

En torno a la intencionalidad

JOSÉ HIERRO S. PESCADOR
(Universidad Autónoma de Madrid)

La atención que los problemas de la mente han despertado en los últimos decenios ha contribuido muy claramente a que el tema de la intencionalidad pase a primer término.

En la filosofía moderna, el tema fue introducido por Franz Brentano en la segunda mitad del siglo XIX, en su obra *Psychologie vom empirischen Standpunkt* (1874). El tema había sido planteado —aunque en un contexto teórico distinto— por San Agustín y por filósofos medievales como Santo Tomás, Duns Escoto y Avicena. Pero si se quiere buscar el origen de este tema, hay que señalar, como para tantas otras cuestiones, a Aristóteles. El lugar exacto está en su *Metafísica*, cuando escribe:

Pues pensable significa que hay un pensamiento que lo piensa; pero el pensamiento no es relativo a aquello de lo que es pensamiento (pues se habría dicho dos veces la misma cosa). Y de modo semejante, la visión es visión de algo, pero no de aquello de lo cual es visión (aunque sea verdadero decir esto), sino que es relativa al color o a alguna otra cosa semejante. Describirla del otro modo, como la visión del objeto de la visión, sería decir dos veces la misma cosa (*Metafísica*, libro V, 1021a31-1021b4).

De este fragmento hay varias interpretaciones. Según la de Bonitz (*Aristotelis Metaphysica. Commentarius*, pp. 261 ss.), lo pensable se pone en el género relativo en cuanto que el pensamiento se refiere a ello, pero esto impide definir el pensamiento como lo que se refiere a lo pensable, pues entonces se repetiría lo mismo. De modo similar, lo visto es el objeto de la visión, pero

no se puede definir la visión como lo que se refiere a lo visto, pues incurriríamos en una repetición como la anterior. Según la interpretación de Schwieger (*Die Metaphysik des Aristoteles*, vol. III, p. 231), el concepto de pensable significa que hay un pensamiento de ello, y el pensamiento se refiere a un objeto pensable, pero no al sujeto pensante. Para Reale, lo que Aristóteles quiere decir está claro, teniendo en cuenta que su preocupación real es ésta: toda forma de conocimiento es relativa al ser y a la realidad, pero no viceversa (*Aristotele. La Metafísica*, vol. I, p. 465). De acuerdo con W. D. Ross (*Aristotle's Metaphysics*, vol. I, pp. 330 s.), la idea de Aristóteles es que lo pensado (o lo visto) debe tener una naturaleza propia aparte de ser objeto del pensamiento (o de la visión), y por ello concluye Ross:

«En el fondo del pensamiento de Aristóteles, aunque no expresado de forma muy satisfactoria, está la convicción de que el conocimiento y la percepción son relativos a la realidad de una manera en que la realidad no es relativa a ellos» (*op. cit.*, p. 331; mi traducción).

El carácter intencional, tanto del pensamiento como de la percepción, puede descubrirse en la idea de que ambos son relativos a la realidad, aunque la realidad no sea relativa a ellos.

La doctrina de la intencionalidad recibió notable desarrollo en Avicena, quien no sólo distinguió entre intencionalidad sensible e intencionalidad intelectual (distinción que ya está presente en el párrafo anterior de Aristóteles), sino también entre intenciones primeras y segundas, así como entre intenciones formales y objetivas. La doctrina clásica de los escolásticos sobre la intencionalidad se encuentra ya en Avicena —como han sostenido los especialistas (*vid.* Miguel Cruz Hernández, *Francisco Brentano*, cap. X)—, y de él dependen en este tema Tomás de Aquino y Duns Escoto, entre otros muchos. El primero de ellos toma como base el carácter intencional de todo acto de conocimiento, tanto inteligible como sensible, y de la voluntad. Duns Escoto, por su parte, también claramente influido por Avicena, analizó con detalle la distinción entre primeras y segundas intenciones, estudiando la doctrina intencional de las categorías y de los universales. Pero la doctrina escolástica de la intencionalidad declina a partir del siglo xiv. La teoría tomista recibe críticas importantes de Durando de Saint Pourçain, y finalmente Ockham prescinde del ser intencional y liquida la teoría escotista. La doctrina de la intencionalidad reaparece, sin embargo, en Suárez, quien, partiendo de Escoto, defiende la analogía de la intencionalidad, y —anticipando el sentido moderno de este concepto— atribuye la intencionalidad al conjunto de las facultades humanas, y no a una facultad por separado. (Para más detalles puede verse Cruz Hernández, *loc. cit.*)

El concepto de intencionalidad fue para Brentano la base sobre la que fundar la psicología. La intencionalidad le sirvió para distinguir los fenómenos psíquicos de los físicos:

«Todo fenómeno psíquico se caracteriza por lo que los escolásticos de la Edad Media llamaban inexistencia intencional (o también mental) del objeto, y que nosotros, aunque no con expresiones del todo inequívocas, podríamos llamar relación a un contenido, dirección a un objeto (con lo cual no hay que entender que sea una realidad) o inmanente objetividad» (*op. cit.*, primera parte, libro 2, c. I, sec. 5).

El objeto puede no existir realmente, pero en todo caso se da como contenido del acto psíquico. Cuando se piensa, se piensa algo, cuando se desea, se desea algo, y cuando se teme, se teme algo. Brentano (*op. cit.*) clasifica los fenómenos psíquicos en representaciones, juicios y fenómenos de amor y odio. Habla de representación siempre que algo se nos aparece, y las representaciones incluyen tanto las percepciones sensibles como las imaginaciones. Entiende por juicio el admitir algo (como verdadero) o rechazarlo (como falso). Como fenómenos de amor-odio considera las emociones en general y los estados de interés. De todos ellos, los fenómenos básicos son las representaciones, porque un fenómeno de cualquier otra clase implica siempre una representación, ya que el concepto de representación está tomado en tan amplio sentido que recuerda el concepto de idea en Descartes.

