

Manuel de Vega
Universidad de la Laguna

Visita al País de Turing. ¿Un viaje de ida y vuelta?

El libro de Ángel Rivière *Objetos con mente* es un magnífico ensayo sobre los fundamentos de la ciencia cognitiva. Rivière realiza una profunda revisión teórica que incluye desde un original análisis de algunos pensadores clásicos, como Descartes o Leibniz, hasta las reflexiones de los modernos filósofos de la ciencia (Pylyshyn, Fodor, Searle, Dennet, etc.) o algunas aportaciones seleccionadas de la psicología cognitiva (Cosmides, Bruner, Johnson-Laird, etc.).

Sin embargo, el libro de Rivière es mucho más que una obra erudita o de mera revisión, es un conjunto de argumentos que pretenden desentrañar la naturaleza de las funciones cognitivas humanas. Rivière realiza un apasionante y apasionado viaje al País de Turing y nos trasmite la fascinación del hallazgo intelectual de Turing: un sencillo dispositivo mecánico (idealizado) puede ejecutar, en principio, cualquier cómputo, y ello también incluye las funciones mentales como la memoria, el razonamiento, o la comprensión y producción del lenguaje. Por extensión, el cerebro humano puede considerarse como un ejemplo particular de máquina de Turing, los ordenadores digitales son otro ejemplo. Es irrelevante el que el cerebro o los ordenadores difieran considerablemente entre sí, o difieran de la propia máquina de Turing en su estructura material. Lo importante es que serían funcionalmente equivalentes a la máquina de Turing. El cerebro y el ordenador operarían con un vocabulario de símbolos (entidades discretas y arbitrarias) de acuerdo con ciertos principios sintácticos. La misión del cerebro como la del ordenador es ejecutar procedimientos efectivos o algoritmos.

Pylyshyn y Fodor, los grandes cancerberos de la ciencia cognitiva, retienen intacto el postulado de Turing. El cerebro y el ordenador son máquinas de Turing, por tanto debemos construir una única ciencia cognitiva: una ciencia formal y unificada del conocimiento, en que los detalles materiales del sistema (p. ej., si es artificial o biológico, si está hecho de neuronas o de latas de cerveza) son irrelevantes.

Rivière señala con toda la razón que esta visión unificadora nunca ha gustado a los psicólogos ya que «los psicólogos no queremos hacer cualquier clase de ciencia cognitiva, sino específicamente psicología cognitiva...» (p. 58). También está Rivière en lo cierto al señalar las debilidades psicológicas que conlleva la identificación entre la mente y la máquina de Turing. Así la máquina universal de Turing es una idealización que no se adecua a los recursos limitados de atención y memoria, ni a otras propiedades bien conocidas de la mente humana. Asimismo, un sistema estrictamente algorítmico como una máquina de Turing es rígido y poco adaptable y en él no parece tener cabida el aprendizaje. En cuanto a la conciencia quedaría relegada a un simple epifenómeno irrelevante desde el punto de vista computacional.

Hasta aquí nada tengo que objetar. Rivière ha analizado con rigor las consecuencias psicológicas de una concepción de la ciencia cognitiva basada en una interpretación literal de la máquina de Turing y a continuación ha señalado las insuficiencias psicológicas de esta fundamentación. Los vientos del País de Turing —parece decir Rivière— pueden ser gratos para algún filósofo interesado en una descripción formal del conocimiento, pero resultan insalubres para un psicólogo cognitivo con pretensiones de científico natural.

Naturalmente lo que uno espera, llegado este punto, son las claves alternativas para fundamentar la psicología cognitiva. En suma, el «billete de vuelta» desde el País de Turing debería suponer la adopción de nuevos postulados meta-teóricos sobre los que asentar, ya firmemente, la ciencia de la mente.

Creo que Rivière hace un intento serio para hallar esta alternativa meta-teórica pero, en mi opinión, sólo presenta argumentos parciales y sugestivos más que una verdadera alternativa a Turing. Así, Rivière parece considerar como argumentos «paradigmáticos» que debilitan la tesis de Turing: 1) el hecho de que las personas tengamos una teoría de la mente que nos permite tener expectativas y realizar inferencias en el entorno psicosocial, 2) el que los procesos mentales tengan un origen biológico, 3) el conexionismo como alternativa computacional al paradigma simbólico-computacional. Sin embargo:

