

**SEMINARIO DE LA CÁTEDRA CTR
SEGUNDA SESIÓN BÁSICA
CURSO 2006-2007**

Temática: Hombre/Neurología

15 de febrero de 2007

**COMPUTACIÓN Y NEUROLOGÍA:
¿ES EL HOMBRE UN COMPUTADOR?**

***Ponente: Prof. Antonio Crespo
Universidad Nacional de Educación a Distancia***

***Discussant: Prof. David Travieso
Universidad Autónoma de Madrid***

DOCUMENTO MARCO:

***Modelos psicológicos-computacionales de la actividad de
la mente: Ordenadores y Neuronas en la explicación de la
mente,
Por Antonio Crespo***

Modelos Psicológicos-Computacionales de la Actividad Cognitiva. Ordenadores y Neuronas en la Explicación de la Mente

Antonio Crespo, PhD
Profesor Titular
Departamento de Psicología Básica II, UNED
C/ Juan del Rosal 10, 28040-Madrid
acrespo@psi.uned.es

En psicología, se descubren dos aproximaciones computacionales al estudio científico de la actividad cognitiva: la simbólica y la conexionista. La primera entiende que la mente puede contemplarse como un sistema funcionalmente análogo a un software, cuyo propósito es la manipulación de símbolos. La segunda se inspira en el funcionamiento cerebral y, sin reducirla a pura fisiología neuronal, la contempla como un sistema de procesamiento distribuido en paralelo (PDP), en donde multitud de unidades básicas de computación actúan al unísono generando complejas estructuras de redes neuronales artificiales (RNAs). Ordenadores y neuronas constituyen, por tanto, el marco teórico de referencia sobre el que se elaboran diversas teorías y modelos psicológicos de la cognición humana.

Ni el cognitivismo inspirado en el ordenador desea realizar explicaciones de cómo funcionan dichas máquinas, ni los conexionistas pretenden elaborar una teoría fisiológica del funcionamiento cerebral.

*¿Es apropiado entender el funcionamiento de la mente como un software? ¿En qué medida el conocimiento de la actividad neuronal sirve para elaborar modelos del funcionamiento mental? ¿Puede una máquina, adecuadamente programada, llegar a disponer de estados mentales y, como tal, realizar actos intencionales semejantes a los de los humanos? La mente, entidad silenciosa y oculta, pero a la par enormemente diligente, ha captado desde siempre nuestro interés. La forma de entender el funcionamiento mental ha sido diversa a lo largo de los tiempos. La psicología, de forma pareja con la inteligencia artificial, la filosofía, la lingüística o la moderna

neurociencia, son las disciplinas que han contribuido decisivamente en la elaboración de modelos y teorías desde los que explicar la cognición.

En este documento, analizaremos dos paradigmas diferentes que, desde la psicología, han regido la forma de elaborar modelos teóricos sobre el funcionamiento mental. La perspectiva computacional simbólica inspirada en el ordenador serial que manipula símbolos, y la conexionista inspirada en el funcionamiento cerebral.

ARQUITECTURAS MENTALES

Los modelos teóricos utilizados para explicar el funcionamiento cognitivo se tienden a representar mediante lenguajes formales abstractos, al igual que un arquitecto representa sobre papel una determinada edificación. Por su similitud, en ciencia cognitiva se habla de *arquitecturas mentales* para referirse a las especificaciones de diseño que cumplen dichos modelos.

* *Este documento ha sido elaborado para la sesión básica del seminario 2006-2007 de la Cátedra Ciencia, Tecnología y Religión de la Universidad de Comillas (Computación y Neurología: ¿Es el hombre un computador?). Incluye pasajes extraídos y adaptados a partir del libro de Antonio Crespo, titulado *Cognición Humana: Mente, ordenadores y neuronas* (2006, 2ª Edición). Madrid: Editorial Universitaria Ramón Areces.*

Un primer colectivo de científicos ha optado por inspirar sus modelos en los ordenadores seriales. De este modo, al igual que un ordenador distingue entre el nivel lógico, programa o software y el sustrato físico o hardware, la mente puede entenderse como un *programa o software que manipula información en formato simbólico* y que se implementa sobre un cerebro que hace las veces de *hardware*. Esta perspectiva se conoce como *metáfora o analogía del ordenador*, y los modelos elaborados desde ella se refieren como arquitecturas simbólicas o clásicas de la mente.

Alternativamente, un segundo colectivo de científicos entienden que usar el ordenador serial como modelo no es apropiado para comprender el funcionamiento cognitivo. Tal es así que prefieren dirigir su atención al modo en el que funciona el cerebro con el propósito de desarrollar modelos explicativos de la actividad mental. Desde esta perspectiva, la mente se entiende como un sistema de computación que se describe recurriendo a *redes neuronales artificiales (RNAs)*, en donde la información se procesa distribuidamente entre múltiples unidades análogas a las neuronas. Es lo que se conoce como *metáfora o analogía cerebral* y constituye el fundamento de las denominadas arquitecturas conexionistas de la mente.

