

# LA ANALOGÍA FUNCIONALISTA

CARLOS PEREDA

INSTITUTO DE INVESTIGACIONES FILOSÓFICAS  
UNAM

JORGE TAGLE

FACULTAD DE FILOSOFÍA Y LETRAS  
UNAM

Según el funcionalismo, estados mentales como las actitudes proposicionales “creer que *p*”, “desear que *p*”, “considerar que *p*”, . . . , *son* “estados computacionales” de la mente/el cerebro. Esto es, la manera más adecuada de pensar la mente/el cerebro es pensarlos como una computadora digital, cuyo “*software*”, cuya “organización funcional”, es nuestra psicología. De esta manera, a la pregunta “¿qué es la mente?” se la interpreta “¿cómo funciona la mente?” y a la pregunta “¿qué es un contenido mental?” como interrogando “¿qué es un estado mental?” (dando por sentado que lo que sucede en el interior de la mente se da en forma de estados y no de otra forma, y faltando aún por aclarar qué sentido tiene aquí el término “estado”); en ningún caso se intenta, pues, responder a preguntas como “¿es la mente materia o espíritu?” o “¿existe una sola mente o existen muchas?”.

En un serie de influyentes trabajos, Hilary Putnam introdujo varias propuestas funcionalistas para pensar la mente/el cerebro y, luego, las ha criticado con vigor (lo que da ya una lección formal acerca de lo que *es* una actitud filosófica genuina: la verdad importa más que los propios pensamientos o teorías). En este trabajo se discutirá al funcionalismo reconstruyéndolo como una analogía, distanciándonos, pues, del funcionalista que lo piensa como algo más: en términos de identidad; esta distancia tiene la ventaja de permitirnos sugerir algunas diferencias entre los estados mentales y los “estados” computacionales (al menos, comprendidos de cierta manera). El plan de trabajo es el siguiente: se comienza por indicar muy brevemente en qué consiste proponer una analogía, o un argumento por analogía, se pasa después a reconstruir la versión de Putnam de la analogía funcionalista y alguna de sus críticas, y, lo que más nos importa, se termina explorando ciertas consecuencias que puede

tener para nuestras concepciones más generales de la racionalidad aceptar ya sea la analogía funcionalista, ya sean sus críticas.

## I

¿Qué es un argumento por analogía? Este tipo de argumentos contienen una analogía en, por lo menos, una de sus premisas. Así, un argumento por analogía apoya cierta conclusión comparando una cosa con otra real o imaginaria; llamémosle a la primera, el análogo y a la segunda, el analogado: en el caso de la analogía funcionalista, el análogo es una computadora y su *software*, y el analogado, el cerebro y la vida mental.

Sobre el análogo, la computadora, se supone que hay más información o acuerdo que sobre el analogado, el cerebro, la vida mental, pero teniendo en cuenta una presunción de similaridad relevante entre ambos, *se da un salto* y se trasporta la información o el acuerdo del análogo al analogado. En general, los apoyos por analogía se basan en la máxima “es racional tratar a los casos similares de manera similar”. Una analogía es, entonces, *buena* si:

- a) el análogo es mejor conocido que el analogado;
- b) hay la presencia de similitudes importantes entre el analogado y el análogo; y
- c) hay la ausencia de diferencias importantes entre el analogado y el análogo.

Es excesivo decir que a), b) y c) son “criterios”, se trata más bien de “variables que deberán tenerse en cuenta” en nuestros argumentos: apoyándose en la presencia de similitudes y en la ausencia de diferencias quien argumenta *apuesta* a que el patrón de las similitudes entre el analogado y el análogo sea genuino, de tal manera que las propiedades conocidas en el análogo, y que interesa conocer en el analogado, se encontrarán también en el analogado. Sin duda, un argumento por analogía es mucho más débil que un argumento deductivo: incluso las analogías más próximas suelen romperse. Por lo pronto, en toda analogía hay dificultades para evaluar la apuesta articulada por las “variables” a), b) y c): para distinguir *cuán* importantes son las similitudes entre los casos, *cuánto* lo son las diferencias, y *en qué sentido* funcionan similitudes y diferencias.

