

Capítulo sexto. La mente es una computadora: el funcionalismo de máquina . . . . .	153
I. La mente es una computadora . . . . .	153
II. La tesis de una máquina Turing universal . . . . .	154
III. Las realizaciones físicas de una máquina Turing . . . . .	156
IV. Los motivos del funcionalismo de máquina . . . . .	158
V. Las intuiciones que subyacen a esos motivos . . . . .	158
VI. La tesis toral: que la mente es una computadora . . . . .	159
VII. La prueba de Turing: ¿en qué consiste? . . . . .	160
VIII. De cuán apropiado es el funcionalismo de máquina . . . . .	161
IX. Escepticismo sobre el funcionalismo computacional: Searle . . . . .	162
X. Limitaciones de los dos tipos de funcionalismo . . . . .	165
XI. Funcionalismo y computación . . . . .	166
XII. Objeciones al funcionalismo . . . . .	172
XIII. Resumen . . . . .	173

## CAPÍTULO SEXTO

### LA MENTE ES UNA COMPUTADORA: EL FUNCIONALISMO DE MÁQUINA

El presente capítulo continúa el desarrollo de la tesis funcionalista de las propiedades psicológicas. El funcionalismo teórico coloca a la causalidad como algo central de las propiedades psicológicas; este funcionalismo puede especificarse, suplementándolo, con una tesis computacional. La tesis que afirma que la mente es (similar a) una computadora intenta llevar a cabo la suplementación de la tesis funcionalista con la tesis computacional; tenemos que examinar entonces cómo se la puede estructurar.

#### I. LA MENTE ES UNA COMPUTADORA

La idea central del funcionalismo de máquina es que las propiedades psicológicas se conciben como propiedades computacionales. Así, por ejemplo, se afirma que el pensamiento es computación. La tesis es que un proceso psicológico central como es el pensamiento se reduce a procesos computacionales o, según dicen algunos de sus críticos, el pensamiento se reduce a manejo sintáctico de símbolos.

En el capítulo precedente examinamos al funcionalismo teórico. Ahora la principal preocupación es examinar cómo propone construir el funcionalismo de máquina a las propiedades psicológicas; cómo, por ejemplo, puede el pensamiento o el razonamiento reducirse a una serie de operaciones entre símbolos. Similarmente, qué sucede con el contenido de las propiedades

psicológicas ¿también se reduce a operaciones intrasimbólicas? Luego, en la causalidad entre los símbolos ¿qué papel juega el contenido? Finalmente ¿cómo acomoda al punto de vista o al *qualia* el funcionalismo computacional? Éstas son algunas cuestiones centrales que debemos tratar en este capítulo.

La tesis que debemos considerar aquí es la tesis de que la realidad de las propiedades psicológicas (léase su eficacia causal) queda asegurada si consideramos que dichas propiedades psicológicas son valores de máquinas Turing y de que la mente (léase el conjunto de propiedades psicológicas) se comprende mejor como un caso de máquina Turing. Aún más: si las propiedades psicológicas se pueden reducir a información computada, entonces la idea de una máquina Turing resultará el mejor aliado de la tesis materialista. Tal es entonces la importancia del presente capítulo.

¿Qué es entonces una máquina Turing? La idea de que la mente es (=) una computadora nos lleva a examinar en primer término la idea de la computación en forma general, es decir, la idea de una máquina Turing universal.

## II. LA TESIS DE UNA MÁQUINA TURING UNIVERSAL

La máquina Turing universal es un autómata probabilista indeterminista que puede computar cualquier operación aritmética. Si tenemos que:

C = Máquina Turing

PM = Propiedades psicológicas

podremos concluir que: C es la mejor teoría de PM.

Esta es la tesis que debemos considerar, la idea de una máquina Turing como una versión del funcionalismo. La máquina Turing tiene los siguientes componentes:

- i. Una cinta dividida en cuadros ilimitada en ambas direcciones.
- ii. Una cabeza que escanea e imprime colocada en alguno de los cuadros.
- iii. Un conjunto finito de estados internos o configuraciones  $q_0, \dots, q_n$ .
- iv. Un alfabeto finito que consiste de símbolos  $b_1, \dots, b_n$ .
- v. Cada cuadro puede tener solamente un símbolo.
- vi. Una tabla de máquina que establece en forma completa y exhaustiva las operaciones o instrucciones admisibles de la máquina (las cuales son deterministas).

Toda computación o procesamiento que puede hacerse por una computadora puede hacerla una máquina Turing.