Independientemente de que se adopte uno u otro paradigma, cuando se usan las metáforas como análogos conceptuales es importante discriminar entre un nivel físico y un nivel funcional. El *nivel físico o de implementación* lo constituyen los componentes específicos de un sistema: chips y circuitos impresos en el caso del ordenador y neuronas en el caso cerebral. Pero este nivel, al menos a priori, no es necesario manejarlo para desarrollar un modelo explicativo. Lo importante es atender al *nivel funcional o lógico*, es decir, al modo operativo que un determinado sistema inteligente adopta, independientemente de su sustrato físico. Dicho de otra forma. Los cognitivistas clásicos y los conexionistas pretenden ambos elaborar modelos del funcionamiento mental, y para ello se inspiran bien en los ordenadores o bien en el cerebro, respectivamente. Sin embargo, no pretenden realizar ni teorías explicativas del funcionamiento físico de un ordenador (parcela que corresponde a la ingeniería) ni teorías explicativas del funcionamiento físico del cerebro (parcela que corresponde a la neurofisiología).

ARQUITECTURAS SIMBÓLICAS: MIRANDO AL ORDENADOR

Las debilidades de la propuesta conductista

En un contexto de enorme tradición empirista y asociacionista, un primer mentalismo tradicional, a mediados y finales del XIX, dirigió su interés bien a los elementos de la conciencia (orientación estructuralista) o bien a su valor de utilidad (orientación funcionalista). El acontecimiento más reseñado fue la fundación del primer laboratorio de psicología en la ciudad de Leipzig por Wundt en 1879. Es cierto que la orientación estructuralista cuantificó los datos observados, sin embargo, el método utilizado para su obtención recurría a la introspección, o auto observación de las propias experiencias mentales, por sujetos que habían sido preparados para ello. El funcionalismo, por su parte, intentó superar ciertas limitaciones del estructuralismo, progresando en los estudios sobre comportamiento animal, desarrollo evolutivo, patologías humanas y psicología aplicada. Además, las limitaciones del método introspectivo fueron superadas por otros medios de recogida de datos, tales como experimentos fisiológicos, pruebas mentales o cuestionarios estandarizados del comportamiento.

En un segundo momento, el advenimiento del conductismo supuso la primera gran revolución en psicología. Su rechazo a la introspección ocasionó que todo tipo de término mentalista fuera descartado o reducido a explicaciones mecanicista de naturaleza reduccionista. El conductismo mantuvo el marco empirista y los principios asociacionistas del mentalismo tradicional, y ello supuso un avance espectacular en el estudio de los procesos de aprendizaje. Además, la contundente crítica al uso de la introspección como método de recogida de datos obligó a que ésta fuera sustituida por el método positivo, lo que permitió a la psicología avanzar hacia una auténtica ciencia objetiva. No obstante, para los conductistas de las primeras décadas del siglo XX, aunque asumían el modelo de arco reflejo pavloviano, no era preciso acudir a analizar el sustrato cerebral. La conducta debía ceñirse a lo observable y cuantificable, y las relaciones entre estímulos ambientales y respuestas públicas era el único objeto de estudio válido. Es lo que se ha denominado metáfora o analogía de la central telefónica (de las de antaño, con clavijas de conexión) en donde el estímulo E del medio es percibido y, tras las conexiones neuronales específicas (las telefonistas insertaría las clavijas necesarias), enviaría la señal nerviosa para emitir la respuesta R.

La idea general subyacente en los estudios desarrollados dentro del paradigma conductista no era otra que la generalización de los principios básicos obtenidos en el estudio del aprendizaje animal, con objeto de hacerlos extrapolables al ser humano. Una de las críticas más habituales fue la concerniente a la artificialidad de las situaciones de laboratorio en el

estudio del comportamiento, aunque esta crítica es aplicable a cualquier otro paradigma experimental de la psicología, incluido el cognitivo. También se ha criticado la supuesta aplicación universal de los principios del aprendizaje y del refuerzo, que se creían válidos para cualquier especie. En este sentido, el desarrollo de los estudios etológicos puso de manifiesto la existencia de pautas fijas de acción en las diversas especies, muy difíciles, por no decir imposible, de modificar o controlar aplicando mecanismos de aprendizaje y refuerzo. Finalmente, las críticas más contundentes e importantes en relación con el contexto en el que nos movemos, y que supusieron la transición hacia el paradigma cognitivo de la psicología, fueron las referentes a la pasividad del sujeto, entendido como un simple receptor de estímulos ambientales, y el destierro o reducción a cadenas de estímulos y respuestas de los procesos psicológicos propiamente humanos como son los procesos cognitivos.

La teoría de la computación y el desarrollo de los primeros ordenadores

En la década de los años cincuenta del siglo pasado, el profundo desarrollo experimentado en disciplinas como la teoría de la computación o la teoría de la información, parecían constituir fuentes de las manaban conocimientos, modelos e incluso un nuevo vocabulario, en cuyo contexto algunos psicólogos (desanimados por las debilidades del aparato teórico conductista) comenzaron a interrogarse sobre si era posible aplicar dichos modelos al estudio de la cognición. A ello unido, la Segunda Guerra Mundial exigió un enorme esfuerzo de desarrollo tecnológico con un objetivo fundamentalmente bélico –cálculo de tablas balísticas y descifrado de mensajes secretos- pero que tuvo como consecuencia el diseño y construcción de los primeros ordenadores que realizaban tareas inteligentes, marcando un nuevo camino que apuntaba hacia la creación de máquinas pensantes que podían ejecutar tareas similares a las que realizaba la mente humana.