## II

La propuesta funcionalista de Putnam consistió, como ya se dijo, en pensar la mente humana mediante una analogía con una computadora digital. Esto es, se piensa que los estados mentales funcionan como los de una computadora. Una computadora —autómata probabilista— se define de manera si-

milar a una máquina de Turing, con la única diferencia de que en el autómata probabilista —computadora— la transición de un estado a otro tiene varias probabilidades, mientras que en la máquina de Turing el paso de un estado a otro es algo totalmente determinado.<sup>1</sup> Independientemente del material de que esté compuesto el autómata probabilista, o la computadora, lo verdaderamente importante es aquello que describe el funcionamiento o “proceder interno” de la computadora, esto es, la “tabla de máquina” (*machine table*).<sup>2</sup> Un ejemplo de “tabla de máquina” sería lo siguiente:

|                             | <i>A</i>                        | <i>B</i>                        | <i>C</i>                        | <i>D</i>                        |
|-----------------------------|---------------------------------|---------------------------------|---------------------------------|---------------------------------|
| ( <i>s</i> <sub>1</sub> ) l | <i>s</i> <sub>1</sub> <i>RA</i> | <i>s</i> <sub>1</sub> <i>LB</i> | <i>s</i> <sub>3</sub> <i>LD</i> | <i>s</i> <sub>1</sub> <i>CD</i> |
| ( <i>s</i> <sub>2</sub> ) + | <i>s</i> <sub>1</sub> <i>LB</i> | <i>s</i> <sub>2</sub> <i>CD</i> | <i>s</i> <sub>2</sub> <i>LD</i> | <i>s</i> <sub>2</sub> <i>CD</i> |
| ( <i>s</i> <sub>3</sub> )   | <i>s</i> <sub>3</sub> <i>CD</i> | <i>s</i> <sub>3</sub> <i>RC</i> | <i>s</i> <sub>3</sub> <i>LD</i> | <i>s</i> <sub>3</sub> <i>CD</i> |

Las hileras de la tabla corresponden a letras minúsculas del alfabeto; las columnas corresponden a los estados *A*, *B*, *C*, etc. En cada cuadro —estado— aparece una *instrucción* que debe ser leída del modo siguiente: ‘*s*<sub>5</sub>*LA*’ quiere decir, primero, que se imprima el símbolo *s*<sub>5</sub> en el cuadro que en el momento se está examinando (una vez que se haya borrado el signo anterior); segundo, que se proceda a examinar el cuadro inmediato izquierdo y, por último, que se cambie al estado *A* de esa misma hilera. (*R*’ quiere decir examinar el cuadro inmediato a la derecha, *C*’ quiere decir “centro”, esto es, continuar con el examen del mismo cuadro).<sup>3</sup> Cada instrucción indica cómo ha de proceder el autómata.

La “tabla de máquina” describe a una máquina sí ésta posee estados internos que corresponden a las columnas de la tabla y, además, sí la máquina “obedece” las instrucciones de la tabla, esto es, sí al examinar un cuadro en el cual aparece un símbolo *s*<sub>1</sub>*B*, lleva a cabo la instrucción en la hilera y la columna apropiada de la tabla.

La “descripción lógica” del autómata probabilista —el *software* de la computadora—, repetimos, no prevé ninguna descripción de la naturaleza

<sup>1</sup> Véase Putnam, 1967a, p. 433.

<sup>2</sup> Alguien se podría preguntar si éste es el modelo que actualmente defienden los funcionalistas. Nos parece que, por lo menos, la idea básica que intentó hacer plausible el funcionalismo está contenida en el modelo que describimos aquí. Nos parece que el formalismo de la máquina de Turing es algo no sólo básico sino *necesario* para entender la analogía funcionalista.

<sup>3</sup> Véase 1960, p. 365.

física de la misma. O bien, el autómata es una máquina *abstracta*, la cual puede estar materialmente hecha de un número infinito de maneras y materiales.<sup>4</sup>

Otra modificación con respecto a la máquina de Turing se refiere al hecho de que el autómata probabilista incluye “entradas sensoriales” (*sensory inputs*) y “salidas motoras” (*motor outputs*). Luego, la “tabla de máquina” determina, para cada combinación posible “estado” + “entrada sensorial”, una instrucción que fija la probabilidad del siguiente “estado” y de las “salidas motoras”.

