

REPRESENTACIÓN E INTERPRETACIÓN

Pascual Martínez Freire. Universidad de Málaga

Resumen: Las principales propuestas de este artículo son las siguientes : 1) la substitución de la familia de términos en torno a «intencionalidad» por la familia de términos en torno a «representacionalidad», 2) la propuesta de una noción conveniente de representación, 3) la defensa de la tesis «no existe representación sin interpretación», y 4) la tentativa de formular algunos criterios acerca de la atribución de representaciones.

Summary: The main claims in this paper are the followings: 1) the substitution of the family of terms around «intentionality» by the family of terms around «representationality», 2) the proposal of a convenient notion of representation, 3) the defense of the thesis «no representation without interpretation», and 4) the attempt of formulating some criteria about the attribution of representations.

I

La filosofía de la mente, que conoce en nuestros días un periodo de esplendor y prestigio (tanto en el marco de la investigación como en el marco de la docencia), tiene sin duda raíces antiguas, que se remontan a los más importantes filósofos antiguos, Platón y Aristóteles. Sin embargo, su consolidación como disciplina filosófica es reciente, pudiendo situarse cronológicamente en torno a 1950. En efecto, es por entonces cuando se hace cada vez más usual hablar de Filosofía de la Mente, básicamente por la influencia de la obra de Gilbert Ryle *The Concept of Mind* (1949) y de la obra de Ludwig Wittgenstein *Philosophische Untersuchungen* (1953), publicada póstumamente pero cuyas ideas ya eran conocidas antes de su publicación.

Ahora bien, la filosofía de la mente (o lo que más exactamente creo que debería llamarse «nueva filosofía de la mente») adquiere nueva fuerza y características a partir de 1977¹ con el desarrollo, ordenado y constante, de las ciencias cognitivas. En efecto, en torno a esta fecha se consolida un campo de investigación interdisciplinar (en el que están interesadas básicamente la lógica, la informática, la neurociencia, la lin-

güística, la inteligencia artificial y la psicología) cuyo tema central es el estudio del fenómeno de la cognición tanto en seres humanos como en máquinas y animales.

En el ambiente intelectual de las ciencias cognitivas, la nueva filosofía de la mente tiene, a mi entender, tres tareas básicas. En primer lugar, debe ser la metaciencia cognitiva, es decir, abordar los problemas generales de las ciencias cognitivas así como la evaluación y comparación de sus teorías, siendo su asunto central la discusión de la naturaleza de la mente. En segundo lugar, la filosofía de la mente constituye la nueva y más eficaz teoría del conocimiento, atenta a los descubrimientos en psicología empírica, neurociencia e inteligencia artificial, desplazando así a la vieja (o metafísica) teoría del conocimiento. Y finalmente, la filosofía de la mente debe proporcionar fundamentos a las ciencias cognitivas a través de la semántica filosófica, a la que preferimos denominar «teoría de la representación». Evidentemente estas tres tareas están relacionadas, ya que son competencia de una misma disciplina, y los tres temas (mente, conocimiento y representaciones) son las preocupaciones centrales de la filosofía de la mente.

Dentro del contexto de la filosofía de la mente como teoría de la representación se sitúan las siguientes consideraciones acerca de los caracteres fundamentales de la representación.

II

Jerry Fodor² se ha planteado que la intencionalidad podría ser o bien algo molesto o bien la clave para lo mental. El propio Fodor defiende la segunda opción, pero personalmente entiendo que explicar los procesos mentales usando la familia de términos en torno a «intencionalidad» es verdaderamente algo molesto.

Es sobradamente conocido que Brentano³ intentó caracterizar los fenómenos mentales, en cuanto opuestos a los fenómenos físicos, por medio de la intencionalidad. Sin embargo, Spiegelberg ha mostrado que «la inexistencia mental, no la referencia a algo como objeto, constituya para él [Brentano] la esencia de la intencionalidad»⁴, entendiéndose a su vez por «inexistencia mental» la inmanencia respecto de la mente humana del objeto de experiencia. Así pues, para Brentano la palabra «intencional» era sinónima de «inmanente».