La fecundidad de Brentano se manifestó primeramente en la filosofía fenomenológica. Aunque también ejerció influencia en la filosofía analítica, su presencia explícita dentro de ésta ha adquirido más relieve en tiempos recientes. Ya en 1957, Chisholm, en su libro *Perceiving: A Philosophical Study*, explicaba el concepto de aceptar (o asumir) en relación con el concepto de intencionalidad tal y como Brentano lo había desarrollado en 1874. Poniéndolo en términos del lenguaje que usamos, Chisholm propone los siguientes criterios, por medio de los cuales distinguir cuándo se usa una oración de modo intencional (*op. cit.*, cap. 11).

En primer lugar, propone considerar intencional una oración declarativa simple cuando en ésta se use una expresión sustantivada (nombre o descripción) de tal modo que ni la oración ni su contradictoria impliquen la existencia o no-existencia de algo a lo que la expresión sustantivada se aplique con verdad. Con este criterio, la oración «Diógenes buscaba un hombre justo» es intencional, pues ni ella ni su negación implican que exista algún hombre justo, pero en cambio, no es intencional la oración «Diógenes estaba sentado en su bañera», pues implica que existía una bañera en la que Diógenes se sentaba.

En segundo lugar, propone considerar intencional una oración no compuesta que contenga una cláusula proposicional cuando ni la oración ni su contradictoria impliquen que la cláusula proposicional es verdadera o que es falsa. Según este criterio, cualquier oración de creencia es intencional, pues una oración de la forma «N cree que *p*» no implica ni que *p* sea verdadera ni que sea falsa, como tampoco lo implica la oración «N no cree que *p*».

En tercer lugar, si tenemos dos nombres o descripciones que designan el mismo objeto y una oración *E* que afirma la identidad de los objetos denotados por ambos nombres, entonces, si tenemos dos oraciones *A* y *B*, tales que cada una de ellas usa uno de aquellos nombres o descripciones, Chisholm propone considerar intencional la oración *A* en el caso de que la conjunción de *A* y *E* no implique *B*. Esto ocurre con las oraciones que podemos llamar cognitivas, como son las que, en general, usan verbos como *conocer*, *saber* y *ver*. En este sentido, la oración (A) «Sé que estoy hablando con el catedrático de Física Teórica de la Universidad X», es intencional porque en conjunción con la oración (E) «El catedrático de Física Teórica de la Universidad X es idéntico al Rector de la Universidad X», que es una identidad verdadera, no implica la oración (B) «Sé que estoy hablando con el Rector de la Universidad X».

Lo anterior está formulado para oraciones simples, pero Chisholm lo aplica igualmente a oraciones compuestas por medio de conectivas proposicionales, siempre que al menos una de las oraciones componentes cumpla con dichos criterios.

Reformulada la tesis de Brentano en los términos metalingüísticos de Chisholm, la tesis viene a ser que para la descripción de los fenómenos psicológicos se requiere oraciones intencionales, mientras que tal clase de oraciones no es necesaria para la descripción de fenómenos meramente físicos. A la propuesta de Chisholm cabe hacer algunas objeciones. Por ejemplo, podemos describir el comportamiento de una máquina diciendo «La calculadora sabe que cinco por siete son treinta y cinco», lo cual constituye un enunciado intencional. La respuesta de Chisholm es que, para esta clase de ejemplos, los enunciados intencionales pueden ser transformados fácilmente en otros que no lo son, de esta manera: «Si se aprieta, por este orden, los botones de la calculadora marcados con los signos» «5, X, 7, =», aparecerá en el visor de la máquina el signo «35». A la respuesta de Chisholm cabría, sin embargo, contra-objetar que se puede hacer un análisis semejante cuando se trata de una persona, y por tanto de un enunciado genuinamente psicológico. Así, si afirmamos «N sabe que cinco por siete son treinta y cinco», podríamos traducir este enunciado como «Si se pregunta a N cuál es el resultado de multiplicar 5 por 7, éste responderá que 35». Naturalmente, N puede equivocarse, o no entender la pregunta porque es persona carente de estudios, pero entonces no diríamos que sabe que cinco por siete son treinta y cinco (en caso de que se equivocara, tendríamos que comprobar que la equivocación no se repite mucho). Lo que quiero poner de manifiesto es que la producción de una respuesta determinada a continuación de ciertos estímulos no es necesariamente un criterio que nos permita distinguir entre estado intencional y estado no intencional. Otra alternativa considerada por Chisholm consiste en describir los estados psicológicos por medio de oraciones no intencionales que incluyan expresiones técnicas creadas al efecto. Por ejemplo, en lugar de

decir «N toma lo que ve por un toro», podríamos decir «El entorno perceptivo de N incluye un toro». El comentario de Chisholm a esta alternativa es que aquí hay expresiones técnicas como «entorno perceptivo» e «incluye» que, presumiblemente (*sic* en el original), no hacen falta para la descripción de fenómenos no-psicológicos, y que, a menos que reformulemos el ejemplo por medio de una oración no-intencional carente de tales expresiones, la versión metalingüística de la tesis de Brentano ofrecida por Chisholm será igualmente aplicable al ejemplo.