La vida mental —lo que sucede en el autómata probabilista— puede especificarse, sea en una situación particular o en general, describiendo la situación de la “tabla de máquina”. Es decir, la “tabla de máquina” nos dará la organización funcional del organismo en cuestión a la manera como un programa de computadora nos indica los pasos que sigue ésta para realizar una operación. Si un organismo *Y* profiere la oración *x* es porque posee una determinada organización funcional (esto es, en la “tabla de máquina” están presentes y relacionados de cierta manera una serie de “estados” *A*, *B*, etc.).<sup>5</sup> O bien, si un organismo *Y* realiza la acción *z* es porque poseía una organización funcional específica; lo que equivale a decir que tuvo lugar en la “tabla” una serie de estados relacionados de cierta forma.

La mente humana, se sugiere, *funciona* de manera similar al autómata: al encontrarse en un determinado estado *X*, por ejemplo, creer, la mente posee una organización funcional específica, o mejor dicho, porque posee una determinada organización funcional (combinación de estados de la “tabla de máquina”) es que tiene la creencia *x*. Además, también en el caso del lenguaje hay una conexión entre estado mental y habla (verbalización). De manera simultánea o previa al hablar deben ocurrir en la mente o en la “tabla de máquina” del autómata probabilista una secuencia determinada de estados *A*, *B*, *C*, . . . , esto quiere decir, que sólo tendrá lugar el habla, si al mismo tiempo o previamente se tiene el estado pertinente/la organización funcional adecuada.

Resolver problemas, pensar, creer, desear —esto es, los contenidos propios de la mente—, *qua* organizaciones funcionales del ser humano pueden ser descritos en términos de secuencias de estados de “tabla de máquina” de cada persona sin hacer referencia a la realización física de tales estados.<sup>6</sup> No importa dónde se realizan, sino cómo es cada uno y cuál sigue a cuál. Lo que le interesa al funcionalista es, por tanto, el programa que suministra el orden y las posibles combinaciones y configuraciones de los estados mentales.

¿Qué decir de esta analogía? Ya señalamos que una analogía es buena si: a) el análogo es mejor conocido que el analogado, y b) hay la presencia de similitudes importantes entre el análogo y el analogado. Sin duda, se satisface

<sup>4</sup> *Ibid.*, p. 371.

<sup>5</sup> Véase 1967a, p. 434.

<sup>6</sup> Véase 1960, p. 373.

la variable a): el análogo, la computadora, es mejor conocido que el cerebro/la vida mental. Examinemos si también se cumple la variable b).

Como una primera similitud hay que mencionar que tanto la computadora como la mente humana parecería que contarán con una función preferencial constante debido a la cual procuran el máximo de utilidad (eficacia) con el mínimo de esfuerzo; siempre buscan obtener lo más, invirtiendo lo menos (sea esto energía, tiempo, etcétera).

Una segunda similitud consiste en que, al parecer, los problemas tradicionales de la discusión mente/cuerpo, o al menos algunos de esos problemas, surgen también para cualquier sistema computacional:

The conceptual issues surrounding the traditional mind-body problem have nothing to do with the supposedly special character of human subjective experience, but arise for any computing system able to construct theories concerning its own nature.<sup>7</sup>

Incluso quien piensa que Putnam exagera con la expresión “*have nothing to do*”, tanto en el caso de los seres humanos como en el de las computadoras, se presenta lo que podríamos llamar una dualidad de procesos: por un lado, lógico-conceptuales y por el otro, físico-químicos. Para ambos surge la pregunta de si los primeros no son en última instancia reductibles a los segundos. Éste es, a primera vista, un punto que habla mucho en favor de la propuesta funcionalista, pues nos dice que al describir/explicar nuestra vida mental, no es necesario referirnos al “material de que estamos hechos”, esto es, no importa la realización física del *software*. Cuando se hace la descripción de la “vida mental” o *software* de los seres humanos no interesa referirnos a las operaciones físico-químicas del cerebro, ni en general a la constitución física del organismo en cuestión.

También hay que anotar una tercera similitud: tanto la mente humana —el analogado— como las computadoras —el análogo— pueden recibir “información” del mundo externo, esto es, ambos cuentan con órganos sensoriales que les permiten obtener datos del mundo exterior. Estos estímulos sensoriales, para utilizar un término más general, pueden influir en el proceder de la mente o la computadora. Con ello queda integrado al modelo funcionalista la influencia que puede tener el medio ambiente en el contenido, la complejidad y la transición de los estados mentales.