Dejando de lado consideraciones históricas, por muy ilustres que sean, la noción de intencionalidad tal como se usa en la literatura contem-

poránea resulta molesta, ya que dificulta una adecuada comprensión de la naturaleza de la mente. En efecto, la noción de intencionalidad es frecuentemente demasiado vaga, es a veces confusa y, finalmente, se usa para reclamar un privilegio para la mente humana.

En cuanto a lo primero, a menudo se define la intencionalidad como la propiedad de los procesos mentales de referirse a algo distinto de ellos mismos, de versar sobre algo, de dirigirse a algo, etc., de tal manera que por intencionalidad se entiende «direccionalidad» o «acerquidad» (*aboutness*). Sin embargo, esta concepción presupone una «flecha misteriosa» que iría desde los procesos mentales hacia objetos indefinidos. Justamente veremos cómo una noción adecuada de representación sustituye tal «flecha misteriosa» por el mecanismo real de la interpretación.

Respecto de lo segundo, a veces no queda claro si la intencionalidad como propiedad de la mente se dice en concreto de los procesos mentales, o bien de los contenidos de tales procesos, o bien de los objetos de esos procesos, o bien de todas esas entidades (procesos, contenidos y objetos) por igual. En efecto, se habla de la intencionalidad de los procesos mentales, pero también de contenido intencional y de objeto intencional. En consecuencia, parece necesario o la introducción de «sutilezas escolásticas» o el recurso a la navaja de Ockham. De momento diré que todo proceso mental es una representación, que es redundante distinguir entre proceso mental y contenido mental y, asimismo, que el objeto de la representación es relativo a la interpretación.

Finalmente, la noción de intencionalidad ha sido empleada para distinguir los seres humanos de otras criaturas, suponiendo la superioridad del género humano precisamente por la posesión de esa propiedad que no sería compartida. La intencionalidad sería no sólo una característica de los procesos mentales sino también exclusiva de la mente humana. Con ello se excluye la existencia de mentes mecánicas (o máquinas con mente) así como la existencia de mentes animales. Sin embargo, hay evidencia empírica de que ciertas máquinas y muchos animales poseen mente, entendida de modo adecuado.

En contraste con esas ideas, sostengo que la «flecha misteriosa» desde los procesos mentales (presupuesta por la intencionalidad) puede abandonarse si usamos la noción de representacionalidad, o capacidad para representar, en vez de la noción de intencionalidad y al mismo tiempo nos damos cuenta de que cualquier representación necesita una interpretación, que es a su vez una representación de segundo orden. Pero en primer lugar debemos definir la noción de representación y posteriormente

introduciremos las distinciones necesarias para resolver los problemas anejos.

III

Parafraseando a Charles Morris⁵, podemos decir que R es una representación de un objeto O para un intérprete I en la medida en que I toma cuenta de O en virtud de la manifestación de R. Por ejemplo, una línea trazada en la arena por una hormiga es una representación de Winston Churchill⁶ para mí como intérprete en la medida en que tomo cuenta de Churchill en virtud de la línea trazada. Otro ejemplo: mi creencia de que Málaga tiene una catedral renacentista es una representación del hecho de que Málaga tiene una catedral renacentista para mí como intérprete en el caso de que tome cuenta de que Málaga tiene una catedral renacentista en virtud de la presencia en mí de tal creencia; o también, mi creencia de que Málaga tiene una catedral renacentista es una representación del hecho de que Málaga tiene una catedral renacentista para tí como intérprete en el caso de que tomes cuenta de que Málaga tiene una catedral renacentista en virtud de la manifestación de mi creencia (vgr. mi preferencia «Málaga tiene una catedral renacentista»).

Debemos observar que toda representación es portadora de información en el sentido ordinario⁷ de que una representación R es capaz de proporcionar conocimiento acerca del objeto O. Es por ello que las representaciones son elementos básicos del conocimiento y, por ende, de la teoría del conocimiento. También debe repararse en que el objeto O no es necesariamente un objeto real, sino que puede ser un objeto ficticio; podemos tener una representación de la mesa sobre la que escribimos, pero también una representación de Santa Claus.

