



Fundamentos de Inteligencia Artificial: Entrevista con Antonio Benítez

Asier Arias

Departamento de Teoría del Conocimiento, Estética e Historia del Pensamiento, Universidad Complutense de Madrid, España

Tipo de artículo: Entrevista.

Disciplinas: Psicología, Filosofía, Inteligencia Artificial.

Etiquetas: tendencias en ciencia cognitiva, filosofía de la mente, embodiment.

Antonio Benítez, profesor de Lógica en la Universidad Complutense de Madrid, ha dedicado buena parte de su labor docente e investigadora a la inteligencia artificial y a la filosofía de la mente y de la ciencia cognitiva. Recientemente ha publicado una obra en tres volúmenes titulada Fundamentos de Inteligencia Artificial (el tercer volumen, dedicado a la inteligencia artificial bioinspirada, aparecerá en los próximos meses). La misma ofrece una guía práctica de introducción a la programación en inteligencia artificial de clara y cabal factura, además de una cuidada introducción a los aspectos teóricos de la disciplina. En esta entrevista nos habla de esta obra y de la ciencia cognitiva en general.



Antonio Benítez. (c) Todos los derechos reservados.

Pregunta: Quisiéramos empezar agradeciendo su trabajo en la elaboración de estos Fundamentos de Inteligencia Artificial y pidiéndole que describa a los lectores de Ciencia Cognitiva el plan general de la obra.

Respuesta: En mis clases de Inteligencia Artificial (IA), desde hace décadas, siempre he hecho hincapié en que el estudiante aprenda a manejarse con programas, preferiblemente haciéndolos él mismo. Por eso, he enseñado a programar según métodos y técnicas de IA. Mi libro tiene esa orientación: de un lado, programas y, por otro lado, he intentado que el estudiante advirtiera que cada programa, cada técnica, cada sistema tiene como supuestos básicos varias hipótesis teóricas y que

éstas remiten a ciertos fundamentos filosóficos. Además he procurado respetar los distintos paradigmas que se han dado en IA: simbolismo, conexionismo y bioinspiración.

P: Puede que a más de un lector no familiarizado con la filosofía le resulte curioso encontrar citados a autores como Kant o Husserl en un texto como éste (referencias, hay que señalarlo, que no dificultan la comprensión del texto). Respecto del recurso a ideas provenientes de la filosofía, nos gustaría que resumiera el papel que en el capítulo 13 (tomo II) juegan algunas ideas de autores tan infrecuentes en textos de ciencias cognitivas como Ortega y Gasset y Alfred Schütz.

R: En mis estudios de filosofía he procurado prestar mucha atención a nuestra tradición filosófica. Creo que en la obra de Ortega y Zubiri hay ideas muy aprovechables para una filosofía de la mente. El punto de arranque es la idea de que en la vida cotidiana nos valemos de tipos de conocimientos que no son, no tienen por qué ser, estrictamente teóricos. Los investigadores en IA adelantaron la idea de que en las situaciones de la vida cotidiana todos echamos mano de un tipo de conocimiento que llamaron “guión”. Hay que entender “guión” como en el cine o en el teatro. Consecuencia de esta hipótesis es que seamos actores. La idea es antiquísima (persona, máscara en el mundo grecorromano). Así, por ejemplo, cuando vamos al banco a sacar dinero de nuestra cuenta corriente hemos de actuar según un guión que vale para todos. Si uno de nosotros va al banco a sacar dinero, puede que no se lo den si actúa siguiendo un guión inapropiado, por ejemplo, de hincha de un equipo de fútbol. Todas esas ideas fueron pensadas por sociólogos clásicos: Weber, Durkheim, Parsons, etc. He recurrido a la tradición weberiana. Schütz y Ortega se integran en esta tradición.

Tenemos entonces una concepción del hombre como actor de distintos y variados guiones: profesor, amante, padre de familia, diputado, etc. Esos guiones no han sido ideados, fabricados, pensados por el hombre actor, sino que los ha recibido de sus mayores. El sistema de conocimientos que son guiones, usos, vigencias, incluye otro elemento: creencias acerca de la realidad, de lo que hay. A esto lo llamo “ontología”. Las ontologías transmitidas cambian y son tan históricas como las vigencias, los usos, los guiones. Basta pensar en los españoles del siglo XI y los de hoy.