La teoría de la computación desarrollada en torno a los años treinta y cuarenta del siglo XX fue decisiva en el surgimiento del cognitivismo simbólico. Un matemático inglés¹, Alan Turing, publicó en 1936 un trabajo en el que exponía cómo era posible descomponer determinados problemas matemáticos en una serie de procedimientos, tal que su resolución era posible mediante la aplicación de un algoritmo. Turing fue más allá del simple papel escrito. Para

concretar sus ideas sobre la computación, ideó una supuesta máquina lógica abstracta –nunca fue construida físicamente– que, siguiendo las instrucciones detalladas en un algoritmo y manejando un conjunto limitado de símbolos, era capaz de resolver cualquier problema que fuera computable en el sentido indicado. Esta es la denominada *máquina Turing* o también *máquina de propósito universal*. En ella, una cinta de longitud ilimitada contiene una serie de casillas en donde la información se representa en forma simbólica, por ejemplo, en formato binario. La cinta puede desplazarse en ambos sentidos mediante un mecanismo de arrastre. Un cabezal puede leer un símbolo, borrarlo o escribirlo sobre la cinta. La idea que se desprende de esta supuesta máquina computacional es que, cualquier tarea que pueda descomponerse en pasos detallados, es solucionable manejando un conjunto de símbolos y aplicando una serie de reglas. Las repercusiones del trabajo de Turing para la emergencia del cognitivismo simbólico son impresionantes. Primero, sus ideas constituyen el antecedente e inspiración más inmediata para la construcción de los primeros ordenadores, pues en realidad el ordenador que usted y yo manejamos es, formalmente, una máquina Turing. Segundo, recurre al manejo de símbolos como medios de representar la información. Tercero, distingue, implícitamente, entre la estructura física del sistema (la cinta, el cabezal, el mecanismo de arrastre, en esencia, el hardware) y el algoritmo que aplica la máquina para resolver la tarea (el software o nivel lógico). Finalmente, y tal vez lo más trascendente con respecto al tema que nos ocupa, sus ideas son el punto de partida para interrogarse acerca de si el funcionamiento cognitivo humano puede ser considerado como una máquina Turing, donde además de un soporte físico, (el cerebro) se descubre un soporte lógico (la mente) que computa sobre conocimientos codificados simbólicamente.

John von Newmann, inspirado en la idea de la máquina Turing, sugirió un computador que incorporaba una CPU que operaba según una lógica binaria (1, 0) y que disponía de posibilidades de almacenamiento de la información y de cálculo lógico y aritmético. Esta CPU actuaría de forma análoga al cabezal en la máquina Turing. A esta CPU se encontraría también interconectado un sistema de memoria, que es donde se depositarían los datos, y las instrucciones propias del programa. Observe que von Newmann materializa las ideas de Turing y esta nueva arquitectura no es sino la que –con mayor grado de sofisticación– implementan los modernos ordenadores digitales. Si Turing fue la inspiración para la construcción de los primeros computadores, a von Newmann se le considera el padre de los computadores modernos. Al igual que para Turing, para von Newmann no pasó inadvertida la analogía entre la

¹ Turing, A. M. (1936). On computable numbers, with an application to the Entscheidungs-Problem. *Proceedings of the London Mathematical Society, Serie 2*, 42, 230-265.

computación de un ordenador y la computación mental. La principal repercusión de su trabajo en el ámbito cognitivo radicaba en contemplar la existencia de mecanismos de almacenamiento o memorias. Esta idea tuvo un impacto espectacular en el diseño de modelos cognitivos basados en el ordenador, pues además de la manipulación de la información simbólica, era preciso contemplar estructuras de almacenamiento y procesos de traspaso de información entre ellas.

Símbolos, computación y ejecutivos centrales

Un ordenador digital es un sistema inteligente artificial que ejecuta un programa o software. Dicho programa recoge un conjunto de instrucciones codificadas una tras otra que el procesador o CPU codifica y *ejecuta serialmente* en su propio lenguaje binario. Esta actividad se reduce a manipular una serie de símbolos, operando sobre ellos según las instrucciones dictadas por el programa o software. El cognitivismo clásico entiende que la mente, al igual que un ordenador, dispone la información recurriendo a símbolos. Un símbolo es una representación de un determinado objeto o conocimiento y puede ser algo muy concreto o eminentemente abstracto. Nuestra mente maneja símbolos continuamente. Ahora mismo, leyendo estas líneas, usted está manipulando mentalmente símbolos, pues es capaz de leer y comprender significativamente el texto conociendo los símbolos del alfabeto. Cuando realiza operaciones aritméticas maneja una serie de símbolos como número o notaciones formales. Hay símbolos tangibles, como una cruz cristiana que simboliza unas creencias religiosas para determinadas personas, o una señal de tráfico que indica un peligro, prohíbe alguna actuación o simplemente informa de algo. La imagen mental que tengo de mi esposa también es un símbolo. El símbolo, entendido desde una perspectiva cognitiva, es un concepto estrechamente vinculado a cómo se representa el conocimiento en la mente.