Por otro lado, una representación no es necesariamente lingüística (ni mucho menos necesariamente de tipo declarativo). Si existe un intérprete, una sentencia en castellano, una proposición en mentalés (el lenguaje del pensamiento postulado por Fodor), pero también una foto o un deseo⁸, son representaciones. Así pues el punto de vista que adoptamos es muy general. Asimismo las representaciones pueden ser externas o internas según que se dé o no la posibilidad de una observación directa intersubjetiva (esto es, compartida por varios sujetos) de ellas; así una fotografía de Greta Garbo es una representación externa, pero mi deseo de nadar es una representación interna.

IV

Ahora bien, ¿qué es un intérprete? De modo sucinto, un intérprete es algo o alguien que da sentido a una representación. A su vez, sentido, de manera general, es una función o un papel, pero no necesariamente (aunque sí posiblemente) un significado preciso ni tampoco una referencia definida. Es decir, con el término «sentido» se indica cualquier valor significativo (semántico) de una representación, tanto si se trata de un papel concreto y particular (incluso no generalizable) como si se trata de significados o referencias, que son generales y sociales. Por ejemplo, dada la palabra «peligro», el intérprete de esta representación puede atribuirle un «simple sentido» concreto, y obrar en consecuencia, aunque nunca haya elaborado un significado preciso (compartido y social) de tal palabra. Semánticamente el «simple sentido» es previo, mientras que significados y referencias son posibles y eventuales elaboraciones.

Dicho de otro modo, hay varios grados de tomar cuenta del objeto O por medio de la manifestación de la representación R. En un primer grado el intérprete sencillamente toma cuenta de la existencia del objeto a través de su representación; en un segundo grado el intérprete puede dar un papel o función (simple sentido) al objeto manifestado por la representación; pero en grados superiores el intérprete puede dar un significado, preciso y siempre social, al objeto, o bien puede darle referencia, esto es, puede situar el objeto en el mundo real. Así pues, la referencia es una interpretación de una representación en virtud de la cual el objeto manifestado en la representación es situado en el mundo real.

Esto último nos lleva a plantearnos, aunque sea brevemente, el tema de la ontología de la representación. Al respecto sostendremos tres tesis: a) *hay* un mundo externo a la representación, b) *existe* el mundo de la experiencia, y c) las representaciones son acerca de lo que existe. Según ello, la actitud o posición básica es el realismo, pero un realismo que se limita a creer (con una especie de «fe adaptativa») en que hay o se da un mundo independiente del conocimiento, aunque podamos esgrimir argumentos científicos y filosóficos que hacen plausible tal creencia. Asimismo se postula una relación causal entre el mundo que hay y el mundo de la experiencia, que es el mundo que existe para las mentes en el sentido de que la única y posible experiencia de realidades se da en las representaciones mentales. Pero, finalmente, estas representaciones mentales, aunque representan realidades distintas de ellas mismas y pueden ser causadas desde el mundo externo, adquieren sentido (en los términos

indicados en los párrafos anteriores) merced a las interpretaciones correspondientes.

V

Según nuestra definición de representación, ésta es relativa a un intérprete, de modo que podemos decir que no existe representación sin interpretación. En otras palabras, la capacidad para representar (representacionalidad) está potencialmente en la representación pero es actualizada por el intérprete. Por tanto, la representacionalidad es siempre algo extrínseco a la representación, no existiendo representacionalidad intrínseca. Obviamente esto no significa una constante acción de interpretación; cuando la interpretación está cumplida puede ser registrada en una memoria y luego puede ser recuperada. Así pues es posible preservar la interpretación.

En relación con las representaciones mismas debemos defender al mismo tiempo el realismo y el convencionalismo. Las representaciones son productos reales pero no son hechos absolutos. El mundo de la experiencia es un conjunto de cosas incluyendo las cosas que son también representaciones, de tal manera que el mundo es un conjunto de cosas y/o representaciones. Sin embargo, las representaciones (externas o internas, lingüísticas o no lingüísticas) son tales en conexión con un intérprete. Ello implica cierta convencionalidad, ya que son posibles distintas interpretaciones de una misma representación, pero esta convencionalidad no es puramente arbitraria, puesto que la interpretación debe dar un sentido a la representación y tal sentido debe desempeñar con éxito una función o un papel.