A partir de una máquina Turing se puede construir una “máquina universal” que es como una computadora de propósito general (no limitada a algún tipo de computación) a la cual se la dota de su tabla de máquina para la función que se desea computar, los valores específicos iniciales o argumentos. Para su entrada se necesitan dos cosas, a saber, la tabla de máquina en alguna notación estándar y los valores que se van a computar. Así se programa a la máquina universal para que lea la tabla de máquina de esta función a computar y lleve a cabo las instrucciones necesarias hasta que concluya la computación.

Si generalizamos la máquina Turing obtendremos la noción de un *autómata probabilístico*. Por ser probabilístico este autómata es indeterminista. Este tipo de máquina puede resultar más adecuado para describir la conducta de un sujeto psicológico. ¿Por qué? Porque en principio puede tomar en cuenta el proceso de decisión del sujeto psicológico.

La máquina Turing es algo abstracto y como tal puede especificarse en forma múltiple. La máquina Turing puede realizarse de muchas maneras diferentes, por ejemplo, en una cinta de plástico o en tarjetas de cartón, etcétera.

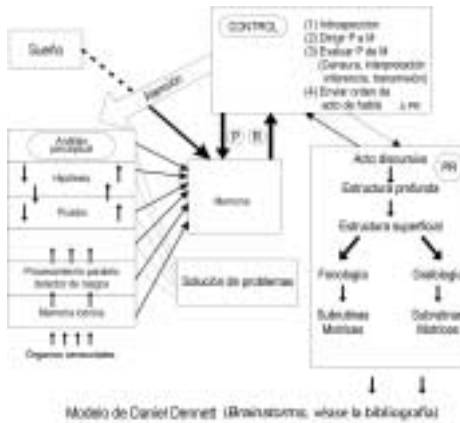
### III. LAS REALIZACIONES FÍSICAS DE UNA MÁQUINA TURING

Una realización física de una máquina Turing incluye lo siguiente:

- i. Los símbolos del alfabeto de la máquina deben adquirir encarnamientos físicos concretos (brochazos de tinta sobre papel o patrones de partículas de acero magnetizadas en una cinta de plástico, etcétera).
- ii. Los símbolos que la máquina imprime en su cinta de salida deben tener formas físicas.
- iii. Los estados internos de la máquina deben realizarse físicamente de la siguiente manera: remplazando las relaciones funcionales o computacionales entre los varios parámetros abstractos mencionados en la tabla de máquina con las relaciones causales apropiadas entre los encarnamientos físicos de esos parámetros.

Desde un punto de vista lógico los estados internos sólo están implícitamente definidos y desde el punto de vista físico no interesan sus naturalezas o su constitución en tanto ésta preserve las conexiones causales correctas. En consecuencia, el estado físico de una máquina Turing es su papel causal. Por lo tanto, se pone de manifiesto una convergencia entre la concepción funcionalista de un estado mental —que vimos antes— como ocupando un papel causal y la idea de un estado físico que realiza el estado interno de la máquina Turing. De lo cual se sigue que la máquina Turing resulta un modelo de la mente. La descripción de la máquina Turing es la relación conversa de la relación de realización o de superveniencia. Y esto significa que dos sistemas que son (*input-output*) equivalentes en sus entradas y salidas pueden no ser realizaciones de (no supervenir en) la misma máquina Turing. Y, por lo tanto, la equivalencia de entradas y salidas no asegura equivalencia psicológica (tal como se apreció en el caso del conductismo expuesto en el capítulo segundo).

Aparte de las entradas y salidas hay el procesamiento interno (algo que el conductismo rechazó por principio) que consiste en la causalidad interna la que se lleva a cabo por mecanismos de procesamiento (misma implementación). Este procesamiento interno se capta en diagramas de flujo. Veamos cómo en el siguiente ejemplo:



El modelo anterior exhibe la ruta que sigue la información del mundo desde los transductores que convierten los datos del mundo en información computable la cual recibe varios tipos de procesamiento que se llevan a cabo en varias cajas negras hasta provocar la salida motora. La pregunta que nos acosará de ahora en adelante es la de si bastan estos diagramas de flujo —y contraflujo— para, por ejemplo, capturar la capacidad de percibir objetos del mundo. Dicho de otra manera, si es suficiente satisfacer un diagrama de flujo como el de Dennett para que ocurra la percepción o, si no basta, y entonces poder percibir es algo extra a la satisfacción de un tal diagrama. Repárese en que estamos hablando de una condición suficiente.

#### IV. LOS MOTIVOS DEL FUNCIONALISMO DE MÁQUINA

Enumeremos algunas de las razones que han inclinado a los teóricos de las propiedades psicológicas a inclinarse en favor del funcionalismo de máquina.

- i. Similaridad.
- ii. Fecundidad.
- iii. Generalidad (no-parroquialismo).
- iv. Carácter científico.
- v. Carácter reproducible.

