

## REDES NEURONALES Y CONEXIONISMO EN LAS NEUROCIENCIAS

di Juan Canseco

El análisis de las relaciones entre el aprendizaje (y otros procesos cognitivos) y sus bases celulares y químicas constituye una cuestión capital para la filosofía de la ciencia contemporánea. La neurociencia más reciente ha emprendido pasos decisivos al momento de identificar correspondencias rigurosas entre la modificación de las conductas (aprendizaje, adquisición-retención de conocimientos), por un lado, y cambios químicos a nivel sináptico, así como cambios morfológicos y funcionales tanto a nivel celular como de redes neuronales. Las neurociencias, en nuestros días, están apuntalando una descripción eminentemente plástica del sistema nervioso. Si el aprendizaje ha sido concebido tradicionalmente como una manifestación de la plasticidad o maleabilidad de la conducta de muchos animales, hoy ya se conoce que a nivel biológico y químico esta propiedad también está presente de forma ya indiscutida (p.e. Hebb, 1949; Kandel y Hawkins, 1983). Recientemente, una aproximación teórica y metodológica conocida bajo el nombre de “conexionismo”, ha propuesto modelos matemáticos de conectividad neuronal, entre los que se incluyen ciertas técnicas y algoritmos que intentan aproximarse al cambio efectivo que tiene lugar en las conexiones entre las neuronas (Rumelhart et al., 1992).

P.S. Churchland y T. Sejnowski (1990) sostienen que en los enfoques clásicos (o simbólicos) del estudio de la cognición se ha simplificado excesivamente el nivel de implementación físico (v.gr. neurobiológico), como consecuencia del enfoque trimembre del cerebro que ha privilegiado la metáfora informática. Los tres niveles contemplados clásicamente son: contenido, estructura e implementación. Pero la neurociencia no constituye un “nivel” homogéneo de materia indiferenciada sobre la que se asientan símbolos y pensamientos. Existen de hecho muchos niveles en la propia neurociencia (op. cit. p. 356). El sistema nervioso es “plástico”, no rígido. Esto significa que se le puede

considerar como un sistema dinámico (se transforma o evoluciona a lo largo del tiempo), y en modo alguno puede considerarse organizado en niveles casi autónomos entre sí, niveles como los propuestos por Marr (1985) en su teoría computacional de la visión. El conexionismo y la neurocomputación dinámica son tendencias que abogan por interacciones múltiples y cambiantes entre todos los niveles. Hay más de un nivel de instanciación. Las neurociencias estudian desde el nivel molecular hasta los niveles sucesivamente más “elevados” de las neuronas y las sinapsis, los circuitos, los sistemas, etc. Además del “sueño de Marr”, acerca de niveles de cómputo, Churchland y Sejnowski proponen “despertarse” de otros dos sueños, el de Boole (el cerebro como máquina lógica que maneja oraciones expresadas simbólicamente), así como del sueño de Ramón y Cajal (el cerebro descrito prioritariamente al nivel de fisiología de la neurona). Además, no hay un programa escrito con las reglas de actuación del sistema sino que, en cierta forma, el sistema dinámico “se programa él mismo” (Churchland y Sejnowsky, op. cit. p. 361), en virtud de los cambios en los “pesos” asignados para el procesamiento. En este sentido se dice que el sistema “aprende” recomponiéndose automáticamente al comparar el patrón deseado con el patrón obtenido en cada fase de cálculos, corrección que se puede lograr por medio de diversas técnicas de modificación de conexiones. Nótese que en los sistemas conexionistas esta palabra de “aprendizaje” se emplea en un sentido de acomodación entre valores numéricos actuales y valores deseados, y no está implicada una teleología por parte de la red. Otro rasgo de la neurocomputación es que no hay representaciones “puntuadas”, ni localizables, sino que la información está distribuida en las propias conexiones. No hay representación en almacenes ni en puntos particulares. El procesado ocurre en paralelo, no secuencialmente, lo que implica la multi-interacción entre niveles. El cerebro funciona de un modo más bien geométrico, que no sintáctico. Una conducta compleja y organizada, como el lenguaje, se debe a cálculos que por sí mismos no exhiben una forma sintáctica.

