

---

## LA NEUROCIENCIA SOCIAL Y EL DESARROLLO DE ROBOTS SOCIALIZADOS

ISRAEL GRANDE-GARCÍA

Aunque mi formación profesional la hice en psicología y filosofía, toda mi actividad académica la he desarrollado en el área más general de las ciencias cognitivas, a pesar de que en México —desafortunadamente— sean muy pocos los que se dedican a este campo de estudio y aún no exista un posgrado o un centro de investigación exclusivos de este campo. Las ideas y reflexiones que siguen las haré entonces en el contexto de las ciencias cognitivas y específicamente desde la naciente neurociencia social.

### ¿DESDE SU DISCIPLINA PROFESIONAL, QUÉ CONOCIMIENTOS DEBEN DESARROLLARSE PARA HACER POSIBLE QUÉ FUTURO(S)?

El catastrofismo al que nos han acostumbrado las películas de Hollywood y conocidas obras de ciencia ficción, hacen pensar que, de seguirse desarrollando la inteligencia artificial, un día las máquinas terminarán dominándonos. Aunque soy amante de la ciencia ficción, personalmente no comparto esta imagen pesimista y creo más en una conjugación armoniosa hombre-máquina para el futuro a mediano plazo (digamos unos cincuenta a cien años) como alguna vez lo planteó Isaac Asimov. El siguiente e inevitable paso será el desarrollo de sistemas artificiales conscientes. Un ejemplo es el agente inteligente IDA (*Intelligent Distribution Agent*) desarrollado por Stan Franklin (Sun y Franklin 2007). El IDA se basa en la teoría de espacio laboral global, de Berard Baars (Baars 2007), y su diseño básico consiste en un espacio competitivo donde muchos 'codelets' (pequeñas piezas de códigos informáticos especializados en tareas simples) provenientes de varios mecanismos de memoria (a corto y largo plazo, episódica, etc.) compiten para obtener un ingreso a la conciencia.

---

Posgrado en Filosofía de la Ciencia (línea terminal, Filosofía de las Ciencias Cognitivas), Facultad de Filosofía y Letras / Instituto de Investigaciones Filosóficas, UNAM, México.  
israel.grandegarcia@gmail.com

Última contribución en *Ludus Vitalis*: "The evolution of brain and mind: a non-equilibrium thermodynamics approach," vol. XV, num. 27, 2007, pp. 103-125.

*Ludus Vitalis*, vol. XVI, num. 29, 2008, pp. 177-180.

Un problema, quizá el principal, con este tipo de enfoques, es que todavía se basan en la idea de que la conciencia y otras funciones cognitivas dependen de manera directa sólo de los sucesos que ocurren en el cerebro. Contrario a este enfoque, un campo relativamente nuevo de las neurociencias y las ciencias cognitivas, la *neurociencia social*, supone que las funciones cognitivas no dependen exclusivamente de los sucesos en el cerebro, sino que el ambiente social es el que en gran medida ha moldeado la arquitectura cerebral. La neurociencia social parte de la idea de que los seres humanos, al igual que otras especies animales, somos seres básicamente sociales, y se ha abocado a la tarea de estudiar las bases biológicas de la cognición social, es decir, de todas aquellas capacidades cognitivas que nos permiten sobrevivir en el ambiente social (Grande-García 2008).

Por lo tanto, las funciones cognitivas no se pueden estudiar al margen de lo social y en consecuencia, aunque la neurociencia social está empezando a desarrollarse, uno de sus impactos inevitables se verá en el desarrollo de sistemas artificiales que interactúen en ambientes sociales con seres humanos y con otros agentes artificiales.

#### ¿CÓMO PARTICIPAN ESTAS IDEAS DE FUTURO EN EL DESARROLLO DE ESOS CONOCIMIENTOS?

Uno de los más importantes hallazgos en las neurociencias en los últimos años fue el de un conjunto de neuronas en la región premotora de monos y humanos que dispara cuando un agente lleva a cabo acciones como asir y manipular, pero también cuando observa a otro agente llevar a cabo movimientos similares, por lo que recibieron el nombre de *neuronas espejo* (Rizzolatti y Sinigaglia 2006). Se cree que este sistema de neuronas espejo (NE) puede ser un mecanismo importante que subyace a la cognición social y en especial a las capacidades de imitación, entendimiento y predicción de las acciones de los demás.