Si lo dicho hasta el momento es correcto, entonces parecería que tienden a satisfacerse las variables a) y b) para evaluar una analogía como buena, esto es, la satisfacción de a) y b) parece ser suficiente para permitirnos transportar la información existente respecto del análogo al analogado. Hay que aclarar,

<sup>7</sup> Putnam, 1967b, p. 408.

sin embargo, que lo se transportaría del análogo al analogado sería la *forma* en que se describe la vida mental y no la descripción misma de la organización funcional. Insistimos, el funcionalismo reconstruido como analogía no intenta utilizar la misma descripción de la organización funcional, el *software* de la computadora para la vida mental, sino que sólo toma la *forma* de la descripción —la noción de “organización funcional”.

Pero ¿qué decir de la analogía funcionalista en relación con la variable c)?, a saber, ¿“hay la ausencia de diferencias importantes entre el analogado y el análogo”? Para responder esta pregunta formularemos algunas objeciones, en gran parte ya planteadas explícitamente por el propio Putnam,<sup>8</sup> que introducen ciertas dudas en relación con las similitudes propuestas por al analogía funcionalista.

### III

#### 1. *El ataque antialgorítmico*

El funcionamiento de nuestra vida mental *no puede reducirse a un algoritmo*; en cambio esta reducción *se puede* llevar a cabo en relación con una computadora. En este contexto, la expresión “carácter no algorítmico de la vida mental” tiene el sentido de no consistente. En efecto, se puede pensar y actuar de manera no consistente y ello puede resultar, en cierta situación, *la respuesta más racional posible*; ¿en qué tipo de situaciones estamos pensando?

Por ejemplo, no es raro que un preferir *A* coexista con un preferir *A* en la vida mental de un sujeto, esto es, un individuo no siempre actúa según la máxima “en las situaciones del tipo  $s_1$ , *P* prefiere siempre el curso de acción que requiera el mínimo de esfuerzo”, sino que, a veces actúa según la máxima “en las situaciones del tipo  $s^1$ , *P* actúa a veces de tal modo que no importa si hay otro curso de acción que implique menor esfuerzo”. No es una afirmación aventurada decir que los seres humanos, en determinadas situaciones, muy razonablemente sacrifican un máximo de utilidad en pro de otro tipo de consideraciones que estiman más valiosas. En su sentido más fuerte esto querría decir que el análogo —la computadora— *siempre* se comporta según las máximas programadas, o en algún sentido permitidas, mientras que en el caso del analogado dicha(s) máxima(s) está(n) en relación no sólo con el contexto, sino con múltiples razones y con la historia personal que posee cada sujeto particular.

Considérese el siguiente ejemplo: la paradoja del prefacio. El prefacio I dice que aunque el autor de cierto libro *B* cree que cada uno de los enunciados contenidos en ese libro es verdadero, al mismo tiempo no duda que el libro contiene algunos errores. Por el contrario, el prefacio II señala que el autor

<sup>8</sup> En tres escritos, principalmente: 1967b, 1973 y 1988.

de ese libro cree en la verdad de cada uno de los enunciados del libro, y punto. El prefacio I hace que la totalidad del libro —prefacio más texto— sea incoherente. Éste no es el caso con el prefacio II. Sin embargo, este último probablemente sea visto como la expresión de una insoportable arrogancia. Tenemos, pues, que elegir aquí, como en muchas otras situaciones, entre varias virtudes y vicios: la honestidad y la incoherencia por un lado, y la arrogancia y la coherencia por otro. ¿Qué hacer? Por lo pronto, hay que tener cuidado de no descalificar esta dificultad como un rompecabezas sin importancia. Si pensamos un poco, de inmediato nos damos cuenta que el dilema del prefacio atañe a todos los saberes, incluyendo a todos los saberes científicos: cada persona piensa —tiene que pensar— que cada creencia que constituye su trama de creencias, su “saber”, es una creencia verdadera, pues si no lo fuese, no lo podría considerar legítimamente como “saber”. Pero, al mismo tiempo, cierta cautela, una metainducción sobre creencias rechazadas, empuja también a formular la razonable conjetura de que algunas de las creencias consideradas como verdaderas son probablemente falsas. De esta manera, comprobamos que la contradicción es *un* vicio epistémico, con el cual a veces tenemos que aprender a vivir si ello nos permite promover otras virtudes epistémicas o prácticas. Éste no podría ser el caso en el *software* de una computadora.