La existencia de una teoría de la mente es compatible con los supuestos de Turing. Ciertamente, el que atribuyamos estados mentales a los demás, infiriendo por ejemplo sus intenciones o emociones, es un hecho de enorme importancia psicológica que no debe en modo alguno minimizarse. Sin embargo, no veo en qué altera esto el presupuesto de la máquina de Turing como modelo psicológico. De hecho Pylyshyn y Fodor, los más estrictos defensores de la identidad formal de la mente y la máquina de Turing, parecen navegar a sus anchas en las aguas de la teoría de la mente. Y ello es porque no la consideran peligrosa en absoluto ya que no atenta contra ningún supuesto fundamental del paradigma simbólico computacional. Todo consiste en reducir la «psicología popular» a proposiciones en el lenguaje de la mente o «mentales» y el resto sigue siendo igual que siempre: cómputo de símbolos mediante algoritmos específicos. Compárese esta actitud despreocupada y hasta complacida ante las nociones mentalistas del ciudadano de a pie, con la actitud beligerante de Pylyshyn contra la noción de imagen mental que por cierto también está muy arraigada en la psicología popular.

El planteamiento biólogo también es compatible con los supuestos de Turing. Tengo todas mis simpatías hacia la visión biólogo de la mente que propone Rivière, que coincide en gran medida con algunos planteamientos míos. Creo además que es un acierto relacionar la teoría de la mente con su fundamentación biólogo: los humanos procesamos información psicosocial con frecuencia, eficiencia y hasta entusiasmo porque estamos «diseñados» biológicamente para hacerlo así. Pero el origen biológico de algunas importantes funciones mentales no implica necesariamente que el cerebro no sea esencialmente una máquina de Turing.

Las investigaciones de Cosmides que cita Rivière son un buen ejemplo de cómo un planteamiento biólogo, por lo demás muy interesante y bien argumentado, es compatible con una noción «modular» del razonamiento en el que

se admite incluso la existencia de «algoritmos» específicos para la resolución de problemas de intercambio social. Estos algoritmos por muy «darwinianos» que pudieran ser en su origen no difieren gran cosa de la noción de «programas» que tan confortables resultan para la ciencia cognitiva.

El conexionismo es una alternativa potencial a la máquina de Turing, pero es todavía muy nuevo y debe someterse al escrutinio cuidadoso de la comunidad científica para comprobar sus posibilidades y límites. Sin duda el conexionismo es una alternativa real al paradigma simbólico computacional y una prueba de ello es la reacción airada de la ciencia cognitiva ante el conexionismo que se ha materializado en lo que podríamos llamar el primer debate paradigmático en psicología en los últimos 30 años. El paradigma simbólico computacional basado en el supuesto de la máquina de Turing hace bien en preocuparse, pues, como señala Rivière, los conexionistas «cambian las reglas del juego de Turing», y además lo hacen muy bien. El cerebro es en efecto una máquina de cómputo —nos dicen— pero no se parece lo más mínimo al juguete de Turing. Desaparecen los símbolos como entidades discretas en el lenguaje mental y desaparece la idea secuencial del procesamiento. En su lugar, nos encontramos con una red masiva de unidades de procesamiento de input, de output y unidades «ocultas».

Lo extraordinario son las propiedades funcionales que emergen de este sistema de procesamiento en paralelo, que semejan tanto a algunos fenómenos psicológicos conocidos. El aprendizaje, la construcción de categorías difusas, y el procesamiento de estímulos degradados, o el hallar soluciones a problemas que implican constricciones múltiples y simultáneas, son algunas fantásticas destrezas de los sistemas conexionistas. Todo esto son buenas noticias para la psicología cognitiva, pero también es cierto que la mente realiza operaciones que no son tan fáciles de emular por el conexionismo. En general, un sistema conexionista es una máquina «inductiva» que puede abstraer regularidades a partir de ensayos con estímulos relativamente variables. No hay que minimizar estas capacidades inductivas, como hace por ejemplo Fodor, ya que la mente es sin duda una buena máquina inductiva. Pero es probablemente algo más. Nuestra actividad mental tiene también características seriales: nuestra memoria, lenguaje, pensamientos y acciones poseen una delicada estructura temporal, que no puede considerarse un epifenómeno. Además, muchos procesos mentales son generativos como siempre ha destacado el paradigma simbólico-computacional. La producción del lenguaje, o el «ensamblaje» de rutinas para resolver un problema nuevo son ejemplos de ello.

Es pronto para decidir si el conexionismo podrá emular todas estas destrezas cognitivas humanas o sólo algunas. La respuesta ha de ser empírica y probablemente llevará su tiempo. De momento parece haber un gran salto entre las funciones inductivas que emulan los modelos conexionistas y las operaciones mentales generativas y de compleja trama temporal que somos capaces de ejecutar con tanta naturalidad las personas.