La unidad de procesamiento “conexionista” es la neurona abstracta. Las operaciones de sus modelos se caracterizan como “inspiradas neuronalmente”. La arquitectura de computación no es de tipo von Neumann, sino que se pretende una aproximación a la arquitectura biológica del cerebro. El grupo PDP, encabezado por Rumelhart y

McClelland, oponen un nivel subsimbólico al nivel superior o simbólico (conceptual). Las huellas de memoria, p. e., se conciben “dinámicamente” como cambios en los pesos de conexión. Escriben Rumelhart y McClelland (1992, p. 304): “Cada huella de memoria está distribuida en muchas conexiones diferentes y cada conexión interviene en muchas huellas de conexión distintas”. La recuperación, asimismo, se concibe como una reinstauración de un patrón de activación anterior, el cual contenía previamente un fragmento que luego servirá para ulteriores recuperaciones. Las unidades están conectadas entre sí según ciertas fuerzas o pesos. Hay toda una serie de “unidades ocultas” que soportan la tarea de ocuparse de las representaciones internas.

Con estos modelos PDP, no sólo se guarda la equiparación “funcional” entre la mente y el ordenador, sino que se da un paso más, y en una cierta dirección. La dirección consiste en reducir lo psicológico a lo que hay dentro de lo cerebral en general (leyes numéricas de conectividad, funciones entrópicas, propiedades estadísticas de la máquina de Boltzman, de Markov, etc.), y debemos subrayar su carácter “genérico”, porque estas leyes estadísticas o numéricas son comunes a todos los sistemas dinámicos compuestos a base de un número elevado de componentes en interacción. Según la caracterización que da Smolensky (1988), el conexionismo “tratado apropiadamente” se puede comparar, (1) en el plano “metodológico” con (1a) los modelos de computación simbólica, ante los que se alza como una alternativa, y (1b) con las metodologías empíricas (experimentos de tiempos de reacción, cómputo de errores o prácticas de fisiología) frente a los que se distingue el conexionismo -en realidad es estrategia de simulación- si bien podemos decir que los límites entre simulación y experimento no siempre están claros. Por otro lado, Smolensky distingue otros dos rasgos que, antes que metodológicos, podríamos considerar “objetivos”, esto es, que tienen que ver con la realidad o realidades a estudiar. Por un lado, (2a) tendríamos modelizaciones conexionistas de la computación neuronal, y por otro (2b) el estudio de la competencia humana. Es decir que, en general, el conexionismo puede entenderse como una investigación que se enfrenta a los dos niveles ontológicos, el neuronal y el conductual, si bien, Smolensky precisa que esta estrategia sigue más de cerca el segundo, esto es, se trata de una teoría/metodología más propiamente cognitiva que neurocientífica: el conexionismo es una derivación interna de la

ciencia cognitiva, no una reducción por parte de la neurociencia. Se describe como un paralelismo con respecto de los niveles reductivos de la neurociencia, y una granulación más fina de los niveles conceptuales (cognición), que se entienden ahora como subconceptuales. Bajo esta interpretación, las redes conexionistas hasta ahora desarrolladas se hallan algo por encima del nivel neuronal, subyacen a lo conceptual y no se identifican (aún) con lo neurobiológico.