Da la impresión de que esas expresiones, que Chisholm considera técnicas llevan el peso de la intencionalidad, y por tanto hacen posible la aplicación del principio de Brentano. La reformulación del principio de intencionalidad por Chisholm presupone un lenguaje verbal, y en esta medida hace ociosas algunas preguntas que, en principio, parecen importantes para la tesis de Brentano. Por ejemplo: ¿tienen intencionalidad los animales? Desde las arañas cuando tejen su tela, y las hormigas cuando construyen su casa, hasta el perro cuando ladra saludando a su amo, hay una amplia gama de comportamientos que, pasando por el pájaro que construye su nido, y por las abejas que danzan para comunicarse en qué dirección hay flores, resultan más claramente intencionales a medida que nos acercamos a los animales con un cerebro más desarrollado.

Con posterioridad a Chisholm, el tema de la intencionalidad ha interesado a diversos pensadores, entre los que destaca Dennett. La teoría de la intencionalidad que éste ha venido desarrollando apareció en 1969 en *Content and Consciousness*, siguió en el artículo «Intentional Systems» (1971), y en varios más que fueron recogidos en el libro *Brainstorms* (1978), continuando con otros muchos, algunos de los cuales fueron reimpresos en su libro *The Intentional Stance* (1987), obra de profunda influencia en los últimos años. En el segundo capítulo de esta última obra, examinando el fenómeno de la creencia, Dennett mantiene que la creencia sólo puede ser percibida si se adopta una cierta estrategia predictiva que él llama «estrategia intencional», y que es la que posee quien se encuentra en la que él denomina «actitud intencional» (*intentional stance*). La estrategia intencional consiste en tratar el objeto cuya conducta se quiere predecir como un agente racional, con creencias, deseos y otros estados mentales que presenten lo que Brentano y otros han venido llamando intencionalidad (Dennett, *op. cit.*, p. 15). La tesis de Dennett es que cualquier sistema (o cualquier objeto) cuyo comportamiento sea bien predicho por medio de tal estrategia es un sujeto de creencias (*believer*), o dicho con otras palabras, un sistema intencional. A la hora de predecir la conducta de una persona hay otras estrategias a las que recurrir. Por ejemplo, la estrategia astrológica, que toma como datos la fecha y hora del nacimiento de la persona. Tal estrategia es deplorable —para Dennett— a causa de «tan buenas razones como tenemos para creer que no funciona» (p. 16). Cuando las predicciones astrológicas resultan verdaderas es por causalidad

—dice Dennett— o porque las predicciones eran tan vagas o tan ambiguas que cualquier curso de los acontecimientos las hace verdaderas (*ibid.*). Sin embargo, esta estrategia —como reconoce Dennett— es popular. ¿Por qué? No es el caso entrar ahora en una discusión de los méritos y deméritos de las predicciones astrológicas, pero debe notarse que atribuir al azar la verdad de una predicción de esta clase presupone que se posee un criterio más seguro y fiable respecto a las condiciones en las que aceptar una tal predicción. Las diferencias entre quienes creen y quienes no creen en predicciones astrológicas no brotan —en mi opinión— del grado de acierto de las mismas, sino del marco teórico en el que ellas tienen su lugar. Hay quienes prefieren aceptar un marco en el que lo que nos ocurra esté ya determinado por el día y la hora del nacimiento, y hay quienes prefieren pensar que somos libres para guiar el curso de nuestra vida en una u otra dirección (aun cuando no dejen de reconocer que en alguna medida estaremos condicionados por la carga genética recibida de nuestros padres).

Otra estrategia, que Dennett menciona, es la estrategia física (*ibid.*), con la cual se predice el comportamiento de un sistema sobre la base de su constitución física, y es la estrategia característica de las ciencias físicas. Más efectivo es, a veces, cambiar a lo que llama Dennett «la actitud del diseño» (*design stance*). Consiste en predecir el comportamiento de un objeto sobre el supuesto de que se comportará tal y como ha sido diseñado para comportarse (p. 17). Con esta estrategia predecimos el comportamiento de un ordenador o el de un reloj despertador (ejemplos de Dennett). No es necesario conocer la estructura física de estos aparatos para predecir con éxito su comportamiento; basta conocer en qué forma están diseñados, esto es, para qué han de servir. Esta estrategia no se limita a artefactos; también vale —según Dennett— para sistemas biológicos como animales, plantas y ciertos órganos de unos y otras, como por ejemplo, los riñones, el corazón y los pistilos (*ibid.*). Se advertirá que la estrategia del diseño no presupone que el objeto haya sido fabricado ni que alguien le haya dado al objeto su diseño. El diseño puede tener un origen meramente natural.

La estrategia que Dennett recomienda para predecir el comportamiento es la estrategia intencional. Para que esta estrategia funcione, tenemos que empezar por tratar el objeto como un agente racional. A la vista de sus propósitos y de su situación en el mundo, imaginamos qué creencias y qué deseos debería tener. Sobre esta base, podemos predecir que tal objeto actuará a la luz de sus creencias para alcanzar sus fines. Nótese que el funcionamiento de esta estrategia requiere la suposición de que el agente es racional, y por tanto que se comporta como debería comportarse en función de sus creencias, sus deseos y sus propósitos. También presupone que atribuimos al agente como creencias todas aquellas verdades que son relevantes para sus deseos y que le ha proporcionado su experiencia (p. 18), y que le atribuimos deseo de todo aquello que es bueno para él (p. 20). Ambos supuestos se reducen a

lo que llama Dennett regla fundamental, y que trivialmente se enuncia así: atribuyamos al sistema aquellas creencias y aquellos deseos que debería tener (*ibid.*). Podemos preguntarnos: que debería tener ¿para qué? Naturalmente, para conseguir sus propósitos, comportándose de una manera racional, o al menos, suficientemente racional. Este último matiz es muy importante, porque podemos partir de un ideal de racionalidad perfecta, asumiendo que el sistema acepta (cree) todas las implicaciones de sus creencias, y que no acepta ningún par de creencias que sean contradictorias entre sí; pero las implicaciones pueden ser infinitas, y esto plantearía un problema a la aplicación de nuestra estrategia. Sin embargo, para que ésta funcione basta considerar aquellas implicaciones que son relevantes para la situación en que se encuentra el sistema en ese momento desde el punto de vista de su comportamiento (p. 21).