Nuestro argumento es, entonces, puramente negativo: no ponemos en duda que una computadora pueda programarse de tal modo que, en muchos aspectos, pueda simular la vida mental. Sólo puntualizamos que, en ese caso, habrá que abandonar el ideal algorítmico de racionalidad como absoluta consistencia y completud, y recurrir a estructuras más complejas (lógicas relevantes, lógicas paraconsistentes, . . .).

## 2. El ataque en contra de la autonomía de la vida mental

El funcionalismo afirma: nuestra vida mental no depende de *our substance*, pues sea lo que sea aquello de lo que estamos hechos, queso, carbono, agua, ello *no importa* cuando queremos describir/explicar nuestra vida mental —creencias, deseos, emociones—; esto es, para saber cómo es y cuál es la forma de nuestra vida mental podemos prescindir enteramente, ya no sólo de aquello de que estamos hechos, sino de si ese *software* está realizándose en algún material particular. Pero ¿esto es verdad?

Tengamos en cuenta ciertos datos y elaboremos con ellos algo así como un argumento parte-todo: el hecho de que tengamos *un* cuerpo es parte del hecho de que tengamos cierta forma de vida y, a su vez, es parte de tener ese cuerpo y de vivir esas formas de vida, el poderme atribuir y atribuir a los demás ciertos deseos, creencias, emociones. Así, hay muchos deseos que tienen que ver de modo inmediato con el cuerpo, tal como éste es en las personas, y que serían incluso ininteligibles sin relación con ese cuerpo, deseos a

la manera de todos aquellos que directa o indirectamente están relacionados con el hambre, la sed o la sexualidad. Por otro lado, una parte importante de las creencias, y hasta de las mitologías, como señala Borges en su poema "Historia de la noche", dependen de esos "tenues instrumentos, los ojos". De alguna manera, algo similar puede decirse de los otros sentidos, piénsese en el papel que juegan los diferentes tipos de comidas en la vida de las personas, o al menos, de muchas de ellas, y su dependencia de las "tenues" posibilidades del gusto humano. La lista de datos en *esta* dirección podría proseguirse. Además, no hay que olvidar el *principio holístico* que rige la vida mental en general y que hace que los ejemplos dados de ciertos deseos y creencias, u otros similares, no puedan reducirse a "ghettos" aislables y, así, a ejemplos que no "comprometen" con sus "influencias" corporales —o si se prefiere, con su "dimensión" corporal— a los otros componentes de la vida mental. Como en la música, "todo" en *algún* sentido, acaba dependiendo de la materialidad del instrumento.

De ahí que en esa vasta trama de deseos, creencias, emociones que constituye la vida mental y, en particular, la vida racional, no haya, en sentido estricto y absoluto, "islas", que directa o indirectamente —a veces, claro, muy *mu*y indirectamente— no reciban las "resonancias" del cuerpo.

Examinemos todavía algunas emociones, como ciertos miedos, digamos el miedo a las alturas o el miedo a las arañas. Hay muchos casos —aunque no todos— en que la emoción de una persona de sentir miedo a las alturas o a las arañas depende de su creencia que atribuye ciertas propiedades a su cuerpo y, también, de su creencia de que ciertos sucesos, objetos o animales pueden causarle daño, esto es, de que su cuerpo puede ser seriamente dañado si se cae de cierta altura o si es infectado por el veneno de las arañas. Por ello, si las personas tuvieran otro cuerpo, otra constitución física, no sentirían esos miedos, esto es, si tuviesen un cuerpo resistente a las caídas de grandes alturas y resistente también a las arañas venenosas, del mismo modo en que el cuerpo resiste ciertos grados de temperatura, entonces, no tendrían las creencias citadas respecto de caerse de las alturas o de ser picado por las arañas, ni sentirían los miedos que estamos discutiendo. Quizá podemos llamar al tipo de emociones a los cuales pertenecen estos miedos, "emociones racionales", emociones respaldadas en creencias verdaderas.

Por supuesto, también tenemos otros tipos de emociones, por ejemplo, se puede tener un "miedo irracional" a las arañas, en el sentido de tenerles miedo incluso aunque se sepa que se está ante arañas no venenosas. Mi sospecha es que este segundo tipo de emociones, "emociones irracionales" o emociones que contradicen nuestras creencias más racionales, depende para existir —aunque *no* en una relación directa, o de uno a uno, sino en general— del primer tipo. Esto es, las personas pueden tener, en el sentido de "tienen

sentido para ellas”, las “emociones irracionales”, porque tienen “emociones racionales”, emociones cognoscitivamente correctas.