#### V. LAS INTUICIONES QUE SUBYACEN A ESOS MOTIVOS

¿En qué sentido es funcionalismo? En el sentido de que un proceso cognitivo de una propiedad psicológica puede analizarse usando una función que mapea ciertas entradas con ciertas salidas. Si se toman estructuras simbólicas como entradas, la función  $f$  que ejemplifica el proceso cognitivo se puede especificar como usando reglas que manipulan símbolos. Pero tal vez la manipulación se dé entre estructuras no-discretas como las imágenes o entre pesos de conexión (*connection weights*) de estructuras algebraicas tensadas (*tensor-algebraic*) capturables por alguna propiedad matemática  $p$ . En este último caso

cuando un patrón de actividad es dado como entrada a la red que es el producto tensado de la realización de una estructura símbolo  $i$ , entonces un patrón de salida que realiza  $f(i)$  se produce (en donde  $f$  es una función simbólicamente especificada). La propiedad  $p$  explica cómo  $P$  ejemplifica  $f$ .<sup>146</sup>

146 *Cfr.*, Smolensky, P., "Connectionsism, Constituency and the Language of Thought", en B. Lower y G. Rey (eds.), *Meaning in Mind: Fodor and His Critics*, Oxford, Blackwell, 1991. Hay aquí una intensa polémica entre los que afirman (Fodor, J., "Replies", *Meaning in Mind: Fodor and His Critics*, *cit.*) que el procesamiento co-

## VI. LA TESIS TORAL: QUE LA MENTE ES UNA COMPUTADORA

La idea de que las propiedades psicológicas de un organismo o su psicología es de naturaleza computacional y que por lo tanto puede representarse por una máquina Turing, es decir, que su descripción o programa es constitutiva de que sea mental. Por lo tanto,

Mentalidad = realización en una máquina Turing

Computadora = cerebro + programa.

Ahora la tesis afirma que la mente es (=) una computadora; es una tesis fuerte pues ha pasado de afirmar que el procesamiento computacional es una condición suficiente a la afirmación de que es a la vez necesario y suficiente.

¿Por qué pensar a la mente como una computadora y a las propiedades psicológicas como los estados de una tal máquina? La idea de una computadora incluye dos elementos constitutivos, a saber, una serie de contenidos codificados y un procesamiento de esos contenidos. Los contenidos pueden ser representaciones con actitudes hacia esas representaciones. El procesamiento puede consistir en operaciones lógicas conscientes o no sobre esas representaciones, mismas que deben respetar la estructura de las representaciones (por ejemplo, su composicionalidad).

Esta idea de las computaciones concretas nos lleva a concebir la mente con sus propiedades psicológicas como un caso o ejemplificación de una máquina Turing. Marr señaló tres requisitos que debe satisfacer una tal máquina, a saber:

- i. Ofrecer un análisis de una propiedad psicológica y justificar por qué lo hace.
- ii.

nexionista no puede ser sensible-a-la-estructura y los que afirman que sí puede serlo (Smolensky, P., "Connectionism, Constituency and Language of Thought", *cit.*



Proveer un algoritmo y un dominio o sistema de representaciones sobre las que operan las computaciones especificadas por la teoría.

- iii. Ofrecer detalles de los mecanismos de implementación neuronal en los que se llevan a cabo las computaciones.

Ahora bien, adviértase que la tesis de la máquina Turing ejemplificada no prejuzga del tipo de representaciones que conforman el dominio (ii). Aquí se abre una importante disyunción pues esas representaciones pueden ser de naturaleza simbólica, como lo quiere la tesis de un lenguaje del pensamiento (LDP)<sup>147</sup> o ser redes construidas a partir de unidades eléctricas procesadoras que llevan a cabo computaciones (en paralelo) y de esta manera simulan las actividades cognitivas en las que consisten las propiedades psicológicas. La primera es la tesis simbolista; la segunda es la tesis conexionista.

Hay una enorme disputa en nuestros días entre los simbolistas y los conexionistas. El trabajo experimental parece apoyar una u otra de las variedades de estas teorías según la propiedad psicológica de que se trate.<sup>148</sup> En este sentido el requisito (ii) de Marr se mantiene neutral respecto de los elementos que lo pueden satisfacer o que pueden constituir un tal dominio.

## VII. LA PRUEBA DE TURING: ¿EN QUÉ CONSISTE?

La prueba de Turing constituye una condición suficiente de mentalidad: si dos sistemas son equivalentes en sus entradas y salidas, tendrán el mismo estatus psicológico; en particular, uno será mental en caso que el otro lo sea.