Por otra parte, una interpretación es una representación de segundo orden. Ello quiere decir que una interpretación es también una representación, pero que versa sobre otra representación, dándole sentido en el modo general de hablar que antes hemos indicado. Podría pensarse entonces que la interpretación entraría en un proceso infinito (es decir, interpretación de interpretación, etc.); sin embargo, la interpretación tiene un carácter terminativo, de modo que una vez cumplida queda al mismo tiempo completa, aunque podamos realizar *otra* interpretación.

Mientras una representación puede ser extramental, toda interpretación es un proceso mental. Un computador interpreta las instrucciones que recibe, un gato interpreta las representaciones de los objetos en torno a él, un hombre interpreta sus creencias, recuerdos, percepciones y voli-

ciones. Justamente *la capacidad de interpretar representaciones es el rasgo básico de la mente*. Y así podemos hablar de la mente en (ciertas) máquinas, en animales y en los seres humanos. Sin embargo, la mente humana es diferente, en complejidad y en recursos, de la mente animal o de la mente de un computador. (En particular, la mente humana posee procesos mentales, tales como los de autoconciencia, volición libre, formación de ideales y de proyecto vital, que son espirituales⁹).

El intérprete de una representación puede ser el propio poseedor de ella o/y un intérprete ajeno a la representación. Cuando interpreto la línea trazada en la arena por una hormiga, la interpretación es ajena a la representación. Cuando interpreto mi creencia de que Málaga tiene una catedral renacentista, el intérprete es el propio poseedor de la interpretación. Si tú y yo intentamos captar el significado de la palabra «democracia», constituímos una red de intérpretes. En general, dado el carácter social de los significados, éstos son interpretaciones establecidas por una red de intérpretes.

Ahora bien, si la interpretación es tomar cuenta de un objeto mediante la manifestación de su representación, parece que toda interpretación es una representación consciente. Pero Daniel Dennett sostiene que «en la medida en que la ciencia cognitiva está interesada, los fenómenos importantes son las representaciones mentales inconscientes explícitas»¹⁰. Por tanto la interpretación (si es consciente) sería irrelevante o de escasa importancia para la ciencia cognitiva. Sin embargo, la situación puede clarificarse fácilmente.

Las representaciones internas, tales como los procesos mentales humanos, los procesos mentales animales y los estados internos de un computador, están incorporados en circuitos neuronales o en circuitos eléctricos, y son muy frecuentemente inconscientes. La manipulación de representaciones inconscientes es en verdad esencial en ciencia cognitiva, pues para ésta el pensamiento en general supone la manipulación de representaciones que no siempre son conscientes. Pero las representaciones son relevantes para la cognición en la medida en que han recibido o recibirán una interpretación consciente. Por ello la interpretación es verdaderamente relevante para la ciencia cognitiva. Dicho de otro modo, aunque buena parte de las representaciones sean inconscientes (permaneciendo entonces como representaciones potenciales), e incluso su manipulación sea también inconsciente, llegan a constituir conocimiento (actualizando su representacionalidad) cuando reciben interpretación consciente. Justamente la necesidad de interpretación para las representaciones excluye identificar explica-

ción física y explicación psicológica, evitando así la reducción de la psicología a la neurociencia.

VI

He dicho anteriormente que el mundo no es sólo un montón de cosas sino también un conjunto de representaciones. El número y tipos de representaciones es realmente muy vasto. Sin embargo, estamos interesados de modo especial en lo que podemos calificar de «representaciones por excelencia», esto es, aquéllas producidas conjuntamente con su interpretación. Por un lado, una foto o una sentencia escrita en castellano son representaciones, pero su interpretación es siempre ajena a ellas. Por otro lado, una creencia o un deseo son representaciones por excelencia, ya que su interpretación puede realizarse por el propio poseedor de ellas. Estableceremos algunos criterios acerca de la atribución de representaciones por excelencia.

Tales criterios de atribución de representaciones resultan de la noción de representación antes establecida así como de las observaciones apuntadas acerca de la interpretación. En general existen básicamente cuatro dimensiones o niveles de esa atribución.