¿Usamos, de hecho, esta estrategia intencional? Sin duda. Según Dennett, ésta es la única estrategia que todos nosotros conocemos hoy para atribuir creencias y deseos, y para predecir la conducta (*ibid.*). Recuérdese que la estrategia astrológica ha sido descartada y que la estrategia física y la estrategia del diseño resultan insuficientes. Nuestro uso de la estrategia intencional es tan habitual que fácilmente podemos olvidarlo. Pero esto no significa que tal estrategia solamente funcione entre personas. Funciona también para toda clase de animales, para las plantas, para fenómenos naturales como el rayo, y para aparatos como los ordenadores. Por lo que respecta a los animales, Dennett recuerda (p. 22) que construimos mejores trampas para atraparlos porque contamos con lo que ellos saben o creen, con lo que prefieren y con lo que desean evitar; y comenta que una almeja sólo aflojará la fuerza con la que mantiene cerrada su concha cuando esté convencida de que ha pasado el peligro que ella creía tener cerca. En cuanto a las plantas, recuerda Dennett aquella variedad de manzanos que son particularmente cautos respecto al momento en que ha llegado la primavera, que es cuando ellos quieren florecer (y puestos a atribuir intencionalidad a una planta, podríamos recordar también el caso de las plantas carnívoras, por ejemplo). Respecto a los rayos, recuerda Dennett que el rayo siempre quiere encontrar el mejor camino a tierra, y en cuanto a los ordenadores, menciona las decisiones que ha de tomar un ordenador programado para jugar al ajedrez. Sin embargo, ésta no es la única máquina para la que funciona la estrategia intencional, pues también podríamos aplicarla a un termostato que apague un radiador cuando «crea» que la habitación ha alcanzado la temperatura adecuada. Pero esto no significa que todas esas entidades sean racionales.

Después de esta amplia variedad de ejemplos de funcionamiento de la estrategia intencional, hay que preguntarse: ¿Existe alguna clase de objetos para los que no sirva la estrategia intencional? Dennett considera el caso del atril que tiene delante (p. 23). ¿Qué nos impediría aplicar a este objeto la estrategia intencional? Parece que podríamos construir este objeto como un

sistema intencional, que tal vez cree que está situado en el centro del mundo civilizado y que desea permanecer así. Sobre esta base, podemos predecir que no se moverá de su sitio, y es lo que acontece. Pero si consideramos el atril como un sistema intencional, entonces no hay objeto que no debiéramos considerar de esta forma, como reconoce Dennett (*ibid.*). ¿Por qué razón debemos rehusar la condición de sistema intencional a objetos como el atril? Porque la aplicación de la estrategia intencional no nos proporciona en este caso un poder predictivo que no poseyéramos antes. La mera estrategia del diseño ya nos permite predecir que el atril seguirá en su sitio, mientras ninguna fuerza exterior opere sobre él, y no se muestra necesario atribuirle creencias o deseos. En el caso de las personas, por el contrario, la única estrategia práctica parece ser la estrategia intencional, pues nos da un poder predictivo que no podemos obtener de otra manera. ¿Y en el caso de los animales y de los ordenadores? ¿No basta la estrategia del diseño para predecir el comportamiento de un ordenador? ¿Es realmente necesario tratarlo como sistema intencional? Recuérdese que, en la tradición de Brentano, la posición intencional implica atribuir estados mentales. Tal vez no resulte extraordinario atribuir mente a los animales, especialmente a aquellos que poseen un sistema nervioso más desarrollado y que desarrollan un comportamiento más variado y más complejo. En tal caso, se insinúa otra cuestión: ¿no habría que distinguir entre la mente humana y la mente animal, o mejor aún, entre la mente propia de los distintos niveles de organización animal? No parece forzado hablar de la mente del perro, del mono, del caballo o del delfín; pero hablar de la mente de la lombriz o de la almeja produce un movimiento de rechazo. El comportamiento de estos últimos animales da la impresión de que no posee variedad ni complejidad suficientes para justificar que se les atribuya creencias o deseos; tampoco la escasa complejidad del sistema nervioso de tales animales, como la lombriz o la almeja, para exigir que se les atribuya estados mentales. El comportamiento de ellos puede predecirse con un alto grado de certeza sin recurrir a la estrategia intencional. ¿A qué estrategia recurriremos?

Podemos imaginar una nueva estrategia que añadir a las estrategias descritas por Dennett. Puesto que se trata de seres vivos, parece razonable adoptar la actitud de predecir su comportamiento en virtud del instinto de supervivencia. Nuestra estrategia será ahora la estrategia de la supervivencia (tanto del individuo como de la especie). Ciertamente que con esta estrategia podemos predecir el comportamiento de la almeja que impide que abramos su concha, así como el de la lombriz que se mete bajo tierra cuando la hostigamos. ¿Pero no podríamos hacer estas predicciones con la estrategia del diseño? Esta estrategia parece tener como paradigma el caso de los artefactos, los cuales han sido construidos para comportarse de una manera determinada al servicio de un propósito específico. Recuérdese, sin embargo, que Dennett aplica esta estrategia a los animales y a las plantas. Podemos pensar que, en el