De si la prueba de Turing es más conductista que el funcionalismo de máquina y entonces solamente crea una presunción

147 Expondré las tesis del LDP y la TRM en el siguiente capítulo.

148 Sobre esto puede consultarse a Cummins, R., *The Nature of Psychological Explanation*, Cambridge, MIT Press, 1983.

de mentalidad, sin asegurar que tiene mentalidad. Solamente alcanza a establecer una presunción de similaridad con una mente, ese es su intento teórico, no es el de reducir el pensamiento a un proceso computacional.

## VIII. DE CUÁN APROPIADO ES EL FUNCIONALISMO DE MÁQUINA

Condición de “apropiado”: no bastan entradas y salidas como veremos posteriormente. No basta decir que hemos capturado una propiedad psicológica si tenemos entradas y salidas aceptables para concluir que hemos descubierto la naturaleza de una propiedad psicológica dada. Conformarse con este requisito es aceptar la tesis del conductismo. Seguramente hay algo más que las entradas y salidas, debe haber, por ejemplo, algún tipo de procesamiento y transformación interno y es en ese programa de procesamiento en lo que consiste la propiedad psicológica. Piénsese en el funcionalismo teórico que habla de “aptitud para causar...”; esa aptitud es la que tenemos que descubrir. Aquí encontramos que la tesis computacional pone de manifiesto la necesidad de algo más que las entradas y salidas al suponer que hay un programa (que el modelo de Dennett ejemplifica) complejo, el cual es una parte principal de la propiedad psicológica en cuestión.

Ahora reparemos en lo de “apropiado”, ¿cómo diferenciar dolor de pulpo de dolor humano? ¿Cuánta complejidad? ¿Hay un criterio? Parece que el funcionalismo de máquina no da respuesta a este tipo de preguntas. Pero ¿puede ayudar a resolverlas? Tal vez pueda mediante una suplementación.

El funcionalismo de máquina no responde esas cuestiones cruciales, por el contrario, las evade y en su lugar propone la ecuación “pensar” = “prueba Turing”. Evidentemente esto no basta;

pero no hay que perder de vista la aportación positiva de este tipo de funcionalismo.

¿Cuándo piensa una computadora? “Cuando hace sus tareas tan bien como las hace una persona”. ¿Qué clase de condición es ésta? ¿Responde la cuestión? Encontrar una similitud no basta. Tiene que responder a las siguientes preguntas:

- i. ¿Es suficiente? No resultan de ella las propiedades psicológicas.
- ii. ¿Es necesaria? No tenemos idea, por ejemplo, de la propiedad de la conciencia.
- iii. ¿Es una condición de pensar o de ser una propiedad psicológica?

## IX. ESCEPTICISMO SOBRE EL FUNCIONALISMO COMPUTACIONAL: SEARLE

Consideremos el argumento del cuarto Chino: Searle imagina el siguiente ejemplo:

Imaginemos que yo, un hablante no-chino quedo encerrado en un cuarto plagado con símbolos chinos en cajas. Se me proporciona un manual de instrucciones en español para hacer corresponder símbolos chinos con otros símbolos chinos y para ofrecer montones de símbolos chinos en respuesta a otros montones de símbolos chinos introducidos en el cuarto a través de una pequeña ventana. Sin que yo lo sepa, los símbolos introducidos por la ventana se llaman preguntas. Los símbolos que produzco se llaman respuestas. Las cajas de símbolos que tengo se llaman bases de datos y el libro de instrucciones en español se llama programa. La gente que me hizo preguntas y diseñó el libro de instrucciones se llaman programadores y yo soy la computadora. Imaginemos que llego a ser tan bueno en producir símbolos y los programadores llegan a ser tan buenos al escribir el programa que eventualmente mis “respuestas” a las “preguntas” llegan a ser indistinguibles de las de un hablante nacido en

China. Apruebo el examen *Turing* para entender chino. Pero igual, no entiendo una palabra de chino —y este es el punto de la parábola— si yo no entiendo chino con base en el programa para comprender chino, tampoco lo logra ninguna computadora digital sobre esa base pues ninguna computadora digital tiene algo que yo no tenga.

Searle resume su argumento de la siguiente manera:

Premisa 1. Los programas son formales (sintácticos).

Premisa 2. Las mentes tienen contenidos (semántica).

Premisa 3. La sintaxis no es suficiente para la semántica.

Conclusión: las mentes no son programas.

De esta forma Searle intenta refutar la tesis fuerte de la inteligencia artificial que dice que las mentes son (=) idénticas a las computadoras o, dicho de otra manera, que las mentes son programas de cómputo implementados en cerebros y que basta un programa de cómputo para tener una propiedad psicológica. La razón profunda de que las mentes no puedan ser computadoras es que la propiedad de ser una computadora no es algo intrínseco sino algo relativo al observador o usuario quien le asigna una interpretación. Es posible asignarle una interpretación computacional al cerebro pero ello no convierte al cerebro en una computadora.