He relacionado la idea de ontologías con la de un sistema agente, de un agente dinámico. El agente está constituido por conocimientos (que incluyen guiones, usos, vigencias, etc.) y creencias heredadas que dictan lo que hay para él en el mundo. La realidad es correlativa a un sistema de creencias. Creo que las investigaciones de Schütz y de Ortega ayudan a entender esa perspectiva mejor que otros autores. En este sentido, he intentado ampliar el acervo de ideas y autores que suelen ser frecuentes en ciencias cognitivas.

P: Es habitual en filosofía de la mente la idea de que el aspecto subjetivo (cualitativo, fenoménico) de los estados mentales conscientes jamás podrá ser abordado por el computacionalismo. ¿En qué cree que se basa la creencia en esta imposibilidad?

R: La imposibilidad defendida se basa en lo que Turing llamó la “objección de Lady Lovelace” que se puede resumir así: cada elemento de un programa que pudiera simular una mente sería un algoritmo y el programa mismo tendría carácter algorítmico. Por tanto, cualquier output ha de estar previsto por el programador. De ahí parece seguirse que dos autómatas -por ejemplo- programados de idéntica manera han de dar el mismo output si están en la misma situación, porque la situación no sería otra cosa para cada uno de ellos que los argumentos con que el programa trabaja.

Mi contraargumento sigue al propio Turing y se basa en cosas conocidas y de las que tenemos experiencia real, no es meramente especulativo. Un agente se puede programar como una clase de objetos (metodología orientada a objetos). Dos instancias de dicha clase serían dos agentes. En la misma situación, pueden tener outputs diferentes. El contenido de las propiedades -idénticas estructuralmente- de cada agente puede ser distinto. Pongo un ejemplo. Dos hombres partícipes de una situación, aunque compartirán bastantes percepciones de elementos de la situación, tendrán algunas otras percepciones que serán subjetivas. ¿Cómo con las mismas estructuras cerebrales y corporales los dos partícipes tienen contenidos distintos? Si pasamos a la IA y suponemos dos autómatas programados como instancias de una clase de sistema complejo, vemos que comparten elementos estructurales y, no obstante, los contenidos pueden ser distintos según la situación.



En la mayoría de las discusiones se pasa por alto que hay programas que desarrollan sistemas que tienen propiedades emergentes y contenidos variables y dependientes de las situaciones, del entorno y del sistema propio.

P: ¿Cuál es su perspectiva acerca del papel en la investigación en ciencias cognitivas de los diferentes marcos históricamente generados (del simbolismo clásico a la cognición corporeizada o “embodiment”)?

R: A mi juicio los tres “marcos históricos” o programas de investigación han jugado distinto papel en la elaboración de una teoría de la mente. La raíz del simbolismo clásico consistió en elaborar una nueva psicología teórica. Una hipótesis esencial del simbolismo es que la mente se puede estudiar, que no es una “caja negra” como quería el conductismo. Que la teoría buscada fuera una psicología llevó al estudio de las funciones mentales superiores: razonamiento, solución de problemas teóricos, lenguaje, etc. El conexionismo constituyó su núcleo de investigación con una propuesta fundamental: simular el cerebro mediante redes de neuronas artificiales porque la vida mental es actividad cerebral. Siempre he mantenido que este giro implica que los programas o la simulación pierden la función de contrastación empírica de hipótesis teóricas que tenían en el simbolismo. La inspiración biológica ha vuelto a perseguir la construcción de una teoría. De nuevo la simulación, los programas, vuelven a ser comprobación empírica de hipótesis teóricas. Creo que este nuevo programa de investigación puede albergar en su seno tanto hipótesis como aplicaciones de los paradigmas anteriores.

Esta “nueva” IA permite aunar neurociencias y modelización computacional. Por último, señalaré que mucho del trabajo en IA bioinspirada constituye un laboratorio de ideas. Si se me permite la exageración: la filosofía tiene por primera vez un laboratorio propio.

Referencias

Benítez, A. (2011) *Fundamentos de inteligencia artificial*. Madrid: Escolar y Mayo.

Listo para publicación el 27 de junio de 2012.