Pero además, el cognitivismo entiende que la mente no sólo contiene información en forma simbólica, sino que además opera activamente sobre dicha información mediante la computación. El término *computación* se ha tomado prestado del lenguaje de las matemáticas. ¿Qué hace un ordenador cuando computa? Simplemente seguir las instrucciones del programa y realizar la tarea exigida. Una calculadora electrónica computa siguiendo un algoritmo con el que ha sido programada. El indicador electrónico de consumo –que tal vez incorpore su automóvil– está computando, en función de la velocidad que registra y de unas tablas de consumo que se encuentran programadas en su memoria, el gasto de combustible expresado en litros por cada cien kilómetros recorri-

dos. Los modernos sistemas de reconocimiento de voz en los ordenadores –utilizados por muchos bancos actualmente para identificar y ofrecer información automática sobre movimientos y transferencias a sus clientes– no son sino sistemas computacionales que siguen las reglas que establece un programa de reconocimiento fonético.

Nuestra mente parece actuar de manera análoga a un ordenador, es decir, usted también realiza tareas de computación. Ahora mismo, leyendo estas líneas está computando el significado de las palabras. Si yo le pido que me diga el resultado de 7×9 su mente computa. Cuando está decidiendo entre dos alternativas para el fin de semana está computando. Si le indico que Juan es más alto que María y que ésta es más alta que Javier, usted representa esta información en su mente y a partir de ella realiza un cómputo que le permite concluir que Juan es más alto que Javier. Computar es realizar una manipulación o cálculo. El cognitivismo clásico entiende que la mente, igual que un ordenador, realiza tareas de cómputo sobre la información simbólica almacenada en la memoria.

Finalmente, el software que se implementa en un ordenador no se ejecuta milagrosamente, ya que necesita de un mecanismo que lo interprete y que se encargue de aplicar las instrucciones del programa. Este mecanismo es la mencionada CPU o procesador central. De forma análoga, los defensores de las arquitecturas simbólicas entienden necesaria la presencia de una especie de CPU o ejecutivo central que determine y aplique las operaciones cognitivas correspondientes a la tarea propuesta.

En conclusión, el cognitivismo clásico considera la mente funcionalmente –que no físicamente– análoga a un ordenador. Como tal, la mente es un *sistema manipulador de símbolos*, y dicha manipulación consiste en aplicar una serie de operaciones cognitivas sobre representaciones mentales que contienen la información conforme a una serie de reglas².

ARQUITECTURAS CONEXIONISTAS: MIRANDO A LA NEURONA

Los conexionistas surgen *oficialmente* en psicología a mediados de los años ochenta, como una alternativa a los modelos cognitivos de procesamiento de la información de corte simbólico. Los primeros en utilizar el término *conexionismo*, tal como aquí lo

² Véase Pylyshyn, Z. W. (1989). Computing in cognitive science. En M. I. Posner (Ed.), *Foundations of cognitive science* (pg 51-91). Cambridge, MA: MIT Press.

entendemos, fueron Feldman y Ballard³ en un trabajo clásico de 1982, donde, de forma alternativa a los modelos representacionales basados en el procesamiento serial, proponen la plausibilidad de modelos que representan la información de forma distribuida mediante la participación de multitud de unidades simples que computan simultáneamente. Sin embargo, existe pleno consenso al considerar la obra *Parallel Distributed Processing: Explorations in the Microstructure of Cognition*⁴, publicada en 1986 en dos volúmenes, y rubricada por el denominado grupo PDP, con Rumelhart y McClelland a la cabeza, la referencia o punto de partida de la moderna psicología conexionista. Este libro vendió el primer año de edición treinta mil ejemplares, inédito en una obra científica, y se ha dado en llamar *Biblia del conexionismo*. No obstante, el conexionismo tiene una larga andadura desde la década de los cuarenta del siglo pasado, e incluso antes. Los manuales de Anderson, Rosenfeld y Pellionitz⁵ son excelentes recopilaciones de los trabajos clásicos sobre neurocomputación que han ido, poco a poco, dando forma al enfoque conexionista.

En 1987 Walter Schneider⁶ afirmaba que el conexionismo estaba llamado a ser el cambio paradigmático más importante que las jóvenes generaciones de psicólogos verían a lo largo de sus carreras y, ciertamente, es innegable que el profundo impacto que esta orientación cognitiva ha tenido en la teorización, no sólo psicológica sino también en otras

muchas disciplinas. Como orientación de la ciencia cognitiva, en el desarrollo del conexionismo confluyen aportaciones de numerosos ámbitos de estudio, fundamentalmente de la matemática, física, inteligencia artificial, neurología y psicología.

El conexionismo irrumpe en el corazón mismo del cognitivismo; como consecuencia de ello comparte el supuesto básico de que el objeto de estudio de la psicología es la mente y la actividad propia de ésta es procesar la información. Ahora bien, si el cognitivismo clásico y el conexionismo asumen que la actividad de la mente es el procesamiento ¿en dónde radican las diferencias? Entre otros, en dos aspectos fundamentales: por un lado, los conexionistas proponen una nueva descripción de la estructura de la mente basada en multitud de unidades de procesamiento interconectadas análogas a las neuronas (redes neuronales artificiales, RNA). Pero además, y como consecuencia de lo anterior, se propone también un nuevo funcionamiento, es decir, una nueva manera o forma en el procesamiento, una nueva manera de computar la información: el *procesamiento distribuido en paralelo* (PDP).