Searle confía en que su argumento del cuarto chino refuta que el cerebro o la mente sea una computadora porque la semántica no es intrínseca a la sintaxis y las computadoras son máquinas sintácticas. En verdad, según vimos antes, la máquina Turing pasa de un cuadro a otro en la cinta en virtud de las propiedades sintácticas del símbolo. Por otra parte, el cerebro, no es intrínsecamente una computadora porque la sintaxis no es intrínseca a la física, es decir, no hay hechos físicos intrínsecos que hagan a un sistema computacional.

Pero Searle concede la tesis débil de que la mente es una computadora, a saber, que podemos interpretar las propiedades psicológicas como propiedades computacionales y representarlas

en diagramas de flujo. Lo que Searle niega es que haya la identidad ontológica <propiedades psicológicas = máquina Turing>; que baste la manipulación sintáctica para que haya mentalidad o psicología.

Entonces, no basta la manipulación sintáctica para tener la propiedad psicológica de comprender un lenguaje. ¿pero quién supone que baste la manipulación sintáctica para que haya lenguaje u otra propiedad psicológica? En el siguiente capítulo veremos algunas versiones de la TRM que recoge la tesis de que la mente es una computadora y veremos que no resulta afectada por el argumento del cuarto chino de Searle.

Hay parroquialismo: ¿las propiedades psicológicas resultan solamente de una organización biológica compleja? ¿Podemos conformarnos con afirmar que el pensamiento se reduzca a una manipulación de símbolos? Pero más importante aún: ¿quién sostiene que baste la manipulación de símbolos para que haya pensamiento? Todo esto reaparecerá en el capítulo sobre la naturaleza de las representaciones mentales.

Tampoco se conoce cómo una organización neuronal compleja (de redes neuronales, por ejemplo) puede producir propiedades psicológicas: ¿se sigue de esto la falsedad del materialismo? Aquí aparece de nuevo la falacia epistemológica: si no conozco ahora que las propiedades psicológicas resultan de la base cerebral entonces concluyo que no son ontológicamente materiales. Este es de nuevo el dogma de que las cuestiones ontológicas se deciden con base en la evidencia epistemológica.

La esperanza del funcionalismo de máquina consiste en reducir la mente a información computada. ¿Cómo puede llevarse a cabo esta reducción? En el siguiente capítulo veremos un tipo de respuesta.

## X. LIMITACIONES DE LOS DOS TIPOS DE FUNCIONALISMO

El funcionalismo teórico deja abierto si debe haber una mediación entre el cerebro y la conducta y cuál tipo de mediación requiere el ejercicio de las propiedades psicológicas. Y si acepta que debe haber tal mediación deben establecer cuál es el elemento psicológico adecuado que debe mediar entre el cerebro y la conducta.

Debemos observar aquí que la teoría funcionalista surgió como vimos antes de la teoría materialista y de las presiones de la realización múltiple. Por ello mismo se conformó con acomodar esas dificultades y no reparó en la necesidad de contemplar una mediación entre cerebro y conducta. La necesidad de una tal mediación surge de otra presión, a saber, de la necesidad de que una teoría de las propiedades psicológicas sea una teoría de seres humanos o personas y no solamente de animales, zombies, autómatas o robots. Estos últimos pueden tener algo similar a las propiedades psicológicas pero no llegan a ser personas. La explicación propuesta dice que lo distintivo de las personas es tener capacidades psicológicas reales, es decir, que esas capacidades medien efectivamente entre el cerebro y la conducta y que además, el conjunto de esas propiedades psicológicas produzca la propiedad superior de la personeadad. A esos mediadores se les llama representaciones. Tenemos por consiguiente que abocarnos a considerar la naturaleza de las representaciones en el siguiente capítulo.

Hay una complejidad constante en ambos tipos de funcionalismo, el computacional y el teórico, a saber, cómo es posible que una propiedad abstracta tenga fuerzas causales. Sobre esto habré de volver.

