

La perspectiva conexionista: hitos y limitaciones

Javier Aróztegui

*Jose M. Prados**

Universidad Complutense de Madrid

Resumen

Hace sesenta y cinco años, apareció publicado en el entonces incipiente *Bulletin of Mathematical Biophysics* un detenido análisis de las propiedades lógicas del funcionamiento nervioso. Tal artículo, firmado en 1943 por el neurofisiólogo Warren McCulloch junto al especialista en lógica Walter Pitts, supuso el inicio de lo que luego se dio en llamar el *enfoque neuronal de la computación*. Partiendo de la noción de computabilidad de Turing, dicho enfoque proponía que el sistema nervioso también realiza *cómputos*, al menos potencialmente y dentro de sus posibilidades espacio-temporales. Ello interesó rápidamente a autores de muy diversas disciplinas, tanto a matemáticos e ingenieros que empezaban a trabajar en *inteligencia artificial*, como a fisiólogos y psicólogos preocupados por el cada vez mayor «vacío» entre la ciencia de la conducta de aquellos años y las bases biológicas del comportamiento. Entre dichos autores cabe citar, entre otros muchos, a Donald O. Hebb (1949), para quien las funciones psicológicas son el resultado de la actividad conjunta de grupos de neuronas conformados mediante el aprendizaje, y a Frank Rosenblatt (1958), quien intentó simular la tarea principal del cerebro, según él, consistente en la clasificación y el reconocimiento de estímulos. Sin embargo, y debido a sus limitaciones, el enfoque neuronal de la computación acabó generando cierto desencanto. Pero poco después, junto a la crisis del enfoque clásico de la computación en torno a los conceptos de representación, racionalidad y procesamiento algorítmico, aparecieron nuevas aportaciones que incrementaron notablemente el potencial de la neurocomputación (redes dinámicas, máquinas Boltzmann, redes alimentadas hacia delante con retropropagación del error, etc.). El objetivo de este trabajo consiste en repasar algunos hitos de este enfoque de investigación así como las diferentes soluciones que ha aportado.

Palabras clave: Conexionismo, neurocomputación, redes neuronales artificiales.

* Correspondencia: Facultad de Psicología (UCM), Campus de Somosaguas, 28223 – Pozuelo (Madrid); Teléfono: 913942939. E-mail: <jmpatienza@psi.ucm.es>.

Abstract

Sixty five years ago, a detailed analysis about the logical properties of nervous system functioning was published by the *Bulletin of Mathematical Biophysics*. The paper written by the neurophysiologist Warren McCulloch and by the logician Walter Pitts, was the start of what is known as the *neuronal view of computation*. This view's starting point was quite simple: If any computation can be physically done, as was demonstrated by Turing, then the nervous system, due to its features, can also potentially perform any computation according to its spatial and temporal possibilities. Many authors from different disciplines quickly showed great interest, from mathematicians and engineers starting to work on Artificial Intelligence, to physiologists and psychologists concerned about the considerable gap between the Behavioral Science and its biological basis. Donald O. Hebb or Frank Rosenblatt were between them. According to Hebb (1949), psychological functions can be understood as the result of the joint activity of neurons groups (or assemblies) shaped by learning. And Rosenblatt (1958) tried to simulate the main task of the brain, according to him: the classification and categorization of stimulus. In spite of those great achievements, the neural view of computation generated some disenchantment due to its limitations. Erroneously, many thought this view had no growth potential. Later, along with the crisis around the classical view of computation (with concepts as representations, rationality and algorithmic processing), new proposals increased considerably the neurocomputation potential and renewed the interest about this ambitious project (*dynamic networks, Boltzmann Machines, feedforward networks with backpropagation of error, etc.*). The target of this paper is to review the milestones and difficulties occurred, and sometimes solved.

Keywords: Connectionism, Neurocomputation, Artificial Neural Networks.