La problemática del procesamiento serial

Si los comparamos con las neuronas, los modernos ordenadores digitales tienen una velocidad de procesamiento que supera con creces la transmisión nerviosa entre las mismas. Sin embargo, el cerebro efectúa cálculos y cómputos con una rapidez impresionante debido a que múltiples unidades de computación individuales –neuronas– trabajan cooperativamente al unísono, superando su relativa lentitud individual. Por ello, usar un ordenador como metáfora explicativa de la mente tiene sus limitaciones. El problema es que el ordenador de von Neumann no se parece en nada al funcionamiento cerebral, pues la forma de operar del primero es serial y, por tanto, ejecuta las instrucciones del programa una tras otra. Es lo que se ha denominado el *talón de Aquiles* de la computación inspirada en el ordenador. Parece pues que, para explicar la cognición, es posible dirigir la mirada a otro tipo de modelos cuyo fundamento debe buscarse en la estructura y funcionamiento cerebral, pero no en el ordenador simbólico serial.

Influencias neuropsicológicas y computacionales en el surgimiento del conexionismo

En 1949, el neuropsicólogo Donald Hebb⁷ llamaba la atención sobre la importancia que adquirían los procesos centrales del sistema nervioso, que actuaban como estructuras mediadoras de la conducta. Es

³ Feldman, J.A. y Ballard, D.H. (1982). Connectionist models and their properties. *Cognitive Science*, 6, 205-254.

⁴ Rumelhart, D. E., McClelland, J. L., y the PDP Research Group (1986). *Parallel Distributed Processing: Explorations in the microstructure of cognition* (Vol.1). Foundations. Cambridge, MA: MIT Press.

McClelland, J. L., Rumelhart, D. E., and the PDP research group (1986). *Parallel Distributed Processing: Explorations in the microstructure of cognition* (Vol.2). Psychological and biological models. Cambridge, MA: MIT Press.

⁵ Anderson, J. A., y Rosenfeld, E. (Eds.) (1988). *Neurocomputing*. Cambridge, MA: MIT Press.

Anderson, J. A., y Rosenfeld, E. (Eds.) (1998). *Talking Nets*. Cambridge, MA: MIT Press.

Anderson, J. A., Pellionitz, A. y Rosenfeld, E. (Eds.) (1990). *Neurocomputing II. Foundations of Research*. Cambridge, MA: MIT Press.

⁶ Schneider, W. (1987). Connectionism: Is it a paradigm shift for psychology? *Behavior Research Methods, Instruments and Computers* 19, (2), 73-83.

⁷ Hebb, D. O. (1949). *The organization of behavior*. Nueva York: John Wiley and Sons.

clásica su distinción entre asambleas neuronales transitorias y permanentes: las asambleas transitorias son circuitos neuronales que actúan conectados durante tiempo limitado. Por ejemplo, en el ámbito de la memoria serían responsables del recuerdo temporalmente breve de información, tal como memorizar un número de teléfono hasta que se marca y posteriormente se olvida. Alternativamente, las asambleas permanentes actuarían a modo de circuitos neuronales fijos y establecidos en el tiempo que, siguiendo con el ejemplo, se responsabilizarían de fijar en la memoria la información durante un tiempo ilimitado. Para Hebb, el aprendizaje cerebral que permite configurar dichos circuitos se produce en términos de cambios en las actividades pre y post sinápticas. El *principio sináptico de Hebb* establece que cuando el axón de una célula A está dispuesto lo suficientemente cercano para excitar a otra célula B y lo hace de forma reiterada se produce algún tipo de cambio en el proceso metabólico en una o en ambas células, tal que la eficiencia de A para afectar a B se incrementa. Es decir, el cambio en la fuerza sináptica no es sino consecuencia de la correlación temporal entre las actividades presinápticas y postsinápticas.

Junto a Hebb, otra influencia importante fue la del neuropsicólogo Lashley⁸. Los conceptos de acción de masa y equipotencialidad disponen de extraordinaria repercusión en el surgimiento de los modelos neuromorfos. En primer lugar, la neurona no debe entenderse como un mecanismo aislado, sino como un sistema unitario en el que complejas agrupaciones neuronales funcionan al unísono, de tal manera que una lesión cerebral va a tener mayores repercusiones cuanto mayor sea la cantidad de tejido destruido. Pero además, las áreas cerebrales son equipotenciales, es decir, exhiben la capacidad de que puede asumir funciones perdidas en otras zonas lesionadas. En este sentido, los trabajos neuropsicológicos han demostrado cómo, por ejemplo, pacientes que por lesiones cerebrales han perdido habilidades cognitivas como el habla, han sido capaces de recuperarlas porque otras zonas cerebrales intactas se han implicado en la recuperación de la función perdida. En septiembre de 1948, la Fundación Hixson organizó un Symposium sobre "Mecanismos cerebrales en la conducta", que se celebró en el Instituto Tecnológico de California. En este foro, el mismo Lashley –ante ponencias como la de von Newmann que defendía una analogía entre los ordenadores digitales y el cerebro– advirtió que dicha comparación no era apropiada, pues entendía que el cerebro

⁸ Lashley, K. (1950). In search of the engram. *Society of Experimental Biology Symposium*, 4, No. 4: Psychological mechanisms in animal behavior. Cambridge, MA: University Press.

era una máquina analógica, pero no digital, llamando también la atención de los participantes hacia las dificultades que presentaba la computación serial del ordenador, ante el funcionamiento unitario que demuestra el cerebro.

