

EDITORIAL

REDES NEURALES EN ANÁLISIS DE LA CONDUCTA: MODELOS, RESULTADOS Y PROBLEMAS

JOSÉ E. BURGOS
UNIVERSIDAD DE GUADALAJARA

El 25 de julio de 1989, el senado y la Casa de Representantes de los Estados Unidos de Norteamérica aprobó una resolución en la que se designó oficialmente a la década de los 90 como la "Década del Cerebro". El propósito principal de esta resolución fue estimular investigación neurocientífica, tanto básica como aplicada, que eventualmente resultara en tratamientos para los diversos desórdenes neurológicos conocidos hasta el momento. Aún queda por determinar hasta qué punto tal objetivo fue satisfecho. Pero la resolución ciertamente estimuló una cantidad sin precedentes de investigación sobre la estructura y función del cerebro. Este tipo de investigación, por supuesto, había estado ocurriendo mucho antes de la resolución en cuestión. Sin embargo, el número de técnicas, métodos, revistas especializadas y publicaciones en las neurociencias se incrementó dramáticamente durante los últimos 10 años. La base de datos resultante es colosal, por decir lo mínimo, y continúa creciendo a cada minuto, no sólo en los Estados Unidos de Norteamérica, sino también alrededor del mundo. El espíritu de la resolución, pues, ha trascendido sus fronteras culturales y políticas inmediatas para convertirse en un esfuerzo humano colectivo hacia un entendimiento del cerebro.

La gran mayoría de los datos neurocientíficos han sido obtenidos mediante la aplicación de una amplia variedad de métodos, técnicas y preparaciones de laboratorio relativamente simples, utilizando cerebros no humanos, tanto de vertebrados como de invertebrados, abarcando desde el nivel molecular hasta el nivel anatómico de organización. Una lección que hemos aprendido de esta investigación es que los cerebros son extraordinariamente complejos, mucho más de lo que cualquiera esperaba. Esta lección se ha hecho tan patente que podría hacernos creer que los cerebros son inherentemente complejos y que tal complejidad es una propiedad fundamental

de los cerebros. No obstante, algunos esfuerzos han sido guiados por un rechazo de esta idea, es decir, por la noción de que tal complejidad puede ser vista como emergiendo de múltiples realizaciones de unos pocos principios generales que son lo suficientemente simples como para permitir la construcción de *modelos matemáticos*. Aunque estos esfuerzos constituyen una minoría dentro del estudio colectivo del cerebro, han sido lo suficientemente sustanciales como para ser notados.

Los modelos matemáticos desarrollados son conocidos por varios nombres, como 'neurocomputación', 'neurociencia computacional', 'conexionismo', y 'modelamiento con redes neurales'. Al igual que la mayoría de las etiquetas, éstas admiten diferentes usos, de tal manera que no hay un acuerdo universal acerca de sus significados. En ciertos usos, algunas de ellas reflejan gruesamente una diversidad de esfuerzos que pueden ser clasificados de acuerdo con un número de criterios diferentes. Bajo lo que parece ser el criterio más utilizado, tales esfuerzos difieren en cuanto intentan reflejar o capturar datos experimentales duros acerca de partes específicas de sistemas nerviosos particulares. El término 'neurociencia computacional' ha sido algunas veces utilizado para nombrar este tipo de esfuerzo, lo cual sugiere que puede ser visto como un subcampo dentro de las neurociencias. Otros esfuerzos están más flojamente guiados (en diferentes grados) por esos datos, y usualmente provienen de la ciencia de la computación/información y de la psicología. Pero, de nuevo, las fronteras disciplinares entre los distintos tipos de esfuerzos son más bien difusas.

El interés científico por el cerebro, pues, ha trascendido las fronteras disciplinares de las neurociencias. No obstante, ello ha relajado los criterios para evaluar modelos de redes neurales. El trabajo de modelamiento fuera de las neurociencias es usualmente considerado como basado sobre una postura filosófica denominada 'conexionismo'. De acuerdo con esta posición, la 'inteligencia', la 'cognición', y la 'mente' (como sea que elijamos definir estos términos conductualmente) deben ser entendidos como propiedades *emergentes* (más que fundamentales) del cerebro. Después de todo, cualquier fenómeno conductual referido por esos términos debe relacionarse de manera crítica a la estructura y función de cerebros, aún cuando la naturaleza exacta de esta relación permanece evasiva.

Sin embargo los conexionistas insisten que no necesitamos capturar los detalles biológicos (moleculares, celulares, anatómicos) de los cerebros para entender aquellos fenómenos. Sólo necesitamos concentrarnos sobre las características definitorias más generales de los cerebros, a saber, representaciones no simbólicas distribuidas y procesamiento en paralelo. Estas características contrastan con el tipo de representaciones centralizadas/simbólicas y procesamiento paralelo de las máquinas de Turing (como las

caracterizaciones teóricas subyacentes a las computadoras personales actuales). Los conexionistas, pues, tienden a rechazar la metáfora de la computadora Turing/von Neumann acerca del cerebro.