El primer criterio es la existencia de una estructura perceptiva del ambiente. Según ello, si un ser dispone de un mecanismo capaz de advertir la presencia de un objeto y de identificarlo en algún grado, podemos suponer que tal ser tiene representaciones acerca del objeto. Este caso simple está muy extendido ya que aparece en animales, ciertas máquinas y en los seres humanos. Sin embargo, algunos animales elementales y artefactos, por ejemplo almejas y termostatos, tienen sólo este tipo de representaciones; no poseen una estructura causal compleja ni tampoco conducta auténtica.

El segundo criterio de atribución de representaciones es la exhibición de poderes causales que producen salidas para diferentes entradas de información. De acuerdo con ello, si un ser tiene una estructura causal suficientemente compleja de modo que, en una manera determinista y mecánica, puede producir salidas correspondientes a entradas de información, entonces podemos suponer que tal ser tiene representaciones acerca de los objetos relacionados con las entradas y con las salidas. Este caso es compartido por animales, computadores y seres humanos. Sin embargo, algunos animales y computadores, por ejemplo ranas y calculadoras, tienen sólo este tipo de representaciones; no poseen conducta auténtica.

El tercer criterio para atribuir representaciones es la posesión de patrones de conducta. Según ello, si un ser despliega una conducta, incluyendo no sólo regularidades entre estímulos y respuestas sino también respuestas imprevistas, y además abarca no sólo resultados mecánicos sino también soluciones inventivas, entonces podemos suponer que tal ser tiene representaciones acerca de los objetos relativos a los estímulos y respuestas. Este es el caso de los animales superiores (por ejemplo gatos y perros), computadores inteligentes y seres humanos.

Pero existe un último criterio, a saber la proyectabilidad de estados representativos. Si un ser, que sabemos que es capaz de desarrollar conducta, está ocultando o fingiendo su conducta (y ello puede inferirse de modo fiable), entonces podemos suponer que este ser tiene representaciones acerca de los objetos concernientes a los estímulos. Este caso es bien conocido en los seres humanos, pero quizás también se dé en los animales superiores y no resulta impensable en los computadores inteligentes. Pero en este caso la atribución de representaciones es proyectada en vista de que una conducta razonablemente esperada no resulta manifiesta o apropiadamente manifiesta.

NOTAS

1. En ese año apareció la revista *Cognitive Science*.
2. Fodor (1986, p. 308).
3. Brentano (1924, libro II, cap. 1, sección 5).
4. Spiegelberg (1976, p. 121).
5. Morris da una definición de signo que es paralela a nuestra definición de representación. Morris (1938, p. 4).
6. Compárese Putnam (1981, pp. 1-2). Putnam sostiene que tal línea no es una representación.
7. Dretske (1981, pp. 44-45).
8. Mi deseo de nadar es la representación de mi eventual nadar en un futuro, y como intérprete tomo cuenta de mi futuro nadar en virtud de la presencia de este deseo.
9. Martínez Freire (1992).
10. Dennett (1987, p. 218).

REFERENCIAS

Brentano, F. 1924: *Psychologie vom empirischen Standpunkt*. 2ª ed. (O. Kraus). Leipzig: Felix Meiner.

Dennett, D. 1987: *The Intentional Stance*. Cambridge, Mass.: MIT Press.

Dretske, F. 1981: *Knowledge and the Flow of Information*. Oxford: Basil Blackwell.

Fodor, J. 1986: «Information and Association». *Notre Dame Journal of Formal Logic*, 27, pp. 307-323.

Martínez Freire, P. 1992: «Procesos mentales y cognitivismo». *Revista de Filosofía* (en prensa).

Morris, Ch. 1938: *Foundations of the Theory of Signs*. Chicago: University of Chicago Press.

Putnam, H. 1981: *Reason, Truth and History*. Cambridge: Cambridge University Press.

Spiegelberg, H. 1976: «Intention and Intentionality in the Scholastics, Brentano and Husserl». In *The Philosophy of Brentano*, L.L. McAlister, ed., London: Duckworth.