Finalmente, en el aludido Symposium de Hixson, un neurólogo llamado Warren McCulloch⁹ presentó una ponencia titulada *¿Por qué la mente está en la cabeza?*, en la que se planteaba entender el cerebro como un sistema de procesamiento. Las ideas presentadas al Symposium recogían los argumentos publicados anteriormente en colaboración con un matemático, Walter Pitts, sobre las capacidades de computación de las neuronas y su implementación en un modelo formal. A juicio de estos investigadores, la mente podía ser entendida en términos neurofisiológicos, inspirados en el funcionamiento neuronal: si la actividad neuronal es un proceso del tipo todo-nada (se activa o permanece desactivada) entonces es posible representarla formalmente utilizando una lógica binaria. Surge así un sistema de computación binario, con inspiración cerebral, que tradicionalmente se ha conocido como neurona McCulloch-Pitts, en el que una supuesta neurona artificial se activará si el conjunto de valores que ingresa supera un determinado umbral, permaneciendo desactivada en caso contrario. Pero dicha neurona puede no actuar en solitario, sino que se interconecta con otras, generando complejas RNAs que constituyen la esencia de las arquitecturas conexionistas.

Declive y resurgimiento

La idea de McCulloch y Pitts fue recogida por Frank Rosenblatt¹⁰ en 1958, quien elaboró uno de los primeros modelos de computación neuronal denominado *perceptrón*. El perceptrón era un sistema que tenía como propósito el reconocimiento de formas visuales y generó un gran impacto cuando fue propuesto. El perceptrón demostraba cómo la información recogida por el sistema visual podía almacenarse en forma de red neuronal, y cómo ésta era capaz de manipular dicha información mediante sistemas de computación semejantes a la neurona McCulloch-Pitts.

Los perceptrones generaron un furor inusitado que fue aplacado radicalmente con la publicación en

⁹ McCulloch, W.S. y Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5, 115-133.

¹⁰ Rosenblatt, F. (1958). The perceptron: A probabilistic model for information storage and organization in the brain. *Psychological Review*, 65, 386-408.

1969 del trabajo de Minsky y Papert¹¹ titulado *Perceptrons*. En este trabajo demostraron las limitaciones de estos sistemas de computación, que a veces fallaban al realizar determinadas tareas, concluyendo que el uso de RNAs, como mecanismos de computación, no era tan poderoso como se había creído.

El trabajo de Minsky y Papert marcó el inicio de una época en la que la utilización de RNAs como modelos explicativos de los sistemas inteligentes pasó a un cuarto oscuro en donde no entraba tan apenas subvención económica con la que financiar proyectos de esta clase; circunstancia que, de forma alternativa, favoreció el desarrollo de la investigación cognitiva fundamentada en el ordenador serial.

Hubo que esperar hasta inicios de la década de los ochenta del pasado siglo a que un físico llamado John Hopfield¹², en 1982, publicara un trabajo en el que retomaba las posibilidades de las RNAs como sistemas de computación. A ello unido, también en el mismo año, Feldman y Ballard¹³ rubricaban un artículo en el que por primera vez utilizaban el vocablo *conexionismo* para referirse a los modelos de RNAs como sistemas de procesamiento de la información capaces de ejecutar sofisticadas computaciones ante tareas complejas. En este trabajo, además, se analizaron las propiedades de dichas redes y las proponen explícitamente como modelos cognitivos que representan la información, no de un modo simbólico, sino distribuidamente entre las unidades que componen la red. Finalmente, en 1986 se produjo la eclosión: Rumelhart, McClelland y el denominado Grupo PDP¹⁴ publicaron *Parallel Distributed Processing: Explorations in the microstructure of cognition*. A partir de ahí, se empezó a hablar de conexionismo con un fervor inusitado; se habló de *revolución conexionista*, otros de cambio de paradigma en el que el ordenador serial había sido desmontado como modelo explicativo de la mente.

Brain-style computation

A juicio de Rumelhart¹⁵, el cognitivismo clásico ha concentrado sus esfuerzos explicativos de la mente

¹¹ Minsky, M. y Papert, S. (1969). *Perceptrons: An introduction to computational geometry*. Cambridge, MA: MIT Press.

¹² Hopfield, J. J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Sciences*, 79, 2554-2558.

¹³ Ver nota 3.

¹⁴ Ver nota 4.

¹⁵ Rumelhart, D. E. (1989). The architecture of mind: A connectionist approach. En M. I. Posner (Ed.), *Foundations of cognitive science* (pgs. 133-159). Cambridge, MA: MIT Press.

en el nivel lógico o funcional, el software, dejando de lado, a priori, la elaboración de modelos de la mente que contemplen un nivel más básico: el nivel de implementación. Es decir, el científico cognitivo clásico ha entendido la mente como un programa de ordenador y ha supuesto que dicho programa se ejecuta en cualquier máquina, sea del tipo que sea, bien con circuitos impresos como los actuales, o bien con válvulas como lo fue ENIAC. A fin de cuentas, el software es lo que interesa y no tanto el hardware de la máquina, pues se entiende que la implementación general, el nivel físico, se basa en sistemas de propósito general del tipo von Neumann.