INTRODUCCIÓN

Es innegable que la simulación conexionista acapara cada vez más el interés de los científicos cognitivos. Pero lejos de haber generado cambio paradigmático alguno, como hace veinte años se predecía de forma grandilocuente, hoy se entiende que ésta no es sino una potente herramienta de investigación más, mediante la cual se intentan emular ciertos fenómenos psicológicos. Su particularidad estriba en no olvidar que dichos fenómenos, en última instancia, son producidos por el cerebro. Pero aunque esté inspirado en el funcionamiento neuronal, un modelo conexionista no deja de ser un modelo abstracto del procesamiento de la información (Rumelhart, Hinton y Williams, 1986).

Gran parte del interés suscitado por la simulación conexionista se debe a su énfasis en el aprendizaje, su interactividad, su adaptabilidad y su interpolación. De igual forma, ha sido evidente la gran capacidad de evolución técnica de este enfoque a lo largo de los años, desdibujando algunas de las críticas que habían recaído sobre él. Además, presenta interesantes implicaciones teóricas (véase, por ejemplo, la reconsideración del nativismo que hacen Elman *et al.*, 1996).

El objetivo de este trabajo es revisar, con una aproximación histórica, los principales hitos y dificultades de este marco de investigación (vid. tabla 1), cuya indudable influencia para nuestra disciplina ya ha sido analizada con anterioridad (vid. Ruiz, Soler, Dasí, Algarabel y Pitarque, 1993).

PRIMERAS APROXIMACIONES

En 1943, la entonces incipiente *Bulletin of Mathematical Biophysics* de Nicolas Rashevsky publicó un trabajo firmado conjuntamente por Warren S. McCulloch (1898-1969) y Walter Pitts (1923-1969). La proyección de dicho trabajo fue enorme y constituyó la pieza seminal de lo que luego se dio en llamar el *enfoque neuronal de la computación*.

Warren S. McCulloch era, ante todo, un filósofo, influido por Descartes, Kant y Leibniz, entre otros. Se había formado en Yale pero sus preocupaciones por la «epistemología experimental», como él la llamaba, lo llevaron a especializarse en Nueva York, en el *Colegio de Médicos y Cirujanos*. Su intención era entender la mente y responder a las preguntas clásicas de la filosofía, pero a partir del cerebro. Así, después de doctorarse y ejercer durante algún tiempo en hospitales (el Bellevue y el Rockland State), pronto se trasladó al Instituto Neuropsiquiátrico de la Universidad de Illinois, en Chicago, donde pudo desarrollar su afán investigador. En esa ciudad, conoció a Walter Pitts, un brillante matemático y especialista en lógica, autodidacta y discípulo del positivista lógico Rudolf Carnap (1891-1970). Y Pitts, rápidamente intuyó la aplicación del formalismo de su mentor a las ideas de McCulloch (Arbib, 2000; Smalheiser, 2000).

El resultado de la colaboración entre ambos fue la formalización del cerebro como una red de neuronas. Y éstas, a su vez, como unidades de procesamiento lógico, como *autómatas* con un umbral, cuyo estado activo o inactivo indicaría un valor lógico de verdad o falsedad, y cuya conexión podría representar operaciones lógicas (como la conjunción o la negación lógicas). Este planteamiento partía de la reciente tesis de la *computabilidad* de Alan M. Turing (1912-1954), y Alonzo Church (1903-1995), y que algunos, como Garrido (2003), no dudan en relacionar con los *Ensayos sobre automática* de nuestro Leonardo Torres Quevedo (1852-1936).

Aunque la tesis de la computabilidad inicialmente supuso una contradicción más, junto con las de Gödel, a las aspiraciones formalistas de David Hilbert (1862-1943), también abrió la puerta al desarrollo técnico y científico por sus numerosas implicaciones. Una de ellas, es que si cualquier cómputo puede ser ejecutado físicamente, el propio sistema nervioso, dadas sus propiedades, también realiza cómputos. Precisamente, eso fue lo que analizaron McCulloch y Pitts, ofreciendo así una respuesta interesante al pertinaz problema del dualismo.