Algunos investigadores en inteligencia artificial han hecho eco del descubrimiento de las NE y han desarrollado sistemas robóticos capaces de imitar y predecir las acciones propias y ajenas, utilizando arquitecturas que intentan igualar el funcionamiento de las NE. Uno de estos modelos (Demiris y Johnson 2003) consiste en un sistema robótico en cuyo sistema motor se implementa un sistema distribuido de modelos que se basan en una predicción precisa como medio para clasificar acciones hechas por otro agente. Los investigadores han mostrado con éxito que una arquitectura de este tipo hace que los sistemas artificiales sean capaces de imitar acciones nuevas viendo a otro agente (en este caso, los propios desarrolladores humanos).

Otro ejemplo es el de un grupo de investigadores del MIT (Breazeal et al. 2005) quienes, inspirados también por el descubrimiento de las NE, han

desarrollado un sistema robótico que puede ser capaz de imitar y aprender expresiones faciales humanas, para inferir reacciones afectivas y usar este conocimiento para guiar su conducta y predecir la conducta de los demás.

Aunque impresionantes *per se*, estos no son sino los primeros pasos para la construcción de robots socializados, son un ejemplo del impacto que está teniendo ya la neurociencia social en otros campos de investigación. El futuro a corto y mediano plazo será, creo yo, nuestra convivencia cooperativa con estos sistemas artificiales socializados y, esperemos, sociables y amistosos.

En su contribución a este foro, el doctor Andrés Moya (2007) ha hecho algunas interesantes reflexiones sobre la relación entre seres humanos y posibles sistemas artificiales muy desarrollados, los *cyborgs*. No quiero aventurarme a reflexionar sobre un futuro más lejano y preguntarme si un día habrá un enfrentamiento con nuestras propias creaciones artificiales y si eventualmente seremos remplazados. Como lo dije al principio, personalmente tengo una visión mucho más optimista y creo que será más interesante llegar un día a casa y que mi “computadora inteligente, autoconsciente y socializada” me pregunte qué me sucede al ver mi rostro apagado.

Por supuesto, puedo estar pecando de un exceso de optimismo y se me puede acusar de estar olvidando los malos y peligrosos usos que este tipo de tecnología podría tener (y de hecho tiene), como en la guerra y el crimen. Sin embargo, creo más en la posibilidad del día en que tales robots socializados y conscientes, más que reemplazarnos, finalmente sean quienes logren hacer mayor conciencia social y mejorar la calidad de vida en nuestro planeta. La conjugación de los esfuerzos entre campos de investigación, como la neurociencia social y la inteligencia artificial, nos pueden dar una pista de cuál puede ser el futuro, al menos a corto y mediano plazo, y claro, menos terrible y frío que como lo vemos en las películas.

## REFERENCIAS

- Baars, B. (2007), "The global workspace theory of consciousness," in M. Veltmans y S. Schneider (eds.), *The Blackwell Companion to Consciousness*. Oxford: Blackwell, pp. 236-246.
- Breazel, C., Buchsbaum, D., Gray, J., Gatenby, D. y Blumberg, B. (2005), "Learning from and about others: towards using imitation to bootstrap the social understanding of others by robots," *Artificial Life* 11: 31-62.
- Demiris, Y. y Johnson, M. (2003), "Distributed, predictive perception of actions: a biologically inspired robotics architecture for imitation and learning," *Connection Science* 15: 231-243.
- Grande-García, I. (2008), "¿Qué es la neurociencia cognitiva social?: los temas fundamentales en el estudio de las bases cerebrales de la cognición social", en J. A. Durand Alcántara e I. Grande-García (coords.), *Psicología y ciencias sociales: extensión temática, Tomo I*. México: Facultad de Estudios Superiores Zaragoza, UNAM, pp. 61-113.
- Moya, A. (2007), Hombres y cyborgs", *Ludus Vitalis* XV (28): 227-230.
- Rizzolatti, G. y Sinigaglia, C. (2006), *So quel che fai: il cervello che agisce e i neuroni specchio*. Milán: Raffaello Cortina Editore.
- Sun, R. y Franklin, S. (2007). "Computational models of consciousness: a taxonomy and some examples," in P. D. Zelazo, M. Moscovitch y E. Thompson (eds.), *The Cambridge Handbook of Consciousness*. Nueva York: Cambridge University Press, pp. 151-174.