caso de los sistemas biológicos, el diseño viene dado por la evolución natural (véase más abajo el uso que hace Dennett de esta idea). Los sistemas biológicos son objetos que, en cuanto tales, están al servicio de un proyecto de supervivencia, con relación al cual podemos predecir su comportamiento. Pero algunos sistemas biológicos —los animales superiores— presentan además tal complejidad en su comportamiento (y en su sistema nervioso), que resulta conveniente recurrir a la estrategia intencional para predecir su conducta, porque la estrategia del diseño se torna insuficiente. Para los sistemas biológicos que no han alcanzado este nivel (como los animales inferiores y las plantas) basta la estrategia del diseño; aunque no hay que olvidar que el diseño tienen carácter natural y viene dado por la propia evolución biológica, a diferencia de los ordenadores y los termostatos, cuyo diseño le ha sido dado por el fabricante. Este es también responsable de las decisiones que tome el ordenador cuando juega al ajedrez. Atribuir estados mentales a este ordenador, o a un termostato, no parece justificado, porque la estrategia del diseño es suficiente para predecir el funcionamiento de ambos artefactos.

Es verdad que Dennett mantiene que la intencionalidad que podamos atribuir a un artefacto es derivada (p. 298), puesto que la tiene por la forma en que lo hemos construido, pero no se ve qué diferencia existe entre atribuir a los ordenadores una intencionalidad derivada y atribuirles simplemente una capacidad de comportamiento que resulta del diseño con el que han sido construidos. La estrategia del diseño es suficiente para predecir el comportamiento de un artefacto, pero no parece que lo sea para predecir la conducta de una persona. El diseño que la evolución natural ha proporcionado a las personas incluye tal variedad de respuestas a las distintas situaciones, que no es posible predecir el comportamiento de una persona en una situación determinada a menos que le atribuyamos creencias y deseos relacionados con esa situación. La diferencia fundamental entre las personas, por un lado, y los demás seres naturales (biológicos o no) junto con los artefactos, de otro lado, está en que las personas son, en gran medida, libres, porque tienen que elegir entre alternativas, y nos vemos obligados a recurrir a sus creencias y deseos cuando queremos predecir su comportamiento, porque la estrategia del diseño no basta. Para los animales más evolucionados puede ser conveniente también recurrir a esta estrategia intencional, atribuyéndoles creencias y deseos. «Sabe que es su amo» o «Quiere salir a la calle» son atribuciones de estados mentales que nos permiten predecir, en cierta medida, el comportamiento de un perro, y por ello facilitan nuestro trato con él. Pero debe notarse que esta estrategia implica la atribución de libertad en alguna medida, por pequeña que sea. Para animales con un elemental nivel de evolución —como la almeja— ni siquiera parece conveniente recurrir a la estrategia intencional. El comportamiento de una almeja es de tan escasa complejidad que toda necesidad de predicción queda satisfecha con la estrategia del diseño. Y con mayor razón aún si el sistema considerado es un artefacto. No se

ve ninguna justificación para hablar de las creencias de un termostato con respecto al radiador que ha de encender o apagar según la temperatura que tenga. El termostato ha sido fabricado para que se comporte de esta manera, y saber esto ya es suficiente para predecir su comportamiento. ¿Qué se gana atribuyéndole creencias?

A la relación que existe entre la organización de un sistema y su entorno se alude diciendo que hay una representación del entorno en la organización del sistema. A esta idea responde la tesis de Dennett, según la cual, cuando un objeto es tal que con él funciona la estrategia intencional, hemos de interpretar algunos de sus estados internos como representaciones internas (p. 32). Por esta razón, Dennett no tiene ningún inconveniente en aceptar que un termostato tiene una representación del mundo, aunque una representación mínima. Robots más complejos tendrán una representación más completa, y ésta lo será aún más si se trata de personas. Cuando hablamos de personas, el concepto de representación se explica, para determinados autores dentro de la ciencia cognitiva, de la siguiente manera: a cada creencia atribuible a la persona corresponde un estado interno de ésta, el cual puede descomponerse en partes funcionales de un modo parecido a como puede descomponerse en palabras la oración que expresa la creencia. Es la hipótesis, defendida por Fodor y otros, de que tenemos en la mente un lenguaje del pensamiento, y que hay que entender la mente como un sistema que manipula símbolos a la manera de un ordenador. Aunque no es una hipótesis aceptada por todos los autores dentro de la ciencia cognitiva, Dennett muestra por ella simpatía (pp. 34 s.), al menos por la importante razón de que el lenguaje es el único modelo que tenemos de un sistema de representación capaz de generar un número potencialmente infinito de representaciones.

Una explicación pragmática de las representaciones recurre a la función que éstas desempeñan en la acción humana. La acción es racional en la medida en que ha de elegir entre alternativas, y la elección se explica en función de las creencias y de las actitudes del sujeto agente. Desde este punto de vista, el objeto de la actitud no es una proposición sino el resultado alternativo de una acción. En una perspectiva distinta, la explicación lingüística de las representaciones subraya que los agentes racionales son esencialmente hablantes, y en consecuencia, se describe el pensamiento como una especie de habla interior, y se explica la representación por analogía con las expresiones lingüísticas (*vid.* Stalnaker, *Inquiry*, cap. 1). En este contexto, el problema de la intencionalidad lo presenta Stalnaker como un problema sobre la naturaleza de la representación: ¿Por qué ocurre que cosas tan dispares como pinturas, nombres, mapas, expresiones proferidas, y ciertos estados mentales, representan, están en lugar de, o son acerca de, otras cosas? La respuesta clásica es que ésta es una capacidad característica de la mente, que distingue a ésta de lo meramente físico. El problema que plantea es que no se puede dar cuenta de lo mental como fenómeno natural a menos que se explique la

intencionalidad en términos naturalistas. La diferencia entre la perspectiva pragmática y la perspectiva lingüística está en que la primera toma como fundamental la intencionalidad de los estados mentales, mientras que la segunda toma como básica la intencionalidad de las expresiones lingüísticas. Stalnaker arguye en favor de la perspectiva pragmática, a diferencia de Hartry Field, quien lo hace en favor del punto de vista lingüístico («Mental Representation»). Esta última argumentación tiene que acabar por atribuir estructura lingüística a los estados mentales.

