de la neurobiología, donde los datos e investigaciones no son nada concluyentes y todavía queda mucho camino para entender exactamente de qué manera el cerebro genera inteligencia.

Por otro lado, aunque conociésemos perfectamente el funcionamiento neurológico, no sería necesario, como hemos señalado, imitarlo realmente. Con una analogía más visual, podemos decir que el hecho de que un reloj de arena dé la hora, no obliga a los creadores de relojes digitales a crear modelos que simulen la caída de la arena para que realice la función de dar la hora.

Y es que la inteligencia que se pretende simular computacionalmente es precisamente “artificial”: su base física no será el cerebro humano y, por tanto, no será indispensable que funcione de la misma manera que éste, sólo que produzca los mismos resultados.

Este punto de vista ha sido ejemplificado por algunos autores con los casos de la luz y el movimiento, entre otros; la luz artificial no está producida por el sol ni, por tanto, está creada de la misma manera que la luz natural, y sin embargo es luz e ilumina; el caso del movimiento de un automóvil es otro ejemplo de artificialidad, ya que está creado por el hombre, pero se considera que su movimiento es real, independientemente de que su proceder sea muy distinto al del movimiento humano.

Ahora bien, la inteligencia “artificial” que estos científicos pretenden crear es tan “real” como la inteligencia producida por nuestros cerebros; la diferencia entre ellas residiría más bien en la dicotomía “artificial-natural”, por lo que podemos defender que la *inteligencia artificial* es *inteligencia real* sin necesidad de ser como la *natural*.

Más allá del camino que escojamos para crear inteligencia, el problema esencial al que se enfrentan todos los que trabajan en este campo es definir en qué consiste ésta, para crear máquinas que exhiban dicha cualidad.

Pero también aquí Hawkins difiere del resto de programadores informáticos de este campo, ya que estima equivocada la consideración del comportamiento como evidencia de la inteligencia de un determinado sistema, pues en el ser humano la comprensión o entendimiento de algún fenómeno no tiene que ir asociado necesariamente a una expresión comportamental.

No obstante no existe contradicción entre el hecho de que el comportamiento no sea la característica definitoria esencial de la inteligencia y que pueda, sin embargo, ser tomada como una prueba de la misma.

Llama la atención el hecho de que estas objeciones expuestas por Hawkins, contradicen, de manera incomprensible, el pretendido funcionalismo que nuestro autor dice compartir.

Si, tal como él mismo afirma, hay que aceptar la posibilidad de crear máquinas inteligentes en tanto que la mente no es una característica inherente inexplicable del cerebro, sino una propiedad que deriva del mismo, no encontramos razón para no aceptar los puntos por él criticados, ya que:

- i) si lo que interesa realmente es la función, las diferencias de arquitectura carecen de importancia, y, en consonancia con esta idea,
- ii) se podría considerar que lo que es indistinguible en su comportamiento es idéntico en su naturaleza.

Como última apreciación, en cuanto a la estructura y funcionamiento de la mente, no parece probable, aunque como dice este autor sea un planteamiento elegante y coherente por su simpleza, la existencia de un mecanismo común, de ese algoritmo cortical, que explique y unifique todos los sistemas.

Más bien parece ser que nuestra inteligencia humana no se basa en un único funcionamiento, sino que existen tanto sistemas que nos permiten adquirir conocimiento a través de descargas sinápticas en distintas conexiones, trabajando en paralelo, como sistemas configurados en serie (como los módulos), pre-programados por nuestra evolución; en este último caso, la investigación ofrece indicios certeros de la existencia de hormas, algunas heredadas por completo, como parece ser la visual, y otras sólo en parte, como parece ser la lingüística, en la que se produce una interacción entre un sistema innato y el contacto con el entorno.

Para concluir, a pesar de las diferencias que se producen en el nivel metodológico, la presente teoría y los planteamientos de la IA comparten las mismas metas cognoscitivas, por lo que sus diferencias pueden resolverse en el denominado, dentro de la jerarquía entre teorías científicas, nivel axiológico. De esta manera, siempre podemos encontrar la ventaja de poder usar nuestro conocimiento de los métodos disponibles de investigación como una herramienta para evaluar la viabilidad de metas cognoscitivas propuestas.

No obstante, más allá de esta contribución, resulta difícil encontrar una razón de peso para apoyar una teoría que, aunque en lo que concierne a la descripción y funcionamiento de la mente humana ofrece muchos puntos esclarecedores, no parece, sin embargo, ser la mejor vía de acceso a la creación de máquinas inteligentes, pues, como se ha señalado, el esfuerzo que implica adentrarse en el terreno de la neurobiología, además de ser innecesario, no reporta necesariamente beneficios para los propósitos de la Inteligencia Artificial.

Es por ello que las actuaciones realizadas hasta el momento en este campo parecen estar mejor encaminadas y es muy probable que, siguiendo su metodología, podamos alcanzar mayores éxitos, con menor costo, en la producción de máquinas que exhiban inteligencia.

Es más, este paradigma nos ofrece un productivo viaje de ida y vuelta de tipo dialéctico, ya que, una vez que la inteligencia humana se sustituya por una máquina, dicha máquina constituirá el mejor modelo para entender el funcionamiento del computador humano: el cerebro. Así, el viaje de ida, de formalización, nos llevará a la construcción de máquinas dotadas de inteligencia, y el viaje de vuelta, de explicación, ofrecerá a los científicos cognoscitivos el modelo más útil provisto hasta ahora para sus investigaciones.

3. Referencias

- FODOR, J. & Z. W. PHYLYSHYN (1988). "Connectionism and Cognitive Architecture: A Critical Analysis", *Cognition*, 28: 3-71.
- GRAUBARD, S.R. (comp), (1988): *The Artificial Intelligence Debate. False Starts, Real Foundations*, Cambridge, Mass. MIT Press. (Versión española, 1993, *El nuevo debate sobre la inteligencia artificial. Sistemas simbólicos y redes neuronales*, Barcelona, Gedisa).
- GUTIERREZ, C., (1993-97): *Epistemología e informática*, <http://www.claudiogutierrez.com/EeI.html>.
- MCCORDUCK, P., "Inteligencia artificial: un aperçu", en Stephen R. Graubard, (ed) 1988.
- PAPERT, S., "¿Una sola IA o muchas?", en Stephen R. Graubard, (ed) 1988.
- PINKER, S., (1998): *How the mind works*, London, Penguin Books.
- SOKOLOWSKI, R. (1988). "Natural and Artificial Intelligence", *Daedalus* 117 (1): 45-64.
- TURING, A.M., (1950): "Computing Machinery and Intelligence", *Mind* 49: 433-460.