



Avatares del paradigma conexionista

Asier Arias

Dept. de Filosofía IV, Universidad Complutense de Madrid, España

Tipo de artículo: Clásicos.

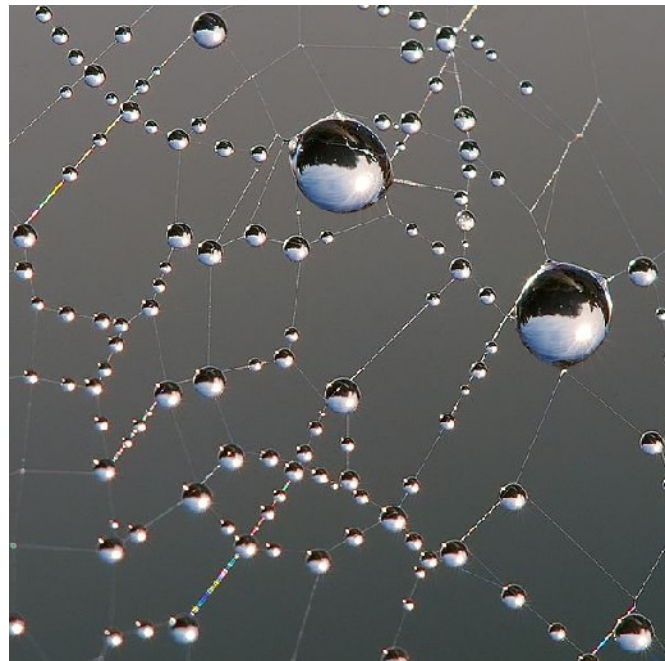
Disciplinas: Psicología, Filosofía, Neurociencias, Inteligencia Artificial.

Etiquetas: conexionismo, embodiment, neurociencia computacional, mente extendida, cognición enactiva.

Hace ahora un año que falleciera David Everett Rumelhart, conexionista pionero y editor junto con James Lloyd McClelland de la llamada “biblia del Procesamiento Distribuido en Paralelo” (PDP). Al mismo tiempo, los dos volúmenes de la referida biblia celebraban su vigésimo quinto cumpleaños. También uno de los padres de la inteligencia artificial, John McCarthy, que bautizara de hecho la disciplina, nos abandonó el pasado mes de octubre a los 84 años. Estas tres efemérides nos sirven de pretexto para un breve recorrido por las derivas y obstáculos a los que se ha enfrentado el llamado paradigma conexionista desde su puesta en escena como alternativa al cognitivismo clásico.

El modelo computacional de la mente sobre el que las ciencias cognitivas se han erigido ha conocido diferentes formulaciones desde sus inicios, los cuales suelen remontarse a las postrimerías de los cuarenta (Gardner, 1985). Documentos, testigos e historiadores indican de consuno que los años cincuenta vieron dar sus primeros pasos a una nueva forma de entender la mente: fue aquélla la década de Newell, Simon, McCarthy, Minsky, el GPS, el LISP, el Perceptron... El suelo estaba abonado ya en aquel momento para la eclosión del paradigma conexionista: McCulloch, Pitts y Rosenblatt habían sembrado ya sus semillas, pero en sus inicios las ciencias cognitivas adoptaron el paradigma simbólico.

La mente computacional fue concebida inicialmente como una mente simbólica al estilo de los ordenadores digitales, es decir, como un procesador serial de información en forma de



(cc) Luc Viatour / www.lucnix.be

símbolos discretos, estructurados sintácticamente y manipulados de acuerdo con reglas ajenas al contenido representado. En esos símbolos y su manipulación de acuerdo con reglas formales estribaría para el cognitivismo clásico la actividad representacional de la mente.

El conexionismo y los modelos PDP (Rumelhart, McClelland et al., 1986; McClelland, Rumelhart et al., 1986) trajeron consigo una profunda reformulación de la noción cognitivista de representación, que viraría de la discreción de los símbolos a la dispersión en redes conexionistas: en ellas las representaciones se hallan no en elementos sino entre elementos; se trata de representaciones distribuidas que dependen de propiedades sistémicas dentro de redes neuronales artificiales. Si para el cognitivismo clásico el proceso paradigmático era la inferencia reglada, para el conexionismo lo son el reconocimiento de patrones y un tipo de aprendizaje que prescinde de reglas previas: las reglas surgen de y con la experiencia.