Sin embargo, independientemente de que esta propuesta sea defendible, la existencia misma de las “emociones racionales”, contribuye a fortalecer el argumento parte-todo en relación con deseos que dependen, incluso para tener sentido, de poseer cierto cuerpo como el nuestro, o de aquellas creencias y hasta mitologías, dependientes de manera directa de los sentidos. Por eso, creemos que no es tan arriesgado afirmar que tal vez Wittgenstein tuvo en mente algunos pensamientos como éste cuando afirmó: “Si un león hablara no lo entenderíamos.” En suma, nuestra vida mental no parece estar del todo desligada del lugar en donde inmediatamente se realiza: el cuerpo humano.

### 3. El ataque “intencionalista”

Por “intencionalidad” se entiende el fenómeno que vincula pensamiento y cosa, la mente o la conciencia y el mundo exterior; es, en muchos sentidos, la manera epistémica de hablar del concepto semántico de referencia, así como contenido mental es, en algunos sentidos, la manera epistémica de hablar del concepto semántico de significado o sentido. La analogía funcionalista parece dar una explicación reductiva de la intencionalidad, o si ello no fuese posible, se pide eliminar este concepto, alegando que, de todos modos, “intencionalidad” es un término del vocabulario de la “psicología popular”, una de las tantas reliquias de nuestras creencias precientíficas, de nuestras supersticiones. En *Representation and Reality*, Putnam rechaza ambas actitudes: la intencionalidad no puede ser reducida ni tampoco eliminada.

Atendamos, primero, los argumentos de Putnam en contra de la reducción. Quien se propone reducir la intencionalidad tiene a su disposición, por lo menos, dos caminos. El primer camino es el mentalismo, en algún sentido inspirado por Chomsky y desarrollado por Fodor, el segundo camino es el reduccionismo físico-químico. Directamente en contra del primero, pero de una manera que también puede extenderse al segundo, Putnam introduce tres razones por las cuales el reduccionismo no puede ser correcto:

1. *El significado es holístico*, las palabras cobran su sentido en relación con otras palabras, en el entramado con las otras palabras. No es posible, pues, reducir el lenguaje a un “vocabulario básico” (por ejemplo, al vocabulario de las sensaciones, como quisieron algunos positivistas).
2. *El significado es, en parte, una noción normativa*, y como los conceptos de creencia justificada, creencia garantizada o creencia razonable no pueden ser reducidos a conceptos fisicalistas.
3. *Nuestros conceptos dependen de nuestro entorno físico y social de un modo que la evolución no podía prever*, esto es, palabras como “carburador”,

“burócrata”, “soneto” o “sinfonía” dependen de las contingencias de la historia, en un sentido que ninguna semántica innatista podría elucidar.

En estas tres observaciones podemos apoyar la propuesta crítica de Putnam en contra del antiintencionalismo:

We cannot individuate concepts and beliefs without reference to the *environment*. Meanings aren't "in the head".<sup>9</sup>

En consecuencia, las actitudes proposicionales no son estados funcionales definibles en términos de una descripción del *software* del organismo. Las actitudes proposicionales *no* son meros estados computacionales del cerebro.

Putnam es consciente de que en contra de su propuesta, se le podría objetar una analogía más abarcadora: la sociedad entera de organismos y partes apropiadas del entorno físico como análogas a una computadora. Putnam responde: la determinación del significado no sólo depende en parte del entorno social y en parte de la naturaleza de las cosas externas, sino que también el significado y la referencia dependen de “descontar las diferencias de creencias”. ¿Qué es esto? Se sabe que la atribución de significados a una persona, esto es, que la comprensión/interpretación de su lenguaje debe realizarse conjuntamente con la atribución de creencias y deseos de la persona en cuestión (tesis del holismo de la vida mental). Pero los deseos y creencias del otro no son idénticos a los nuestros, de ahí que considerar dos palabras como sinónimas o no depende de nuestros criterios de razonabilidad y de qué consideramos, en ese momento, que es una interpretación razonable. Pero como son imposibles conceptos formales de racionalidad y de interpretación razonable, como es imposible un *algoritmo de la interpretación*, son imposibles también definiciones funcionalistas de sinonimia y correferencialidad y, por lo tanto, son imposibles definiciones funcionalistas de significado y referencia.