Los conexionistas consideran que no es momento de abrir brechas con la tradición precedente, y por ello entienden (metafísicamente) la relación entre la mente y una red de neuronas en términos de emergencia. Sin embargo, otros autores, (Ramsey et al., 1990) sostienen que el conexionismo está ligado a una filosofía eliminacionista, y que tales simulaciones resultan incompatibles con los constructos mentalistas y los términos de la psicología popular. A diferencia del cognitivismo clásico, que no era incompatible, dicen, con una teoría popular en términos de actitudes, creencias, deseos, etc., el conexionismo moviliza una nueva ontología no-conservacionista, formada por entidades nuevas que destruyen a (o son incompatibles con) las entidades de la teoría que se ha dejado atrás (como fue el caso de las teorías físico-químicas que destruyeron entidades supuestas como el éter o el flogisto). Parece que esta identificación entre conexionismo y eliminacionismo, viene unida a la crítica basada en la irrelevancia de las redes neuronales para la psicología-popular. Un mismo estado mental (una creencia, un deseo) puede ser representado por un número indefinidamente elevado de redes conexionistas, lo cual será como decir que son irrelevantes para asuntos de “alto nivel” cognitivo. Se viene a decir que las redes neuronales son meras simulaciones de los “substratos” y las causas de la vida mental, pero ninguna de ellas logra construir clases naturales de esos estados mentales de alto nivel, lo cual es como sostener que son irrelevantes para la cognición, a diferencia de las redes clásicas, p.e. de tipo proposicional. Esta identificación del conexionismo con una filosofía eliminacionista no es compartida por Smolensky, autor que ha procurado explicitar lo que él denomina un adecuado tratamiento del conexionismo. Según Smolensky, los modelos conexionistas sirven para refinar los modelos simbólicos, eliminando las distancias que los separan por medio de una “traducción” de los sistemas de símbolos discretos en computaciones continuas. Es decir, se aboga por un programa

de reconstrucción de las entidades precedentes (simbólicas) por medio de una dinámica subsimbólica. Por otro lado, Smolensky nos recuerda que importantes algoritmos incorporados a los sistemas conexionistas, como p.e. el algoritmo de retropropagación y la máquina de Boltzmann, se han desarrollado por medio de consideraciones estrictamente matemáticas y no, en cambio, de tipo neuronal. Hay que tener en cuenta que nuestro conocimiento acumulado sobre el sistema nervioso es, principalmente, de tipo estructural (v.gr. anatómico) pero no tenemos un conocimiento igualmente profundo de su dinámica.

Los defensores del paradigma conexionista advierten acerca de las diferencias existentes entre el cerebro real, y las simulaciones “inspiradas neuronalmente”. La morfología de las neuronas es muy compleja. Las neuronas no sólo son unidades que reciben o envían un valor numérico. Las neuronas reales son mucho más lentas, pero en cambio hay un número enorme de ellas. Las neuronas reciben entradas desde un elevadísimo número de sinapsis. La base del conocimiento está en esa masa cambiante de conexiones, y no en las unidades mismas. Estas unidades, en la simulaciones, no son sino unidades arbitrarias o “neuronas formales”. Por otro lado, hay varias clases de señal neuronal, mientras que las señales entre unidades conexionistas son de tipo único.

La idea de que los límites del lenguaje son los límites del sistema cognitivo, es la que usan Fodor y Pylyshyn (1988) para emprender su crítica contra las arquitecturas conexionistas. La mente real está sometida a las mismas restricciones que un lenguaje real, desde el punto de vista del paradigma simbólico. A saber: (1) Productividad: como ya señalara Chomsky, no es posible suponer que una mente almacene todos los pares que definen una función de correspondencia entre representaciones y las proposiciones que expresan esas proposiciones. Se supone que la mente posee una capacidad finita de almacenamiento y, sin embargo, su facultad generadora de oraciones es infinita. Los sistemas conexionistas se enfrentan con este problema de la infinitud. Para aproximarse a la enorme productividad real de la mente, necesitarían de una memoria expandible *ad infinitum*, la cual no es necesaria en las arquitecturas clásicas. (2) Sistemática: puede entenderse como la conectividad que unos enunciados (“mentales”) tienen con respecto a