Los neurocientíficos computacionales, por supuesto, estarán de acuerdo en que la inteligencia, la cognición y las mentes deben ser entendidas como propiedades emergentes, y que la representación no simbólica/distribuida y el procesamiento en paralelo son, en efecto, características del cerebro. También rechazarán la metáfora de la computadora Turing/von Neumann acerca del cerebro. Sin embargo, los neurocientíficos computacionales sostienen que aquéllas características son demasiado generales para ser relacionadas de maneras específicas a cerebros específicos, puesto que son compartidas virtualmente por todos los cerebros. Los neurocientíficos computacionales arguyen que necesitamos modelar los detalles biológicos de la estructura y funcionamiento de cerebros reales. La construcción de modelos de redes neurales, pues, debe estar apretadamente restringida por la evidencia experimental dura sobre cerebros reales, en los niveles molecular, celular, y anatómicos. Los conexionistas están en desacuerdo, arguyendo que dichos modelos pueden (y deben) ser validados conductualmente, es decir, sobre la base de cuán bien simulan ciertas ejecuciones de interés.

La distinción anterior entre neurocientíficos computacionales y conexionistas, por supuesto, es sumamente idealizada y tajante. En la realidad, estas categorías representan extremos de un continuo. Ciertamente, tales extremos no son vacíos. Sin embargo, esfuerzos ubicados entre esos extremos son también muy reales. De cualquier modo, la controversia acerca de cómo deben ser evaluados los modelos de redes neurales permanece sin resolver. No obstante, existe un acuerdo general sobre la idea de que la complejidad estructural y funcional de cerebros reales es el resultado de muchas realizaciones particulares de un mismo conjunto de principios básicos. Cualquier función compleja que deseemos atribuirle al cerebro (llamémosla 'inteligencia', 'cognición', o 'conducta') debe ser considerada como emergiendo de colecciones de elementos que son estructural y funcionalmente mucho más simples que los cerebros mismos que tales elementos constituyen. Esta idea crucial ha estado en la base del modelamiento con redes neurales desde sus inicios, en los trabajos seminales de McCulloch y Pitts (1943) y de Rashevsky (1938). Es frecuentemente expresada diciendo que dicho modelamiento sigue una estrategia *fondo-arriba*. De acuerdo a esta estrategia, simulaciones de cualquier fenómeno conductual de interés debe emerger de sistemas cuyos elementos constituyentes funcionen como neuronas. De nuevo, cuán cercanamente deberían tales componentes emular neuronas reales, o cuán cercanamente deberían tales sistemas emular la anatomía de circuitos neuronales reales está todavía (y, sospecho, permanecerá por un largo tiempo)

en discusión. Pero la noción de una metodología fondo-arriba representa un tema unificador que corre a lo largo de los esfuerzos de modelamiento con redes neurales, apartándolos de otros esfuerzos para entender la conducta, como aquéllos provenientes de la inteligencia artificial no conexionista y de la psicología cognoscitiva tradicional.

La "Década del Cerebro" ya finalizó. Sin embargo, originó un ímpetu de investigación que muy posiblemente permanecerá por muchos años venideros, por lo menos durante la siguiente década. El modelamiento con redes neurales (tanto dentro como fuera de las neurociencias) comparte este ímpetu. Todavía permanece como un esfuerzo minoritario en las neurociencias y en la psicología, aunque se está haciendo cada vez más popular en la ciencia de la computación/información. Actualmente, con el inicio de una nueva década, este tipo de esfuerzo es reconocido como un campo de investigación firmemente establecido y legítimo.

En 1998, la Asociación Psicológica Americana propuso designar esta nueva década como "La Década de la Conducta", una propuesta cuya resolución, hasta donde sé, aún espera aprobación oficial. Si llega a ser aprobada, se espera que estimulará una gran cantidad de investigación en la ciencia de la conducta, al menos tanta como la que fue estimulada en las neurociencias por la resolución de "La Década del Cerebro". Aquella propuesta, en combinación con el ímpetu dejado en el modelamiento con redes neurales por esta resolución, nos proveen de un *zeitgeist* ideal para una incorporación más sistemática y completa de principios conductuales en el modelamiento con redes neurales. Después de todo, la conducta juega un papel central en la validación fenoménica de estos modelos, para neurocientíficos computacionales y conexionistas por igual. De nuevo, los investigadores permanecen en desacuerdo acerca de hasta qué punto la validación conductual es suficiente. Sin embargo, creo que la mayoría estaría de acuerdo en que por lo menos es necesaria.