## XI. FUNCIONALISMO Y COMPUTACIÓN

Será bueno recordar aquí algo de lo dicho anteriormente acerca del funcionalismo para enlazarlo con el funcionalismo de máquina. El funcionalismo es una tesis acerca de las propiedades psicológicas y dice que las propiedades psicológicas son (=) funciones, propiedades funcionales. Luego dice que las propiedades funcionales están constituidas por las relaciones causales de unas propiedades psicológicas con otras y con las entradas sensoriales y las salidas conductuales tal como se exhibe, por ejemplo, en el diagrama de Dennett que incorporamos anteriormente. Además de los dos tipos de funcionalismo, a saber, el funcionalismo teórico y el funcionalismo de máquina hay dos variedades adicionales, a saber, el psicofuncionalismo que es una teoría empírica acerca de las propiedades psicológicas y el funcionalismo analítico o conceptual que es una teoría *a priori* acerca de los términos mentales o psicológicos. De acuerdo con el funcionalismo analítico la propiedad psicológica ya está capturada por el análisis desde el comienzo y lo que falta es encontrar su(s) forma(s) de realización en el cerebro o en una estructura alambreada. En este tipo de funcionalismo no hay hiato explicativo pues al analizar la propiedad psicológica del dolor, por ejemplo,<sup>149</sup> el siguiente paso será encontrar la propiedad física que lo realiza y con ello se habrá satisfecho la pregunta

Q: ¿Por qué la estructura cerebral C realiza el dolor y no una cosquilla o un calambre o nada?

La respuesta será contundente, a saber, “porque eso quiere decir «dolor»”. En cambio, para el psicofuncionalismo aún no conocemos lo que es la propiedad psicológica y para llegar a conocerla tenemos que recurrir a algún tipo de evidencia empí-

149 Recuérdese el análisis causal del dolor en el capítulo anterior.

rica; contamos con las notas o estereotipo que el sentido común atribuye al dolor pero esas notas no nos dan su naturaleza, como tampoco nos la da para el caso del agua el estereotipo líquido incoloro, transparente, que apaga la sed, etcétera. Para el psicofuncionalismo sí surge el hiato explicativo y por ello esta tesis busca trazar un puente entre el estereotipo de la propiedad “dolor” y su realización de modo que haya una respuesta informativa a la cuestión Q; en el caso del psicofuncionalismo la respuesta será sustantiva e informativa, a saber, “porque ‘dolor’ tiene esta estructura cognitiva y/o neurofisiológica según la evidencia E”.

Propongamos el ejemplo de un autómatata que detecta paridad con dos estados E1 y E2; dos entradas “1” y “-” y dos salidas “par” e “impar”. La tabla del autómatata describe dos funciones una de entrada y estado a la salida y otra de entrada y estado al siguiente estado. Cada cuadro encierra dos condicionales especificando la salida y el siguiente estado, dados ambos, el estado actual y la entrada. Así, por ejemplo, el cuadro superior de la izquierda nos da los dos condicionales siguientes: (1) Si la máquina comienza en S1 y ve un “1”, dice “impar” (indicando que ha visto un número impar de “1”); y (2) Si la máquina comienza en S1 y ve un “1” pasa al estado S2. Toda la máquina consta de ocho condicionales.

Si preguntamos qué es S1, la respuesta será que S1 es ese conjunto de relaciones que aparecen completas en la tabla. La naturaleza de S1 está constituida exhaustivamente por esos ocho condicionales. Así podemos decir que S1 es (=) estar en los dos primeros estados que se relacionan uno con otro y con las entradas y salidas como sigue: estar en uno de los estados y recibir una entrada “1” resulta en ir al segundo estado emitiendo “impar”; y así en lo demás, en lo que resta de la tabla. Piénsese de manera similar en el diagrama de la percepción de Dennett que apareció antes.

Ahora podemos establecer los siguientes puntos:



1. La naturaleza de una propiedad psicológica o de un estado mental es semejante a la naturaleza de un autómata en que se agota en las relaciones con otras propiedades psicológicas y con las entradas y salidas.
2. El método para definir los estados de un autómata sirve para definir a las propiedades psicológicas.
3. La propiedad psicológica es una propiedad abstracta, de segundo orden, que consiste en tener o dar cabida a otras propiedades, por ejemplo, neurofisiológicas o electrónicas, que mantienen relaciones unas con otras. A estas otras propiedades se las llama las realizaciones de las propiedades funcionales o alternativamente, que las propiedades psicológicas supervienen de esas propiedades neurofisiológicas, etcétera.
4. Una propiedad psicológica puede realizarse o supervenir de diferentes maneras de manera que S1 puede ser un estado mecánico si su tabla consta de tuercas, ruedas, etcétera, o puede ser un estado electrónico si su tabla consta de alambres, circuitos, electricidad, etcétera.
5. Entonces S1 no es ni químico, ni físico, etcétera, sino un estado abstracto tal como lo definimos con base en su tabla.<sup>150</sup>
6. Hay tres elementos a considerar, a saber, el estado mental o propiedad psicológica con contenido, el algoritmo que computa una función y la realización de esa función en algún compuesto físico, biológico, etcétera.
7. Así como una propiedad psicológica puede realizarse en diferentes tipos de compuesto o supervenir de diferentes tipos de “materia”, un estado físico puede realizar diferentes estados funcionales en diferentes máquinas; o bien, un estado físico puede supervenir en diferentes estados funcionales de máquinas diferentes.
8. Por otra parte, una misma función puede ser computada por varios algoritmos diferentes.