El concepto de intencionalidad viene así generando interminables discusiones. ¿Es tan alto su valor explicativo para justificar tanta atención? Si la intencionalidad es el rasgo definitorio de la mente, como mantuvo Brentano, entonces ¿hay objetos que, en alguna medida, no tengan mente? Frente a otros objetos, las personas son «prototipos de la categoría de objetos con funciones mentales», en palabras de Angel Rivière (*Objetos con mente*, 2.1, p. 41). Lo serán menos los animales, y tanto menos cuanto más simples sean. ¿Y las plantas? Cuando se orientan hacia la luz, ¿no versa su comportamiento acerca de la luz? Cuando una planta realiza la función clorofílica, ¿no versa esta función acerca de algo? ¿Y los objetos inanimados? Siempre que una propiedad física consista en una relación entre objetos, podremos decir que esa propiedad versa acerca del otro objeto; por ejemplo, la atracción magnética.

Quine ha mantenido que se puede tomar la tesis de Brentano en dos sentidos: o bien como mostrando que las expresiones intencionales son indispensables, o bien como mostrando que carecen de base (*Word and Object*, p. 221). Mientras que Brentano parece proponer su tesis con el primero de estos sentidos, Quine la toma en el segundo. El intento —como puede apreciarse muy bien en la filosofía fenomenológica— ha sido construir una ciencia de la intencionalidad, pero el resultado ha sido volver al idealismo cartesiano. Todo era un intento desesperado por encontrar manifestaciones externas de la conciencia refleja, pero la intencionalidad en el sentido de Brentano, el tratar-acerca-de (*aboutness*, como dicen los anglosajones), es tan general, tan vago e indeterminado, que difícilmente puede construirse una ciencia sobre esta base. ¿Y la ciencia cognitiva? En la medida en que ésta analice la mente tomando como modelo el ordenador, y considere las funciones mentales como funciones de computación, hay que tener en cuenta que la intencionalidad del ordenador es una intencionalidad derivada, mientras que la intencionalidad de la persona parece ser intencionalidad originaria (*vid.* Dennett, *The intentional stance*, cap. 8). ¿Cómo puede explicarse la intencionalidad originaria tomando como modelo la intencionalidad derivada?

Las máquinas fabricadas por nosotros —como los ordenadores— tienen intencionalidad derivada de la nuestra. También es así la intencionalidad que poseen las señales de humo o los libros; es la intencionalidad que podemos atribuir a signos en uso (*tokens*) siempre que sean humanos, y aunque no

sean signos de un lenguaje verbal. A diferencia de ésta, la intencionalidad de la mente humana parece ser originaria, pues no es recibida ni derivada de ninguna otra. Esta es la tesis de Fodor, Searle, Dretske y otros, y contra ella arguye Dennett en la siguiente forma (*The intentional stance*, cap. 8). Primeramente, describe el siguiente caso hipotético: supone que alguien deseara permanecer en hibernación hasta cierta fecha en el siglo xxv (el año 2.401, por ejemplo) a fin de ser despertado en esa fecha y así experimentar las condiciones de vida en tal siglo. La primera condición es construir una cápsula adecuada en la que permanecer en hibernación durante cuatro siglos. No es difícil imaginar que esto puede hacerse. La segunda exigencia es proteger la cápsula contra posibles ataques del exterior y asegurar que recibe la energía necesaria durante el tiempo mencionado. Para cumplir con esta exigencia, sugiere Dennett dos procedimientos alternativos. El primero consiste en encontrar un lugar ideal donde instalar la cápsula, de modo que reciba el suministro adecuado de agua, luz, energía y lo que haga falta. El principal inconveniente de este procedimiento reside en que la cápsula no puede ser movida del lugar que ocupa, en la eventualidad, por ejemplo, de que se decida realizar alguna obra en ese lugar. El segundo procedimiento consiste en suministrar a la cápsula una unidad móvil con la que la cápsula pueda cambiar de lugar en la busca, por ejemplo, de nuevas fuentes de energía. Se trataría de un gigantesco robot, dentro del cual se halle alojada la cápsula. Ambos procedimientos reproducen los dos grandes reinos de los seres vivos; el primero recuerda a las plantas, mientras el segundo recuerda a los animales. En este último caso, Dennett supone que el robot está diseñado de forma que tome decisiones en relación con la protección de la persona que cobija y a cuyos intereses sirve, utilizando para ello los sistemas de percepción y de locomoción de que está dotado. El robot posee, por consiguiente, un cierto grado de autocontrol, que puede permitirle calcular las ventajas de cooperar o de competir con otros robots. Diversos autores defenderían que semejante robot no tiene intencionalidad originaria, porque su intencionalidad deriva del papel que tiene como protector de la persona que lleva dentro, que es la función para la cual ha sido diseñado. Por lo mismo, cuando hace planes o toma decisiones, el robot actúa *como si* hiciera planes o tomara decisiones, pero no posee realmente los correspondientes estados mentales.