Hay dos grandes tipos de redes conexionistas. Las denominadas localistas poseen estructuras cuyos nodos representan entidades o conceptos y las conexiones entre nodos trazan nexos entre ellos. Como en el paradigma clásico, cada nodo representa una entidad o propiedad concreta. La diferencia estriba en que los nodos pueden estar activados en distinto grado y transmiten parte de esa activación a sus nodos conexos, los cuales suman la activación recibida, agregando así la evidencia proveniente de distintas fuentes (Figura 1).

En cambio, en otros modelos, denominados redes neuronales subsimbólicas, ningún nodo representa nada en sí mismo. Es el patrón de activación de las unidades y el patrón de conectividad entre ellas el que lo hace. Las representaciones son ahora patrones de activación y conexión, y las computaciones son transiciones entre tales patrones (Figura 2).

El conexionismo comenzó su auge a mediados de la década de los ochenta y se extendió la opinión de que un nuevo marco para la comprensión de la mente se abría paso con él, desplazando al simbolismo clásico y ofreciendo una alternativa explicativa radicalmente diferente. No obstante, algunos defensores del modelo computacional tradicional rechazaron esta opinión y propusieron que el avance era más de forma que de fondo. Así, por ejemplo, Fodor y Pylyshyn (1988) argumentaron que los modelos conexionistas no suponen una alternativa real al paradigma simbólico: no serían más que formas de implementar el carácter computacional-simbólico de la mente, que no por ser, aparentemente, más similares al funcionamiento del cerebro se hallan en realidad más cerca de las bases neurobiológicas de la mente.

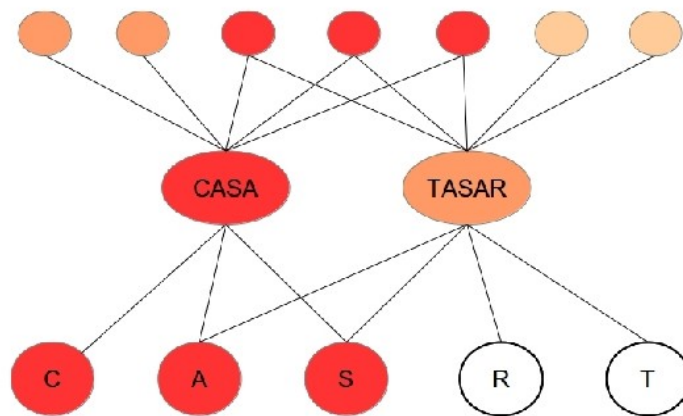


Figura 1.- Red localista con tres niveles. Cada círculo representa una unidad o neurona artificial y las líneas conexiones entre las mismas. El inferior es el nivel de input: las unidades son sensibles a la presentación de diferentes letras, en este caso la palabra “casa”. Este nivel envía señales a un nivel intermedio de unidades léxicas almacenadas. En el nivel superior, el sistema ofrece como output conceptos semánticamente relacionados con los nodos léxicos. En este ejemplo, las unidades 3ª, 4ª y 5ª, relacionadas con las dos unidades intermedias, representan el concepto de “lugar donde se vive”, mientras que las unidades 1ª y 2ª representan significados alternativos de la palabra “casa” (p.ej., “linaje”). Las unidades de la capa intermedia de nuestro ejemplo reciben e integran señales procedentes de las unidades de input: si la suma de las señales recibidas por una unidad excede un determinado valor de umbral, ésta se activa transmitiendo su activación a los nodos de la capa de output. En la figura, el nivel de activación se indica mediante la intensidad del color.