Aunque el trabajo de McCulloch y Pitts fue criticado por su excesiva simplificación del sistema fisiológico (Bishop, 1946), por sus errores, y por la oscuridad de algunas de sus partes debido a la notación carnapiana (Arbib, 1964), lo cierto es que tuvo un gran impacto. Por un lado, influyó en los matemáticos e ingenieros involucrados en la producción de lo que a partir de 1956 se llamó *inteligencia artificial*; de hecho, John von Neumann (1903-1957) y su discípulo Marvin Minsky (nacido en 1927), comenzaron emulando las neuronas formales con tubos de vacío (como el famoso SNARC). Por otro lado, el trabajo de McCulloch y Pitts también influyó enormemente entre los fisiólogos y psicólogos preocupados por el cada vez mayor «abismo» entre la ciencia de la conducta de aquellos años y las bases biológicas del comportamiento. Ese fue el caso del canadiense Donald O. Hebb (1904-1985), formado junto a Wilder G. Penfield (1891-1976) y al eminente Karl S. Lashley (1890-1958). Para Hebb, inspirado también en el trabajo de Rafael Lorente de Nó (1902-1990), brillante discípulo de Ramón y Cajal, el aprendizaje podía entenderse como el establecimiento de cambios estables en los niveles sinápticos

interneuronales. Esto es, si con un estímulo inicial se activan conjuntamente algunas neuronas, con cada presentación repetida del mismo se refuerza la conexión entre dichas neuronas hasta el punto de seguir excitándose unas a otras cuando cesa el estímulo. La formación y la función de estas «asambleas neuronales» serían, en última instancia, las responsables de las funciones cognitivas (Hebb, 1949).

Pese a sus incorrecciones, esta sugerente especulación del funcionamiento neuronal ha gozado de notable influencia; no en vano, uno de los principales grupos de algoritmos de aprendizaje empleados por las distintas arquitecturas de redes neuronales artificiales está basado en la *regla de Hebb*. Además, esta propuesta permitía responder al problema de cómo aprende un sistema a pesar del ruido o de un daño parcial del mismo.

Otra aportación de interés fue la del psicólogo estadounidense Frank Rosenblatt (1928-1969), quien después de estudiar Psicología Social en la Universidad de Cornell, y doctorarse en Psicopatología Experimental, intentó simular la tarea principal del cerebro, según él, consistente en la clasificación y el reconocimiento de estímulos. Dicha simulación, bautizada con el nombre de *Perceptrón* (1958), se realizó en un IBM 704 y suponía la unión entre las neuronas formales de McCulloch y Pitts y el modelo llamado *Pandemonium*, del matemático Oliver Selfridge, diseñado originalmente para aprender a descifrar palabras y letras en código Morse.

La propuesta de Rosenblatt fue interesante por introducir la cuestión de la memoria en las redes neuronales, y también porque no consideraba adecuado el tratamiento del sistema nervioso desde la lógica formal (se anticipaba así al actual movimiento de la *lógica difusa*); mientras la lógica clásica es bivalente, emplea únicamente dos valores de verdad (0 y 1), las variaciones de comportamiento de las neuronas son variaciones en la tasa de disparo sináptico, para lo que parece más apropiado emplear un rango de valores.

Pese a estos y otros muchos avances, el enfoque neuronal de la computación acabó generando cierto desencanto ya que sus limitaciones no permitían diseñar procedimientos para aprenderlo «todo». Y erróneamente, muchos creyeron que dicho enfoque apenas podría progresar, frente a la inteligencia artificial simbólica que sí parecía hacerlo en aquellos momentos (Minsky y Papert, 1969).