De la idea de que un artefacto solamente posee intencionalidad derivada, unida al supuesto de ciencia ficción que acabamos de ver, extrae Dennett (*op. cit.*, p. 298) la conclusión de que nuestra intencionalidad es exactamente como la del robot, puesto que la narración que hemos considerado está construida sobre la tesis, mantenida por Dawkins (*The Selfish Gene*, 1976), de que las especies biológicas, incluida la humana, no son sino artefactos diseñados por la naturaleza para prolongar el futuro de los egoístas genes. No somos sino máquinas de supervivencia para proteger nuestros genes, y nuestra intencionalidad deriva de la intencionalidad de nuestros genes (Dennett,

loc. cit.). ¿Cómo puede atribuirse intencionalidad a los genes? A diferencia de las personas, no son entidades conscientes ni libres. ¿Son acaso los autores del diseño de las personas? El autor del diseño para cada especie, el responsable último, es —concluye Dennett— «el largo y lento proceso de la evolución por selección natural» (*The intentional stance*, p. 299). Esta tesis encierra consecuencias que no parecen fáciles de aceptar: no se explica cómo podemos determinar —si solamente somos máquinas evolutivas— el significado de nuestros pensamientos, pues éstos, en definitiva, vienen determinados por la intencionalidad de la evolución natural, de modo análogo a como el significado de los estados internos del ordenador está determinado por la intencionalidad del artífice del programa. Estamos aquí en presencia de lo que Dretske ha llamado «significado funcional», el significado que algo tiene en razón de las funciones que se le ha asignado, y que, naturalmente, es un significado que deriva de las intenciones y las creencias de quien le asigne esas funciones. ¿Quién asigna funciones? En el caso de un artefacto cualquiera, o de un ordenador, una mente humana. En el caso de un sistema biológico, la evolución natural. Y esto vale igualmente para el ser humano, y por tanto, para la mente humana, cuyos estados reciben su significado del gran artífice, la Naturaleza. La conclusión es que el significado de un estado mental es derivado y, por consiguiente, indeterminado. Dretske intenta evitar esta consecuencia asumiendo que, a través del correspondiente proceso de aprendizaje, un organismo puede establecer un tipo de estado interno que posea una función única y definida, adquiriendo así significado funcional. Aplicándolo a la narración anterior, la idea es que el robot construido para proteger la hibernación del sujeto puede adquirir, por un proceso de aprendizaje, ciertos estados internos que tengan un significado funcional, que no es asignado, sino natural. La tesis, en consecuencia, es ésta: es lógicamente posible crear un artefacto que adquiera por aprendizaje intencionalidad originaria, y que no se limite a poseer una intencionalidad derivada de su fabricante y limitada a las funciones que éste le ha asignado. Es decir: que la intencionalidad originaria puede adquirirse por las relaciones con el entorno, sin ser recibida de ningún artífice. Si la intencionalidad originaria puede adquirirse así, entonces también de ese modo se adquirirán estados psicológicos o mentales. De qué manera ocurra esto sigue siendo misterioso.

Más claro resulta en la perspectiva evolucionista de Dennett, donde la intencionalidad es recibida de la evolución natural. Pero esto se reduce a la trivialidad de afirmar que la naturaleza ha producido seres con intencionalidad, y por tanto, con mente. ¿Puede, por ello, afirmarse que la intencionalidad de esos seres es derivada de un artífice? Decir que esa intencionalidad deriva de la que posee la naturaleza supone representar a la naturaleza como un organismo o sistema intencional, y por consiguiente, con estados mentales. ¿En qué sentido tiene intencionalidad la naturaleza? Solamente en el sentido de que hay entidades naturales con intencionalidad, y añadir que la han adquiri-

do de la naturaleza es lo mismo que decir que la poseen de forma natural, o lo que tanto da, que puesto que no son entidades fabricadas (artefactos), no hay ninguna mente previa que se la haya dado. Y si es concebible que un artefacto adquiriera intencionalidad por aprendizaje, no existe razón ninguna por la que se pueda negar que algunas especies biológicas adquieren la intencionalidad por un proceso de aprendizaje, que es en definitiva en lo que consiste su relación con el entorno. La vaguedad de la tesis de Brentano permite atribuir intencionalidad a todos los seres vivos, e impide tomar la intencionalidad como la característica distintiva de lo psicológico. A menos que se acepte que hay grados de intencionalidad, y grados también en la posesión de mente, en cuyo caso podríamos atribuir estados psicológicos de nivel tanto más bajo cuanto menor es la complejidad del organismo en su estructura y en su comportamiento.

La atribución de intencionalidad a la naturaleza implica —por la tesis de Brentano— atribuirle mente, y ésta es una consecuencia contraria a la concepción científica de la naturaleza que tenemos heredada de Darwin. A pesar de ello, la actitud intencional resulta tan útil —como comenta Dennett (p. 314)— que los propios biólogos no pueden sustraerse a ella, y en consecuencia hablan de que las enzimas *discriminan* entre aminoácidos y también de que *corrigen* sus errores (L. Stryer, *Biochemistry*, 1981). El propio Dennett toma esta forma de hablar tan sólo como una interpretación mentalista de estos procesos, «como si» fueran intencionales, aunque realmente no tienen intencionalidad ni originaria ni derivada. Se trata de una metáfora, cuyo valor es semejante al valor de la metáfora empleada por Dawkins cuando habla del «egoísmo» de los genes: hablar de entidades biológicas simples en términos intencionales nos ayuda a describir su comportamiento de forma más breve y más sencilla, y nos permite atribuir funciones a las entidades biológicas. De modo congruente, Dennett habla de las razones que presiden la evolución natural, añadiendo:

«La naturaleza ha apreciado estas razones sin representárselas. Y el propio proceso del diseño es la fuente de nuestra intencionalidad propia» (p. 317).