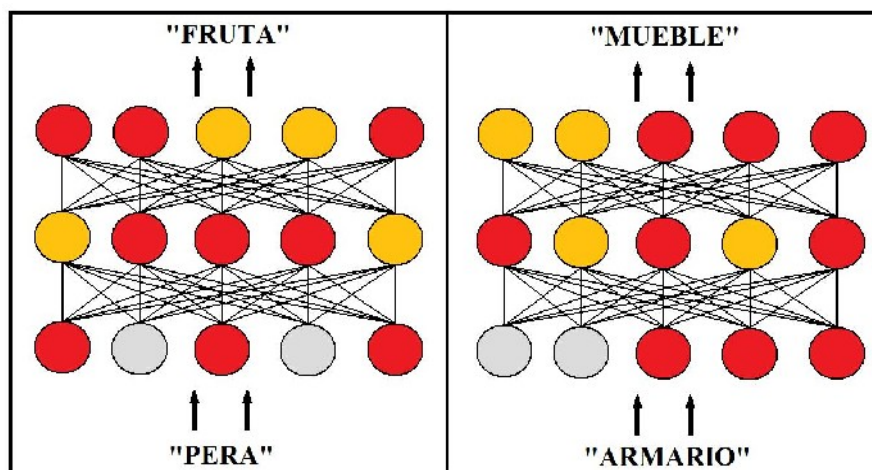


Figura 2.- En las redes conexionistas subsimbólicas la unidad de representación es el propio patrón global de activación y conexión entre las unidades: ninguna unidad representa por sí misma entidades o conceptos concretos. La fuerza de las conexiones entre unidades varía en función del “entrenamiento” de la red mediante su exposición repetida a determinados estímulos hasta lograr una relación apropiada entre input y output. El sistema, dada su arquitectura, su flujo de información y su historia de aprendizaje, ofrece como respuesta a un patrón de input un determinado patrón de activación en las unidades de output. En este ejemplo, el objetivo de la red es la categorización de palabras. Los estímulos visuales “pera” o “armario” son convertidos a patrones de activación en las unidades de input (en la parte inferior). Tras propagarse la activación en la red (en mayor cantidad si atraviesa conexiones más fuertes) se generan diferentes patrones en el nivel de output (arriba), correspondientes en nuestro ejemplo a los conceptos o categorías supraordinadas de “fruta” y “mueble”.

Otro tipo de acometida anticonexionista vino más tarde desde concepciones como la cognición corporeizada (Varela, Thompson & Rosch, 1991; Clark, 1997; Lakoff & Johnson, 1999), la cognición enactiva (Varela, Thompson & Rosch, 1991), la cognición extendida (Clark & Chalmers, 1998) o la robótica situada (Brooks, 1991), desde las que también se pone en tela de juicio que el enfoque conexionista sea un marco teórico esencialmente diferente del simbolista. Esta corriente crítica, por su parte, viene enfatizando el hecho de que la modelización computacional corre el riesgo de perder en su abstracción el contacto con el modo en que los procesos cognitivos se hallan conectados con los cuerpos, los sentidos y el mundo externo en que actúan los organismos.

Como balance general, frente a las representaciones simbólicas propias del cognitivismo tradicional, las representaciones conexionistas cuentan no sólo con la habitualmente señalada virtud de una mayor plausibilidad biológica, sino con la ventaja de ofrecer una explicación flexible de la cognición. Esto se debe a que permiten hacer uso de información parcial contenida en patrones incompletos. Por otra parte, tanto simbolistas como conexionistas conciben la cognición, en último término, de un modo abstracto y centralista, escindiendo los sistemas cognitivos del medio y los organismos en que tiene lugar la cognición. Frente a esa abstracción vienen las concepciones situadas y corporeizadas reivindicando el carácter concreto de la cognición (vid. Bietti, 2011, <http://www.cienciacognitiva.org/?p=258>, y Gomila, 2009, <http://www.cienciacognitiva.org/?p=69>).

Referencias

- Brooks, R. (1991). Intelligence without representation. *Artificial Intelligence*, 47, 139-159.
- Clark, A. (1997). *Being there: Putting brain, body and world together again*. Cambridge, MA: MIT Press. (Trad.: Barcelona: Paidós, 1999).

- Clark, A., y Chalmers, D. J. (1998). The extended mind. *Analysis*, 58, 7-19 (Trad.: Oviedo: KRK, 2011).
- Fodor, J. A., y Pylyshyn, Z. W. (1988). Connectionism and cognitive architecture. *Cognition*, 28, 3-71.
- Gardner, H. (1985). *The mind's new science. A history of the cognitive revolution*. New York: Basic Books (Trad.: Barcelona: Paidós, 1988).
- Lakoff, G., y Johnson, M. (1999). *Philosophy in the flesh. The embodied mind and its challenge to Western thought*. Nueva York: Basic Books.
- McClelland, J. L., Rumelhart, D. E., and the PDP Research Group (1986). *Parallel distributed processing: Explorations in the microstructure of cognition. Vol. 2*. Cambridge, MA: MIT Press.
- Rumelhart, D. E., McClelland, J. L., and the PDP Research Group (1986). *Parallel distributed processing: Explorations in the microstructure of cognition. Vol. 1*. Cambridge, MA: MIT Press.
- Varela, F. J., Thompson, E., y Rosch, E. (1991). *The embodied mind: Cognitive science and human experience*. MIT Press. (Trad.: Bracelona: Gedisa, 1992).

Manuscrito recibido el 16 de octubre de 2011.

Aceptado el 6 de febrero de 2012.