RESURGIMIENTO

Pero poco después, junto a la crisis del enfoque clásico de la computación en torno a los conceptos de representación, racionalidad y procesamiento algorítmico (Cobos, 2005), fueron sucediéndose nuevas aportaciones que incrementaron notablemente el potencial de la neurocomputación e hicieron renacer el interés por este ambicioso proyecto. Algunas de ellas fueron las *redes dinámicas* del físico John J. Hopfield (1982), que permitían entender el procesamiento de la información como un proceso continuo que evoluciona con el tiempo; las *máquinas Boltzmann* (Hinton y Sejnowski, 1983), las cuales mejoraban las anteriores con una regla de aprendizaje capaz de solucionar problemas complejos; y, en un lugar muy destacado, dada su enorme influencia, las redes alimentadas hacia delante con *retropropagación del*

error (Rumelhart, Hinton y Williams, 1986), que supusieron un aumento considerable de la capacidad de aprendizaje de las redes neuronales respecto a lo que habían conseguido hasta entonces (Anderson y Rosenfeld, 1998; Medler, 1998; Widrow y Lehr, 1990).

Y el conexionismo volvió a suscitar interés. En un análisis del *Social Sciences Citation Index* y del *Sciences Citation Index*, Ruiz *et al.* (1993), encontraron que, tan sólo en la década de los ochenta, David Rumelhart había sido citado 3397 veces (con 784 citas más que Herbert A. Simon), y James McClelland, 1571 veces.

De esa forma, en los siguientes años no tardaron en aparecer modelos conexionistas para un sinnúmero de fenómenos: entre otros, el aprendizaje discriminativo (Pearce, 1994), el aprendizaje de categorías (Kruschke, 1992), la adquisición de estructuras silábicas (Vroomen, van den Bosch y De Gelder, 1998), el reconocimiento de palabras y el nombrado (Seidenberg y McClelland, 1989), el razonamiento condicional (Leighton y Dawson, 2001), la atribución causal (Van Overwalle, 1998), etc. Incluso, empezaron a desarrollarse acercamientos a algunos ámbitos de crítica clásica, como el de las representaciones estructuradas (Doumas y Hummel, 2005), y las propias redes conexionistas comenzaron a ser utilizadas para poner a prueba hipótesis acerca de la función de algunas áreas cerebrales (Rolls y Treves, 1998).

Entre las limitaciones y retos aún pendientes están el problema del escalamiento, el problema de la representación de la estructura abstracta de los objetos y el de la modelización compleja, entre muchos otros.

LAS REDES NEURONALES EN ESPAÑA

En nuestro país, las redes neuronales han interesado e interesan a científicos y profesionales de diversas disciplinas (vid. Martín y Sanz, 2001, para una aproximación desde la ingeniería). Dentro del ámbito de la Psicología, cada vez son más los equipos de investigación españoles que utilizan la simulación conexionista (Martín-Valdivia, Ureña-López y García-Vega, 2007; Montaña, Palmer y Fernández, 2002; Pitarque, Ruiz y Algarabel, 1995; Sopena, Ramos, López-Moliner y Gilboy, 2002). Y en el plano divulgativo, Jose A. García Madruga ha sido uno de sus principales mentores, desde la comunicación titulada «Conexionismo y desarrollo intelectual», presentada en 1990 al Congreso del Colegio de Psicólogos celebrado en Valencia; la selección de capítulos que realizó en 1992 del clásico *Parallel distributed processing* de Rumelhart y McClelland (1986); y sus sucesivos libros de texto, en los que realiza sencillas introducciones al tema (p. ej., García-Madruga, Gutiérrez y Carriedo, 2002). Otro autor es Pedro Luis Cobos, cuya reciente monografía *Conexionismo y cognición* ha conseguido aunar divulgación y rigor de forma espléndida, con la esperanza de colaborar a la difusión de este enfoque en la Psicología española.