Alternativamente, los conexionistas dirigen su mirada al *nivel de implementación*, pues a partir de él se podrán diseñar más adecuadamente algoritmos que se puedan utilizar para explicar los fenómenos mentales. Así, si se admite que la mente se implementa sobre un sustrato físico que es el cerebro, un conexionista elaborará modelos de la cognición fundamentados en una arquitectura neuronal. Por este motivo, se afirma que los conexionistas han sustituido la metáfora del ordenador por la *metáfora cerebral* de la mente, y han abandonado la computación de estilo simbólico por una computación inspirada en el funcionamiento cerebral (*brain-style computation*). Es por ello que la distinción entre un *hardware* (*la sustancia dura*) y un *software* (*la sustancia blanda*) propuesta desde el ámbito del cognitivismo clásico haya dado paso ahora al *wetware* (*la sustancia húmeda*, al cerebro como modelo).

Pero ¿por qué resulta apropiado dirigir la mirada al cerebro cuando se buscan modelos explicativos de la mente? Esta pregunta puede parecer baladí para algunos lectores; sin embargo, piensen que durante décadas los científicos parecen haber mirado antes a una tecnología como el ordenador que a un sistema orgánico como el cerebro, para explicar la mente. Algunos pueden llegar a afirmar, según sus creencias u opiniones, que sin cerebro no existe ese fenómeno o conjunto de fenómenos que denominamos mente, o que incluso, a fin de cuentas, la mente en última instancia no es sino puro funcionamiento cerebral. Frente a estas consideraciones, la pregunta no pretende tener, al menos desde el punto de vista conexionista, las concepciones ahora expuestas. La pregunta se refiere, más bien, a por qué orientar nuestra atención al cerebro para explicar la mente; o dicho de otra manera, qué aporta el cerebro para un modelo de la mente que no haya aportado el ordenador serial que manipula símbolos.

Para contestar esta pregunta, el lector debe tener presente que *el conexionismo pretende diseñar sistemas artificiales inteligentes inspirados en el cere-*

bro, pero no una teoría fisiológica del funcionamiento cerebral. Así, dirigir la mirada al cerebro para elaborar modelos generales del funcionamiento mental, implica recurrir a un tipo de estructura o arquitectura análoga al de la *sustancia húmeda* (*wetware*). De esta forma, los modelos conexionistas entienden que la unidad o elemento básico de procesamiento es la *neurona*, pero considerada ahora como un elemento computacional abstracto organizado en forma de *redes neuronales artificiales*. Por eso, se afirma que estos modelos están *neuronalmente inspirados* o que proponen una concepción *neuromorfa* de la mente. Muy sintéticamente, la aproximación conexionista al estudio mental puede sintetizarse en los siguientes puntos:

- Primero, los modelos del conexionismo son modelos neuromorfos porque explican el funcionamiento de la cognición utilizando como referente el sustrato físico que subyace en los procesos cognitivos: el cerebro. En estos modelos múltiples unidades simples de computación, a modo de neuronas artificiales, se interconectan de forma excitatoria o inhibitoria, constituyendo complejas redes de procesamiento (RNAs). Por este motivo también se denominan *modelos de la microcognición o de microcaracterísticas*.
- Segundo, la información no se procesa de forma serial –tal como postulaban los modelos simbólicos– sino que la red ejecuta un procesamiento distribuido en paralelo (PDP). Por tanto, el foco de atención no se dirige a la actividad de las neuronas artificiales contempladas aisladamente, sino al patrón de conexiones global que generan.
- Tercero, las arquitecturas conexionistas admiten representaciones; aunque ahora la información no se almacena en forma simbólica, sino como patrones de actividad distribuida entre las neuronas de la red. De ahí que a los modelos conexionistas se les denomine también modelos subsimbólicos.

¿QUÉ PODEMOS ESPERAR?

Desde mediados de los ochenta, los modelos conexionistas han tenido un fuerte impacto, no sólo en psicología sino en otras muchas disciplinas. En la ingeniería de computadores y la inteligencia artificial las RNAs suelen utilizarse como sistemas de computación en el reconocimiento artificial de patrones, lenguajes artificiales, etc. En matemáticas, estadística o economía se utilizan como modelos predictivos, alternativos a los modelos tradicionales de regresión lineal, por ejemplo, para realizar predicciones en los mercados bursátiles basadas en el análisis técnico o chartista. En neurociencia, estos modelos son entendidos como micro aproximaciones

al estudio de la computación neuronal, altamente simplificadas, que ayudan a comprender determinados aspectos del funcionamiento cerebral. Finalmente, para los psicólogos el lenguaje de las RNAs y su forma de funcionamiento suponen una nueva metáfora, alternativa al ordenador serial, para explicar el funcionamiento de la cognición humana.