otros. No obstante, los críticos del conexionismo, Fodor y Pylyshyn deben argumentar que el estudio etológico de organismos infraverbales deja abierta la cuestión de la sistematicidad cognitiva en otras especies, como una cuestión empírica y no sujeta a argumentaciones *a priori*. Por último, (3) el principio de la compositividad de las representaciones: en la medida en que un lenguaje es sistemático, se asume que un elemento léxico debe realizar (aproximadamente) la misma contribución semántica a cada expresión en la que ocurre este elemento. Este principio supone un amplio margen para una semántica independiente del contexto. Pero Smolensky ha replicado que sólo es legítimo al hablar de una compositividad “aproximada”. Los distintos niveles influyen recíprocamente y de distintas maneras. Así, el nivel sintáctico se vería influido por la específica implementación que le subyace, pero también por niveles semánticos que están actuando sobre las propiedades estructurales. El enfoque sintáctico de la mente se basa en una lógica-matemática discreta, que consta de símbolos generados por reglas de formación y transformación. Smolensky, por su parte, contrapone unas matemáticas “continuas”, de tipo estadístico y vectorial. Estas otras técnicas de formalización contrastan con las versiones atomísticas de la semántica y sintaxis clásicas. Los sistemas conexionistas se presentan, pues, como modelos altamente interaccionistas, lo cual contrasta con el enfoque modular. Según Fodor y Pylyshyn los sistemas conexionistas no iluminan, ni siquiera mínimamente, la estructura (sintaxis y semántica) de la cognición, y sólo simulan cómo podría implementarse a nivel neurológico y causal, pero no de una forma esencial, sino de una forma irrelevante (habría indeterminación de múltiples implementaciones para cada estructura).

Los modelos conexionistas, a nuestro juicio, poseen la ventaja de salir al paso de los toscos modelos de cajas y flechas de flujo. Asimismo suponen una verdadera destrucción de la metáfora del ordenador, no por medio de un abandono o negación (por una mera posición actitudinal o ideológica de los investigadores), sino precisamente por ampliación, por profundización de las nuevas posibilidades derivadas del desarrollo de la ciencias y las tecnologías del cómputo. De igual forma, las célebres jerarquías de niveles (físico---> simbólico--->semántico--->pragmático) pueden destruirse en virtud de la múltiple virtualidad de interconexión de unos mismos elementos u objetos básicos (términos

fisicalistas), que entran en relaciones de distinta clase por referencia precisamente a la diversa operatoriedad del agente. Así, en el conexionismo, los niveles jerarquizados se desvanecen, mas no por medio de una mezclanza de estos subsistemas o módulos en un sólo sistema indiferenciado u holista (mente, cerebro), sino precisamente por medio de las múltiples escalas o planos de intercalación, idea ésta que es una de las más fértiles de entre las aportadas por el conexionismo.

Muchos otros modelos o metáforas han sido propuestos en la historia de las neurociencias. Estaciones telefónicas, tendidos eléctricos, organigramas empresariales, topologías y grafos, hologramas, fractales, etc. En general, una cosa nos parece cierta: el cerebro es un sistema que requiere, como cualquier otro, una aproximación científica en términos organizativos. Además de unas estructuras anatómicas, se hace preciso apelar a funciones (fisiológicas, conductuales). Las correspondencias entre éstas y aquéllas no son simples. Es más, el mismo término “correspondencia” nos parece completamente inservible con vistas a una reconstrucción gnoseológica de las neurociencias. Las formas -o pautas organizativas de la materia- no son otra cosa que el funcionamiento o actividad de la materia misma implicada (moléculas, células, tejidos, cuerpos, etc.). Debemos precisar más aún: ninguna ciencia agota un objeto o una jurisdicción de objetos. Hay unos límites (difusos y empíricamente relativos) en los que el estudio del tejido nervioso es y debe ser el estudio de la conducta, y recíprocamente, hay momentos prácticos y empíricos en los cuales la investigación conductual (y lingüística) involucra a los tejidos, moléculas, etc. Asimismo, las descripciones no tienen la menor utilidad ni guardan la menor relación con las demarcaciones entre ciencias, y en ese sentido se debe rechazar de lleno esa epistemología ingenua de “abrir la cabeza y observar qué acontece dentro”, como forma de concebir un programa de reduccionismo neuronal aplicado a la psicología. Muchos filósofos y científicos se aferran a esta epistemología ingenua, a este materialismo vulgar, cuando despachan todo un campo empírico de investigación por medio de alusiones misteriosas y expeditivas como estas: “hechos físicos del cerebro”, “eventos neuronales”, “descripción reduccionista”, “causalidad física en ultima instancia”, etc. En la filosofía de la mente y de la psicología, diversas tendencias que se autodefinen como materialistas y funcionalistas asumen que el problema filosófico por excelencia