Con lo cual parece reconocerse que la intencionalidad brota del diseño evolutivo de la naturaleza, y en consecuencia habría que aceptar que la actitud del diseño, aplicada al devenir natural, es más básica que la actitud intencional. Lo cual es coherente con esta conclusión de Dennett:

«nuestra intencionalidad es altamente derivada, y precisamente de la misma manera en que la intencionalidad de nuestros robots (e incluso de nuestros libros y mapas) es derivada (...). Por tanto, si ha de haber alguna intencionalidad originaria —en el sentido de no ser derivada de

ninguna otra fuente ulterior— la intencionalidad de la selección natural merece este honor» (p. 318).

Pero en mi opinión, si se trata de explicar la mente humana entonces hay que reconocer como su característica distintiva la conciencia refleja, esto es, la posesión de estados mentales acerca de estados mentales propios; es una intencionalidad sobre los estados psicológicos del propio sujeto, por tanto, una intencionalidad de segundo orden, que podemos llamar *meta-intencionalidad*. Esta, y no la simple intencionalidad de la que hablaba Brentano, es la auténtica marca característica de lo psicológico humano frente al resto de la naturaleza. Por esta razón, Roger Penrose se ha planteado la cuestión de cuál es la ventaja selectiva que la conciencia aporta a los objetos que la poseen (*The Emperor's New Mind*, 1989, cap. 10). Lo que esta pregunta implica, Penrose se inclina a formularlo en términos del principio antrópico: la naturaleza del universo obedece a la exigencia de que existan en él seres sentientes que lo observen (pp. 405-406). Estar consciente —comenta Penrose— es estar consciente de algo, que puede ser una sensación (dolor, calor, visión), o bien un sentimiento (asombro, felicidad), o bien un recuerdo, o la comprensión de una idea, o la intención de hacer algo... (pp. 406-407). Como todos estos son estados mentales, y por consiguiente caracterizables (en el enfoque tradicional) por la intencionalidad, estamos presentes ante lo que acabo de llamar *meta-intencionalidad*. Pero ésta puede darse en un grado aún más alto, pues se puede ser consciente de la propia conciencia (Penrose, *ibid.*).

¿Cómo se manifiesta la conciencia en los organismos que la poseen? Responder a esta pregunta parece una condición necesaria para que se pueda hacer una ciencia de la conciencia. La gran disparidad de opiniones acerca de la respuesta ha sido subrayada por Penrose, quien, por su parte, se manifiesta dudoso de que haya algo de conciencia en un gusano o en un insecto, pero no duda en atribuírsela a los mamíferos en general (p. 408). Tal vez estos sean dos extremos claros en una taxonomía psicológica, pero tendremos que preguntarnos dónde colocar los peces, las aves y los reptiles mientras no sepamos cuál es el criterio que debemos aplicar. No quiere esto decir que no exista un criterio, sino únicamente que no hay un criterio generalmente aceptado. La tesis de Penrose es que una acción consciente es tal que no puede ser descrita por un algoritmo, a diferencia de lo que acontece en el caso de una acción inconsciente; la formación de juicios nuevos, que es característica de la conciencia, es algo que no sabemos programar en un ordenador (pp. 411-412; y también *Shadows of the Mind*, 1994, c. 1).

Antes de considerar un criterio positivo para caracterizar la conciencia, debe aceptarse que hay grados de conciencia como hay grados de complejidad en la estructura y en el comportamiento de un organismo. Sin perder de vista esta gradación, propongo considerar que podemos atribuir conciencia (y por tanto, mente) a un organismo o sistema, en la medida en que podamos

afirmar que tal organismo predice las consecuencias de su comportamiento, proyecta su conducta ulterior, y elige entre alternativas. En suma, nuestra interpretación de un sistema u organismo al que atribuyamos conciencia debe ser tal que le reconozcamos libertad. Y aunque debamos aceptar grados de libertad parejamente a grados de mente o de conciencia, lo que distingue al mundo mental del mundo físico es la libertad.

Referencias bibliográficas

Aristóteles, *Metafísica*.

Bonitz, H., 1849, *Aristotelis Metaphysica. Commentarius*, Bonn (reed. 1960, G. Olms, Hildesheim).

Brentano, F., 1874, *Psychologie vom empirischen Standpunkt*, Leipzig.

Chisholm, R., 1957, *Perceiving: A Philosophical Study*, Cornell University Press.

Cruz Hernández, M., 1953, *Francisco Brentano*, Universidad de Salamanca.

Dawkins, R., 1976, *The Selfish Gene*, Oxford University Press.

Dennett, D., 1969, *Content and Consciousness*, Routledge, Londres.

—, 1978, *Brainstorms*, Harvester Press, Sussex.

—, 1987, *The Intentional Stance*, MIT Press, Cambridge, Mass.

—, 1991, *Consciousness Explained*, Little Brown, Boston.

Field, H., 1978, «Mental Representation», *Erkenntnis*.

Penrose, R., 1989, *The Emperor's New Mind*, Oxford University Press.

—, 1994, *Shadows of the Mind*, Oxford University Press.

Quine, W., 1960, *Word and Object*, MIT Press, Cambridge, Mass.

Reale, G., 1968, *Aristotele. La Metafísica*, L. Loffredo, Nápoles.

Rivière, A., 1991, *Objetos con mente*, Alianza, Madrid.

Ross, W. D., 1924, *Aristotle's Metaphysics*, Oxford Clarendon Press.

Schwegler, A., 1847-48, *Die Metaphysik des Aristoteles*, Tübingen.

Stalnaker, R., 1984, *Inquiry*, MIT Press, Cambridge, Mass.

Stryer, L., 1981, *Biochemistry*, Freeman, San Francisco.