Dentro del ámbito específico de la psicología, Walter Schneider afirmaba que el conexionismo suponía un cambio paradigmático, de igual forma que ocurrió con el paso del conductismo a la psicología cognitiva. Es más, afirmaba que el conexionismo sería el cambio paradigmático más grande que las jóvenes generaciones de psicólogos verían a lo largo de sus carreras. La argumentación de Schneider recurría a la noción kuhniana de cambio paradigmático, aplicada ahora al conexionismo. Ciertamente, desde la década de los cincuenta, el paradigma cognitivo de la psicología instauró un periodo de *ciencia normal* en el que proliferaron numerosas teorías inspiradas en el ordenador simbólico serial. Paralelamente, como es habitual en ciencia, surgían anomalías difíciles de interpretar desde las teorías vigentes, lo que ocasionó en numerosas ocasiones la elaboración de modelos *ad hoc*. Sí que es cierto que el paradigma cognitivo clásico adolece de una enorme dispersión conceptual y, probablemente, en un futuro no muy lejano, su tarea de elaboración de supuestos teóricos pueda agotarse, tal como indicaba Schneider. No es menos cierto que, desde mediados de los ochenta, el conexionismo marcó un periodo de *ciencia revolucionaria*, que incorporó un nuevo vocabulario y un renovado tratamiento de los fenómenos cognitivos.

Personalmente, creo que el interés conexionista de mediados de los ochenta y principio de los noventa no ha constituido un cambio paradigmático tan *revolucionario*, como fue el paso de la psicología de corte conductista a la psicología cognitiva del procesamiento de la información. Ahora, dos décadas después de esta supuesta revolución conexionista en psicología, los modelos explicativos de la cognición, fundamentados en arquitecturas de RNAs, conviven en el seno de la psicología y ciencia cognitiva con las propuestas clásicas del ordenador serial que manipula símbolos. Si el funcionamiento de un determinado proceso mental sigue unas reglas bien establecidas es posible explicarlo recurriendo a un modelo simbólico serial, pues la buena definición ayuda a descomponer el procesamiento y a la elaboración, si se desea, de un algoritmo que permita la simulación. El problema es que, la cognición humana, no está tan rígidamente determinada como para afirmar que sobre la mente actúan una serie de leyes o reglas mandatorias que se aplican siempre y sin excepción, contrariamente a lo que parecen sostener

Fodor y Pylyshyn¹⁶. El ser humano debe tomar decisiones en situaciones de información parcial, debe manejar la incertidumbre del medio, experimenta estados emocionales que afectan a su forma de percibir el mundo y, en circunstancias de este tipo, donde las reglas que rigen la cognición pueden ser vagas o intuitivas, es cuando los sistemas conexionistas demuestran su supremacía.

Nadie, entre los científicos cognitivos actuales, pone en duda que el estudio neurofisiológico del cerebro ayudará a descubrir numerosos aspectos de la naturaleza de la mente. Es más, en la actualidad es exigible que una teoría cognitiva no sólo ofrezca datos provenientes de un laboratorio experimental, sino también validación de los mismos mediante técnicas de neuroimagen cerebral, estudio de funciones alteradas en cuadros de lesiones cerebrales u observaciones clínicas. Sin embargo, explicar el complejo funcionamiento de la actividad cognitiva y sus aspectos de contenido, será mucho más dificultoso y la moderna neurociencia cognitiva, como aproximación multidisciplinaria, deberá realizar un gran esfuerzo en las próximas décadas para generar aportaciones decisivas e innovadoras. Evidentemente, toda experiencia cognitiva tiene un sustrato neurofisiológico; pero esta afirmación no contradice en absoluto que la materia cerebral se organice de forma emergente, y las propiedades que rigen los fenómenos cognitivos humanos sean diferentes de las que determinan el funcionamiento biológico cerebral. El neurólogo portugués, Antonio R. Damasio¹⁷, recientemente incorporado como Director del *Brain and Creativity Institute* en la *University of Southern California*, afirmaba lo siguiente:

Creo que podemos arriesgarnos a decir que para el año 2050 tendremos suficiente conocimiento de los fenómenos biológicos para suprimir el dualismo tradicional entre cuerpo y cerebro, cuerpo y mente, cerebro y mente. Algunos podrían temer que con la determinación de la estructura física de algo tan valioso y digno como la mente humana, ésta quede degradada, si no desechada. Pero, explicar los orígenes y operaciones de la mente en el tejido biológico no comportará su eliminación (A. R. Damasio, Op. Cit., pg. 71).

Tal vez, en el 2050 tengamos una contestación clara al atractivo, a la par que espinoso, asunto de la cognición humana. Personalmente, intuyo que no será simple. Sin embargo, seguro que siempre guardaremos un pequeño rincón para nuestra amada mente y ella, como agradecimiento, nos seguirá permitiendo percibir el mundo de forma significativa, razonar, comprender y comunicarnos con los demás, atender a lo que nos interesa, aprender conscientemente y, tras décadas, recordar de forma increíblemente exacta y con multitud de detalles el día más feliz de nuestra vida.

A. Crespo, PhD
Enero 2007

¹⁶ Fodor, J. y Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture: A critical analysis. *Cognition*, 28, 3-71.

¹⁷ Damasio, A. R. (2000). Creación cerebral de la mente. *Investigación y Ciencia*, 280, 66-71.