Pero, ¿por qué el conexionismo no es tan popular en España como en otros países? Aunque el objetivo de este trabajo no era responder a esta pregunta, creemos que una relectura de *Los tónicos de la voluntad* de Ramón y Cajal podría ayudar a desentrañar algunas de las complejas razones; quizás no hayan cambiado tanto las cosas como parece.

Referencias

- Anderson, J. A. y Rosenfeld, E. (1998). *Talking Nets: An oral history of neural networks*. Cambridge, MA: MIT Press.
- Arbib, M. A. (1964). *Brains, Machines and Mathematics*. Nueva York: McGraw-Hill.
- Arbib, M. A. (2000). Warren McCulloch's search for the logic of the nervous system. *Perspectives in Biology and Medicine*, 43, 193-216.
- Bishop, G. H. (1946). Nerve and synaptic conduction. *Annual Review of Physiology*, 8, 355-374.
- Cobos, P. L. (2005). *Conexionismo y cognición*. Madrid: Pirámide.
- Doumas, L. A. A., y Hummel, J. E. (2005). A symbolic-connectionist model of relation discovery. En B. G. Bara, L. Barsalou, y M. Bucciarelli (Eds.), *Proceedings of the Twenty-Third Annual Conference of the Cognitive Science Society*. Mahwah, NJ: LEA.
- Elman, J. L., Bates, E. A., Johnson, M. H., Karmiloff-Smith, A., Parisi, D., y Plunkett, K. (1996). *Rethinking innateness: a connectionist perspective on development*. Cambridge, MA: MIT Press.
- García Madruga, J. A., Gutiérrez, F. y Carriedo, N. (Eds.) (2002). *Desarrollo cognitivo y lingüístico* (Vol. I). Madrid: UNED.
- Garrido, M. (2003). El autómata con discernimiento de Torres Quevedo: Un antecedente del modelo de Turing. *Teorema*, 17, 1-7.
- Hebb, D. O. (1949). *The organization of behavior*. Nueva York: Wiley.
- Hinton, G. E., y Anderson, J. A. (Eds.) (1981). *Parallel models of associative memory*. Hillsdale, NJ: Erlbaum.
- Hinton, G. E., y Sejnowski, T. J. (1983). Optimal perceptual inference. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition* (pp. 448-453).
- Hopfield, J. (1982). Neural networks and physical systems with emergent collective computational abilities. *Proceedings of the National Academy of Science*, 79, 2554-2558.
- Kruschke, J. K. (1992). ALCOVE: An exemplar-based connectionist model of category learning. *Psychological Review*, 99, 22-44.
- Leighton, J. P., y Dawson, M. R. W. (2001). A parallel distributed processing model of Wason's selection task. *Cognitive Systems Research*, 2, 207-231.
- Martín, B., y A. Sanz (2001). *Redes neuronales y sistemas borrosos*. Madrid: Ra-Ma (2ª Ed.).
- Martín-Valdivia, M. T., Ureña-López, L. A. y García-Vega, M. (2007). The learning vector quantization algorithm applied to automatic text classification tasks. *Neural Networks*, 20, 748-756.
- McCulloch, W. S., y Pitts, W. (1943). A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, 5, 115-133.
- Medler, D. A. (1998). A brief history of connectionism. *Neural Computing Surveys*, 1, 18-72.
- Minsky, M. y Papert, S. (1969). *Perceptrons*. Cambridge, MA: MIT Press.
- Montaño, J. J., Palmer, A., y Fernández, C. (2002). Redes neuronales artificiales: abriendo la caja negra. *Metodología de las Ciencias del Comportamiento*, 4, 77-93.