150 Adviértase la similitud con el funcionalismo teórico del capítulo anterior.

9. La tesis completa dice que de entrada hay que funcionalizar<sup>151</sup> la propiedad psicológica; una vez funcionalizada hay que encontrar un algoritmo que la compute y que da por resultado la propiedad psicológica con todas sus características (una duplicación de la propiedad psicológica) y finalmente hay que descubrir el compuesto o materia en el que se realiza la función.<sup>152</sup>

Supongamos que T es la teoría psicológica que define “dolor” como:

x tiene dolor sii  $\exists P Q$  [cortándose con un punzón causa P & P causa ambos Q y proferir “ayy” y x tiene P]

Tener dolor es estar en el primero de dos estados, el primero de los cuales consiste en cortarse con un punzón, el cual a su vez causa el estado de experimentar una sensación aguda/punzante y proferir “¡ayy!”

Y más generalmente, si en la teoría psicológica T T17 es el término mental de “dolor” y F designa otros términos mentales, tendremos la siguiente definición:

x tiene dolor sii  $\exists F1... \exists Fn$  [T(F1...Fn) & x está en F17]

El funcionalismo caracteriza lo mental en términos de lo no-mental, en términos que involucran cuantificación sobre la realización de otros estados mentales (sin mención explícita de ellos). El funcionalismo caracteriza lo mental en términos de estructuras abstractas que tocan la realidad al nivel de las entradas y salidas justo como lo hacían las oraciones Ramsey en la teoría causal de Lewis.<sup>153</sup>

151 Usaré este adverbio para indicar que las propiedades psicológicas se reducen a una función.

152 Estas son las tres tareas señaladas por Marr, según vimos antes.

153 Según vimos en el capítulo anterior.

Las propiedades psicológicas funcionalizan o se reducen, de la manera anterior, a propiedades funcionales, es decir, a propiedades abstractas o de segundo orden que pueden realizarse o no en propiedades primarias como en lo físico (incluido lo biológico y lo químico) o lo electrónico, lo espiritual, etcétera. En esta medida el funcionalismo no nos da una respuesta que compita con el dualismo o el materialismo pues no está comprometido a un tipo de compuesto (*stuff*) sino que se mantiene neutral entre los varios tipos de compuestos. En este sentido el funcionalismo no es una afirmación metafísica que responda a la pregunta “¿qué es?” o “¿de qué está hecha o compuesta *p*?” sino una elucidación a la pregunta “¿cuál es la constitución o estructura de la propiedad *p*?” o a la cuestión “¿cómo es la propiedad *p*?”.<sup>154</sup> De esta suerte podemos tener un materialismo funcionalista o un espiritualismo funcionalista dependiendo de otra evidencia adicional que se ofrezca para poder afirmar que una propiedad funcional *p* se realiza en una materia espiritual o en una física.

En tratándose de las propiedades psicológicas tenemos entonces varios niveles de descripción, a saber:

1. Un nivel físico de estados particulares, por ejemplo, de un cerebro.
2. Un nivel biológico o neurofisiológico.
3. Un nivel abstracto con la funcionalización de la propiedad psicológica y el algoritmo que la computa.
4. Un nivel comunitario con el contenido que tiene la propiedad psicológica en la psicología popular.

Recordemos que en el funcionalismo teórico se utiliza la tesis funcional para analizar el contenido de las propiedades psicológicas y para posteriormente llevar a cabo la identificación de la

<sup>154</sup> Pero en otro sentido sí es metafísica pues nos dice algo que tiene que ver con lo que es la propiedad y eso que nos dice no es epistemológico ni semántico.

propiedad psicológica con una propiedad física. El propósito allí fue de análisis del contenido. Pero puede avanzarse y llevar la tesis del funcionalismo al nivel más ambicioso de proponerla como una teoría empírica computacional y/o informacional.<sup>155</sup> A esta última se la conoce como la teoría representacional de la mente (TRM) que ha aparecido antes y será considerada en el siguiente capítulo.