Reconocer la validación conductual como un criterio necesario para evaluar modelos de redes neurales plantea el problema de exactamente qué tipos de fenómenos conductuales deberían tomarse como referentes empíricos para aplicar tal criterio. No sorprenderá que los investigadores también difieran en este punto. El rango de fenómenos es tan vasto (desde la conducta más simple en invertebrados hasta la más compleja en humanos) que es futil intentar identificar (mucho menos imponer) un objeto único de estudio, o dedicar una cantidad significativa de tiempo al estudio de una porción significativa de ese rango. Es de esperarse que dicho rango se hará aún más vasto, especialmente si esta nueva década es oficialmente designada como la "Década de la Conducta".

Lo mejor que podemos hacer, pues, es ser (muy) selectivos. Son posibles varias pautas para seleccionar fenómenos de interés. De acuerdo con una pauta, deberíamos utilizar como criterios de validación *principios* más que fenómenos. Los principios constituyen pautas de selección poderosas, ya que usualmente se refieren a unos pocos fenómenos paradigmáticos. También, se supone que un principio es válido para la mayoría de (sino todas) aquellas instancias de esos fenómenos paradigmáticos que no hayan sido directamente estudiadas en el laboratorio. Sobre esta base, podemos concentrarnos en el tipo de fenómenos que son estudiados en el análisis experimental de la conducta (ampliamente concebido, para incluir la investigación tanto en condicionamiento operante, como Pavloviano), como disciplina cuyo objetivo principal ha sido y aún es, precisamente, derivar principios conductuales en el laboratorio.

En vista de todo lo anterior, no es coincidencia que este Número Especial del Año 2000 de la *Revista Mexicana de Análisis de la Conducta* sea acerca del modelamiento con redes neurales. Una característica que separa estos esfuerzos de otros, sin embargo, es que están siendo hechos por investigadores que han sido entrenados y trabajan en la tradición analítico-conductual. Esta característica es extraordinaria, si consideramos que los analistas conductuales tradicionalmente se han concentrado sobre el estudio de la conducta como objeto de estudio por derecho propio, rechazando cualquier descripción o explicación que apele a niveles de análisis no conductuales. No obstante, la investigación en redes neurales dentro del análisis conductual se ha convertido en una realidad innegable que puede muy bien cambiar la geografía conceptual, teórica y metodológica de nuestra disciplina. El presente Número es testimonio de ello.

El Número contiene siete contribuciones. En la primera contribución, Potter y Wilson nos ofrecen una introducción al modelamiento con redes neurales en el análisis conductual, mediante una revisión de algunos modelos. También discuten maneras específicas en las cuales el análisis conductual y el modelamiento con redes neurales pueden beneficiarse mutuamente. En la segunda contribución, ofrezco una interpretación de la superstición, basada sobre simulaciones de este fenómeno utilizando redes neurales de selección. También, discuto algunos problemas filosóficos que subyacen a criticismos tradicionales hechos al modelamiento con redes neurales en el análisis conductual. En la tercera contribución, Delepouille, Preux, y Darcheville describen simulaciones de cooperación en una situación social mínima por parte de agentes artificiales cuyo funcionamiento es descrito por diferentes modelos de aprendizaje por reforzamiento. En la cuarta contribución, Jozefowicz, Darcheville, y Preux describen trabajo relacionado sobre una aproximación operante al dilema iterado del prisionero. Sus simulaciones muestran la

emergencia de cooperación luego de un reforzamiento indirecto de conducta no cooperativa en agentes de aprendizaje artificiales. En la quinta contribución, Kemp y Eckerman, guiados por la investigación acerca de los efectos de la dopamina sobre células piramidales hipocampales, presenta datos de simulación sobre la formación de patrones conductuales en redes neurales que no reciben estimulación exteroceptiva antecedente. En la sexta contribución, Matt Morris describe simulaciones de varios fenómenos de condicionamiento operante, tales como adquisición, extinción, readquisición, y ejecución de razón e intervalo variable, utilizando su modelo "Artie". Finalmente, en la séptima contribución, Jackson Marr nos provee de una serie de reflexiones críticas sobre la naturaleza de los modelos de redes neurales, qué pueden ofrecer al análisis conductual y bajo qué condiciones.

Quiero expresarle a Carlos Bruner mi gratitud por haberme dado esta oportunidad única como editor invitado para este Número y a Laura Acuña por su invaluable ayuda durante el proceso de armarlo. También quiero extender mi más profunda apreciación a los otros seis autores por haber aceptado mi invitación a participar en este esfuerzo colectivo, y por haber hecho tan excelente trabajo en sus respectivas contribuciones. Su paciencia y consideración ciertamente hicieron mi trabajo como editor un verdadero placer. Aún queda mucho trabajo por hacer. Espero que este Número sirva para sensibilizar a los analistas conductuales a las posibilidades que los modelos de redes neurales nos ofrecen para un entendimiento científico más preciso y comprensivo de la conducta. Por supuesto, si este Número inspira a más analistas conductuales para realmente hacer investigación en redes neurales, pues mucho mejor.

REFERENCIAS

- McCulloch, W.S., & Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5, 115-133.
- Rashevsky, N. (1938). *Mathematical biophysics: Physico-mathematical foundations of biology*. University of Chicago Press.