Putnam todavía examina la siguiente objeción: las dificultades anteriores demuestran que “no hay esperanzas” de tener en la práctica un algoritmo de la interpretación pero ¿no puede existir ese algoritmo en principio, aunque ningún ser humano pueda disponer de él o siquiera comprenderlo?

Cuando se formulan preguntas como éstas, hay que distinguir entre:

- a) preguntas que, dadas las circunstancias actuales, son humanamente indecidibles, pero que sabemos en qué podría consistir una respuesta correcta (por ejemplo, preguntas como “¿hay habitantes fuera de nuestra galaxia?”), y
- b) preguntas con respecto a las cuales no tenemos la menor idea de en qué podría consistir su respuesta o pseudopreguntas.

<sup>9</sup> Putnam, 1988, p. 73.

La pregunta acerca de la posibilidad en principio de un algoritmo de la interpretación pertenece, creemos, a la clase b) de preguntas o pseudopreguntas. Señala con razón Putnam:

The claim that there is something finitely specifiable that all cases of correct interpretation have in common is one that we have simply no reason to believe.<sup>10</sup>

Ello posee graves consecuencias. Una actitud ingenua acerca de las relaciones entre la mente y el cerebro proponía que las descripciones de actitudes proposicionales corresponden de manera biunívoca a estados del cerebro, independientemente de la especie; el funcionalismo rompió con esto para proponer que las descripciones de actitudes proposicionales corresponden biunívocamente a estados del cerebro en cada organismo individual. Pero, según Putnam, esto es sólo “ciencia popular” (en el sentido peyorativo de la expresión). ¿Por qué? Si bien los seres humanos son todos, en el momento de nacer, computadoras de la misma clase, los diferentes seres humanos pasan por distintas secuencias de estados, que dependen de sus particulares biografías, cuando fijan sus diferentes creencias. De ahí que nuestras interpretaciones de los otros no puedan reducir la intencionalidad a estados neuroquímicos o computacionales, pues no proceden, ni aislando “estados neuroquímicos”, ni mediante un algoritmo de la interpretación, sino “descontando las diferencias”.

Detengámonos todavía en la posibilidad de eliminar la intencionalidad. Para los eliminativistas no constituye una dificultad grave que resulte imposible hacer una reducción de las actitudes proposicionales ya que, en especial, las creencias y los deseos son “entidades cuasimitológicas”, que forman parte de una colección de prejuicios primitivos articulados en la “psicología popular”. Por ejemplo, para Paul y Patricia Churchland una verdadera explicación de aquello que llamamos “mente” o “fenómenos mentales”, como las “creencias” y los “deseos”, sólo se puede dar en términos del funcionamiento de las neuronas. Como señala Putnam, las consecuencias del eliminativismo son más graves de lo que podría pensarse a primera vista, incluso para la posición desde la que argumentan los eliminativistas: si conceptos como el de creencia y traducción han de rechazarse, hay que rechazar también los conceptos de referencia y verdad. Pero los conceptos de referencia y verdad son los conceptos fundamentales de la lógica. Entonces, si hay que hablar con desprecio de una “psicología popular”, de manera similar habría que hablar de una “lógica popular”. Sin embargo, ¿estamos dispuestos a pagar *este* precio? Putnam piensa con razón que es un precio excesivo e irrazonable, y que

<sup>10</sup> *Ibid.*, p. 93.

quienes están en dificultades no son quienes defienden el concepto de intencionalidad sino quienes intentan reducir o eliminar este concepto, así como otros conceptos de la “psicología popular”, o de la “lógica popular”.

#### IV

¿Qué podemos concluir de toda esta discusión en pro y en contra de la analogía funcionalista? Por lo pronto, la disputa en torno al funcionalismo deja indirectamente entrever *otra* disputa, todavía más decisiva: la disputa acerca de la naturaleza de la racionalidad. Un funcionalista tenderá a pensar que los procesos más importantes de la mente son los procesos racionales (argumentar deductivamente, etc.) y que un algoritmo permite, básicamente, reconstruir tanto los propios procesos de pensamiento como interpretar el funcionamiento de la mente de los otros; el resto de la vida mental, para cuya reconstrucción tengamos que apelar a estructuras más complejas (lógicas relevantes, lógicas paraconsistentes. . .) tendrá menos interés. En este sentido, se pensará a la racionalidad como una *racionalidad dura* y a todo lo que quede fuera de ella se le descalficará como procesos irracionales o arracionales, procesos de poca monta. (Por lo demás, no sólo los funcionalistas sino también los reduccionistas y los eliminativistas de la vida mental apuestan por igual a un concepto de la racionalidad dura.) Quienes defienden, pues, un concepto de racionalidad dura, directa o indirectamente, tendrán que defender un algoritmo de la comprensión/interpretación.