Mientras tanto la máquina de Turing como modelo formal de la mente sigue aparentemente en pie. Hay algunas alternativas esotéricas a Turing que no creo deban tomarse en serio. Por ejemplo, Searle propone que hay ciertos elementos no-computacionales en la mente que se relacionan con las propiedades

materiales de las neuronas. Más recientemente Penrose en su libro *La nueva mente del emperador*, propone que el pensamiento creativo y la conciencia no son funciones computables, no pueden ejecutarse en una máquina de Turing. Por el contrario, se relacionan con ciertas propiedades cuánticas de alto orden que existirían en la materia. Tanto Searle como Penrose coinciden en negar el principio de equivalencia funcional de los sistemas de cómputo, ya que ciertos sistemas físicos, como el cerebro, añadirían un plus cualitativo a la cognición que no podrían reducirse al mecanicismo computacional. Lamentablemente, ese plus neuronal (Searle) o cuántico (Penrose) se nos antoja demasiado misterioso. Se parece a los residuos vitalistas que poblaron la biología durante el siglo pasado.

Quizá la mejor respuesta ante el desafío de Turing es no responder a él. No hay que derribar la máquina de Turing, basta simplemente con dejar de prestarle una atención obsesiva. El problema, en mi opinión, no es si la máquina de Turing, y su prima hermana la metáfora del ordenador, son descripciones formales adecuadas de los sistemas de cómputo. Sin duda lo son. Para la psicología cognitiva la cuestión es más bien si ello tiene alguna relevancia intelectual, ¿nos permite el postulado de Turing avanzar en el conocimiento del sistema cognitivo humano? Mi respuesta a esta segunda cuestión quiero que sea clara: NO, la máquina de Turing (y la metáfora del ordenador) como fundamenteo teórico de la psicología cognitiva es inadecuada.

El paradigma simbólico-computacional ha definido a priori qué es el «cómputo» y ello impone unas estrechas restricciones a las ciencias cognitivas, incluida la psicología. Sólo sería legítimo estudiar aquellos procesos que tienen un sustrato simbólico y sintáctico. Pero, ¿qué hacer con las funciones de la conciencia, el carácter semántico o referencial de las representaciones, o la posibilidad de que existan algunas representaciones analógicas? Estas líneas de investigación las considera el paradigma simbólico-computacional como sospechosas, irrelevantes o equivocadas.

Creo que el concepto de computación universal derivado de la máquina de Turing es un buen ejercicio teórico, pero tiene una aplicación limitada en una ciencia natural que debe descubrir y hacer explícitas las funciones cognitivas de ciertos sistemas biológicos de procesamiento. La investigación en psicología cognitiva es una tarea abierta, hasta tal punto que quizá haya que reformular la propia noción de computación como consecuencia de nuestros hallazgos. Eso es precisamente lo que parece estar haciendo de modo esperanzador el conexionismo. Como asegura Rivière, el conexionismo está sustituyendo la «metáfora clásica de Turing, la metáfora del ordenador, por la metáfora del cerebro» (p. 223).

Me inclino a creer que el conexionismo con sus unidades «sub-simbólicas» de input, de output y ocultas, sus procesos de activación e inhibición, su regla delta, etc., es efectivamente la gran alternativa a Turing. En cualquier caso también el conexionismo deberá acomodarse a nuestro conocimiento empírico del sistema cognitivo. No basta con una descripción «universal» de los principios de microcomputación en redes quasi-neurales. El conexionismo deberá instanciarse en modelos psicológicos que se acomoden a nuestro conocimiento funcional del sistema cognitivo. Dicho de otro modo, el conexionismo debe llenarse de contenido psicológico.

El cerebro es con gran probabilidad una máquina conexionista, pero sus características funcionales son todavía muy enigmáticas. En este sentido, algunas de las dificultades del paradigma simbólico-computacional quizá persistan con el planteamiento conexionista. No está claro, por ejemplo, que las funciones de regulación y evaluación de la conciencia se puedan comprender o modelar en una red conexionista. Por otra parte, seguirá siendo necesaria durante mucho tiempo una descripción de alto orden de las funciones de la mente, en términos de «módulos», «mecanismos», «memoria operativa», etc.

Objetos con mente es una obra incompleta pero sólo en la medida en que la actual ciencia cognitiva lo es. Sólo podemos intentar comprender el pasado y el presente y, haciendo un ejercicio de perspicacia, asomarnos algo al futuro. Todo ello lo hace Ángel Rivière con gran maestría, ayudándonos a comprender mejor lo que nos traemos entre manos —o en mente— los que pretendemos hacer ciencia cognitiva.