- Pearce, J. M. (1994). Similarity and discrimination: A selective review and a connectionist model. *Psychological Review*, 101, 587-607.
- Pitarque, A., Ruiz, J. C., y Algarabel, S. (1995). Simulation of dissociative effects on explicit/implicit memory tasks from a connectionist model. *Psicothema*, 7, 339-350.
- Ramón y Cajal, S. (1941). *Los tónicos de la voluntad*. Madrid: Espasa-Calpe.
- Rosenblatt, F. (1958). The perceptron: a probabilistic model for information storage and organization in the brain. *Psychological Review*, 65, 386-408.
- Rolls, E. T. y Treves, A. (1998). *Neural networks and brain function*. Oxford: Oxford University Press.
- Ruiz, J. C., Soler, M. J., Dasí, C., Algarabel, S., y Pitarque, A. (1993). Estudio histórico de la influencia del conexionismo en la ciencia cognitiva actual. *Revista de Historia de la Psicología*, 14, 43-57.
- Rumelhart, D. E., Hinton, G. E., y Williams, R. J. (1986). Learning internal representations by error propagation. En D. E. Rumelhart, J. L. McClelland, y el grupo PDP (1986). *Parallel distributed processing: explorations in the microstructure of cognition* (Vol. 1, pp. 318-362). Cambridge, MA: MIT Press.
- Seidenberg, M. S., y McClelland, J. L. (1989). A distributed, developmental model of word recognition and naming. *Psychological Review*, 96, 523-568.
- Smalheiser, N. R. (2000). Walter Pitts. *Perspectives in Biology and Medicine*, 43, 217-226.
- Sopena, J. M., Ramos, P. J., López-Moliner, J., y Gilboy, E. (2002). Composicionalidad, cómputo de estructura y redes neuronales. *Estudios de Psicología*, 23, 175-235.
- Van Overwalle, F. (1998). Causal explanation as constraint satisfaction: A critique and a feedforward connectionist alternative. *Journal of Personality and Social Psychology*, 74, 312-328.
- Vroomen, J., van den Bosch, A., y de Gelder, B. (1998). A connectionist model for bootstrap learning of syllabic structure. *Language and Cognitive Processes*, 13, 193-220.
- Widrow, B. y Lehr, M. A. (1990). 30 years of adaptive neural networks: Perceptron, Madaline and Backpropagation. *Proceedings of IEEE*, 78, 1415-1442.

TABLA 1
Cronología sucinta de la perspectiva conexionista

Año	Suceso
1943	Warren McCulloch y Walter Pitts realizan una formalización de la actividad neuronal
1949	Donald O. Hebb acuña la expresión «asamblea neuronal» y define el aprendizaje de forma neurológica
1951	Marvin Minsky simula redes neuronales con SNARC
1958	Frank Rosenblatt simula en un IBM 704 el aprendizaje perceptivo
1959	Se edita la revista <i>Computer Journal</i>
1961	Se edita la revista <i>Kibernetik</i>
1969	Minsky (y Papert) critica severamente el «perceptrón»
1977	James A. Anderson propone un modelo lineal (asociador lineal)
1981	Geoffrey Hinton y James A. Anderson editan una monografía de modelos de memoria asociativa
1982	John Hopfield propone una arquitectura dinámica
1983	Hinton y Sejnowski proponen una nueva estructura neuronal; se edita la revista <i>Connectionist Symbol Processing</i>
1986	Se publica el influyente libro editado por David E. Rumelhart y James L. McClelland junto al grupo PDP
1987	Primera conferencia internacional en redes neuronales celebrada en San Diego (EE. UU.); aparece la International Neural Network Society (INNS), cuyo primer presidente fue Stephen Grossberg; se edita la revista <i>Complex Systems</i>
1988	Primer congreso de la INNS en Boston (EEUU); se editan las revistas <i>Neural Computation</i> , <i>Neural Networks</i> , <i>Network: Computation in Neural Systems</i> , y <i>Connection Science</i>
1989	Se edita la revista <i>International Journal of Neural Systems</i>
1997	Se edita la revista <i>Neural Computing Services</i>
1998	Aparece la European Neural Network Society (ENNS); se edita la revista <i>Neural Computing Surveys</i>