Por el contrario, quienes toman en serio argumentos antifuncionalistas o, más en general, antirreductivistas y antieliminativistas, como los tres argumentos reconstruidos en la parte III, podrán defender un concepto de racionalidad más abarcador, más complicado: una *racionalidad incierta*. Examinemos este punto.

A partir del ataque antialgorítmico, podemos afirmar que una persona es racional aunque su pensamiento, en algún sentido, contenga contradicciones, lección que podemos extraer de la paradoja del prefacio. Toda racionalidad incierta es una racionalidad *incompleta*, en el sentido de “incompletamente algorítmica”.

Por otro lado, desde el ataque en contra de la autonomía de la vida mental, puede indicarse que gran parte de los intereses de la razón se configuran en relación con ese *otro* de la razón que es el cuerpo. Toda racionalidad incierta es una racionalidad *encarnada*.

A su vez, el ataque intencionalista nos permite reafirmar que la acción racional necesariamente se vincula con su entorno. Toda racionalidad incierta es una racionalidad *relacional*.

La polémica entre racionalidad dura y racionalidad incierta, que puede reconstruirse como una consecuencia, lateral pero importante, de la polémica

en torno a la analogía funcionalista de la vida mental, es, pues, una polémica entre pensar la razón como reconstruible por un algoritmo no encarnado y pensar la razón como una “racionalidad de razones”, como un despliegue de debates. En el primer caso, el paradigma de racionalidad es una computadora; en el segundo, una persona que frente a una perplejidad, un conflicto o un problema sopesa diferentes razones, *de la manera más abarcadora posible*, incluyendo, por ejemplo, argumentos por analogía como los argumentos funcionalistas y antifuncionalistas. A favor del segundo concepto de razón, de un concepto de razón implicado, o al menos sugerido por los ataques antifuncionalistas, puede alegarse que este concepto de razón *no* desencadena ninguna “sofística de la irracionalidad”. ¿De qué hablamos? A partir de un concepto de racionalidad dura se desencadena una “sofística” previsible se comienza por defender procesos racionales *tan* estrictos, *tan* fuertes que, tarde o temprano, el incumplimiento de estas promesas desproporcionadas hace sospechar de cualquier intento de elucidación racional, por humilde que éste sea. Una vez más asistimos a los efectos de la peligrosa lógica o, mejor, “sofística” del “todo o nada”: o disponemos de una racionalidad dura, de modelos criterios de reconstrucción, de modelos con criterios precisos, fijos y generales, y de programas fundamentalistas de justificación, de justificaciones según las cuales toda creencia justificada se apoya en otras creencias también justificadas hasta llegar a creencias últimas que se autojustifican, que son evidentes en sí mismas. . . o no nos queda más que el “todo vale”. Esto es, el relativismo en sus diferentes versiones, cuando no, el escepticismo. Sin embargo, no hay por qué aceptar esta alternativa, ni el ejercicio de la razón tiene que reducirse a lo indubitable, al cálculo en exacto, a respaldar conclusiones de manera necesaria, al algoritmo, a la “razón funcionalista”.

## BIBLIOGRAFÍA

- Block, Ned (ed.), 1980, *Readings in Philosophy of Psychology*, Harvard University Press, Cambridge (Mass.).
- Putnam, Hilary, 1960, “Mind and Machines”, en Putnam, 1975.
- , 1967a, “The Nature of Mental States”, en Putnam, 1975.
- , 1967b, “The Mental Life of Some Machines”, en Putnam, 1975.
- , 1973, “Philosophy and Our Mental Life”, en Putnam, 1975.
- , 1975, *Mind, Language and Reality. Philosophical Papers II*, Cambridge University Press, Cambridge.
- , 1988, *Representation and Reality*, MIT Press, Cambridge (Mass.).