reside en la búsqueda de algún esquema de conexión fiable entre esos eventos materiales (físicos) y espirituales (cognitivos). De forma enteriza, y con una ignorancia imperdonable hacia la práctica científica implicada en los problemas (la práctica de los neurocientíficos, pero también la de psicólogos, etólogos, lingüistas), los esquemas conectivos toman la formas de emergentismo, reduccionismo, epifenomenalismo, y entre otros más, el esquema reciente de la mente “sobrevenida” (*supervenience*) a partir de hechos físicos. En toda esta literatura se plantean preguntas acerca de si es posible una mente habitando un sistema físico, si una mente puede tener poderes causales sobre la conducta o sobre otros sistemas físicos. Incluso hay marcadas tendencias hacia la ciencia-ficción: baste recordar, por ejemplo, los debates acerca de si un ordenador puede llegar a pensar, o en torno al caso de un marciano con substrato neuroquímico completamente diferente al nuestro pero dotado de una capacidad cognitiva equivalente a la humana (nuevas versiones, éstas, del bizantinismo especulativo). Tales desvaríos -con respecto a la investigación empírica rigurosa- sólo sirven a los efectos de legitimar ideológicamente el mentalismo, el espiritualismo y la creación *ex nihilo* de unos pensamientos que vuelven a adornar el mundo grosero de la materia. Precisamente, la persistencia del fisicalismo, como doctrina filosófica en las actuales “teorías materialistas de la mente” ha servido para remozar viejísimas concepciones metafísicas de la materia. La materia es entendida de nuevo como un homogéneo sustrato brutal y desorganizado (que obedece solamente a las “leyes físicas”) carente de formas y de pensamientos. Este materialismo grosero es el mejor aliado y la más notoria consecuencia del moderno espiritualismo. Esta vía de degradación de la materia (materia “meramente física”, como a veces se dice, como si con ello pudiera implicarse que la materia adoleciese de algo), es la mejor propaganda para una vergonzante rehabilitación de la idea de espíritu.



## Bigliografía

- Churchland, P.S. y Sejnowsky, T. (1990): "Neural representation and neural computation", pps. 343-382, en Tomberlin, J. (ed.), *Philosophical perspectives*, 4.
- Fodor, J.A. y Pylyshyn, Z. (1988): "Connectionism and cognitive architecture: a critical analysis", *Cognition*, 28, 3-71.
- Hebb, D. (1949): *The organization of behavior*, Wiley: Nueva York.
- Kandel, E. y Hawkins, R. (1993): "Bases biológicas del aprendizaje y de la individualidad", pps. 54-63 en V.V.A.A., *Mente y cerebro*, Prensa Científica:Barcelona.
- Marr, D. (1985): *Visión*, Alianza: Madrid.
- Ramsey et al (1990): "Connectionism, eliminativism, and the future of folk psychology", pps. 499-533 en Tomberlin, J. (ed.), *Philosophical perspectives*, 4; *Action Theory and Philosophy of mind*, Ridgeviews: Atascadero.
- Rumelhart et al. (1992): *Introducción al procesamiento distribuido en paralelo*, Alianza: Madrid.
- Smolensky, P. (1988): "On the proper treatment of connectionism", *Behavioral and Brain Sciences*, 11, 1-23.



Sesto San Giovanni (MI)  
via Monfalcone, 17/19

© Metábasis.it, rivista semestrale di filosofia e comunicazione.  
Autorizzazione del Tribunale di Varese n. 893 del 23/02/2006.  
ISSN 1828-1567



Esta obra está publicada bajo una Atribución-No Comercial-Sin Derivadas 2.5 Italia de Creative Commons. Para ver una copia de esta licencia, visite <http://creativecommons.org/licenses/by-nc-nd/2.5/it/>.