Consideremos al funcionalismo conceptual que pretende darnos el contenido o el significado de los términos que designan a las propiedades psicológicas. En la definición de dolor expuesta antes supongamos que aceptamos que dolor = S17 en donde S17 puede ser un estado psicofuncional o un estado cerebral. Ambos lados del signo de identidad aprehenden al referente según un modo de presentación diferente. ¿Cuál será el modo de presentación asociado con “dolor”? Se dirá que una propiedad mental, como el aspecto fenomenal del dolor. Este aspecto a su vez será algo físico o funcional. Si no acepta que ya tiene el contenido del dolor tendrá que aceptar una nueva identidad como “aspecto fenomenal = S13”. Pero tenemos de nueva cuenta dos aspectos y de nuevo la pregunta “¿cuál es el modo de presentación de ‘aspecto fenomenal’?” que conlleva la amenaza de un regreso vicioso. Si se quiere detener ese regreso hay que aceptar una identidad *a priori* entre las propiedades psicológicas y las propiedades funcionales: este es el funcionalismo analítico. Pero entonces ese tipo de identidades *a priori* no serán aceptadas por un fisicalista o un psicofuncionalista.

Una forma de evadir las identidades *a priori* consiste en introducir una teoría de la referencia holista la cual establece que el significado de los términos mentales o psicológicos se dan en términos de toda la teoría psicológica. Otra salida consiste en afirmar con Loar<sup>156</sup> que los diferentes modos de presentación son

155 Este es el desafío que enfrenta Fodor, J., *Psychosemantics*, Cambridge, MIT Press, 1987, *The Elm and the Expert*, Cambridge, MIT Press, 1994.

156 Véase, Loar, B., “Phenomenal States”, en J. Tomberlin (ed.), *Philosophical Perspectives 4*, California, Ridgeview Press, 1990.

los diferentes roles funcionales de “dolor” y de “S17” de modo que la relación de referencia se da no-mediada por algún modo de presentación.

¿Quiere esto decir que los términos funcionales de las propiedades psicológicas no denotan clases naturales (nómicas y proyectibles)? Las clases naturales surgen de las propiedades funcionales realizadas en algún compuesto físico. El funcionalismo *a priori* o analítico que nos da las descripciones abstractas de segundo orden se limita a ofrecer una teoría de los conceptos que pretende limitar o desplazar al psicofuncionalismo empírico.

Pero si el funcionalismo nos da el contenido o significado de los términos de propiedades psicológicas es decir, una semántica procedual o una semántica del papel conceptual, esta teoría se verá expuesta a las objeciones de Putnam<sup>157</sup> y Fodor.<sup>158</sup>

## XII. OBJECIONES AL FUNCIONALISMO

Recordemos que hasta ahora han aparecido dos tipos principales de funcionalismo, a saber, el funcionalismo teórico y el funcionalismo de máquina. Luego, hay dos variedades: el funcionalismo analítico y el psicofuncionalismo. A continuación, para concluir enumeraré ocho objeciones al funcionalismo pues sus respuestas ayudarán a discriminar y evaluar las dos variedades de teoría funcional de las propiedades psicológicas, a saber, entre el funcionalismo analítico y el psicofuncionalismo:

1. Que es una teoría parroquial. Esta objeción dice que la construcción funcional o funcionalización de las propiedades psicológicas las vuelve parroquiales pues no permite otras variedades por descubrir puesto que ya ha definido a la propiedad psicológica de que se trate. Esto no es verdad

157 Putnam, H., *Representation and Reality*, Caambridge, MIT, 1988.

158 Fodor, J., “Replies”, *cit.*

del psicofuncionalismo pues esta tesis dice que la propiedad psicológica es aquello que la evidencia empírica encuentre y si encuentra varias propiedades diferentes deberá ajustar su taxonomía.

2. Que por el contrario es una teoría permisiva.
3. Que tiene problemas con la intencionalidad. Algo hay de esto y lo examinaremos en los capítulos finales.
4. Que no puede capturar las propiedades fenomenales. Esta es una objeción central y le dedicaremos un espacio especial en el penúltimo capítulo y también en el capítulo final.
5. Que origina el hiato explicativo. Como vimos más arriba, esto sólo es verdad del funcionalismo analítico.
6. Que las propiedades funcionales son propiedades de segundo orden y como tales no pueden hacerle justicia al carácter causal y explicativo de las propiedades psicológicas. La objeción repara en que una propiedad abstracta no parece poder contar con fuerzas causales y en verdad no puede pero es la propiedad especificada la que cuenta con fuerzas causales; de nuevo esta objeción parece ir en contra del funcionalismo analítico que es el que “cierra” la propiedad psicológica a toda nueva evidencia.
7. Que implica holismo. Es el funcionalismo analítico el que está propenso a caer en holismo; pero también el tratar de evitar las objeciones cae en holismo.
8. Que requiere de una noción de contenido estrecha.<sup>159</sup>

### XIII. RESUMEN

Hemos partido de la teoría computacional para desembocar en una nueva evaluación de la tesis de las propiedades psicológicas como propiedades computacionales. Un sinnúmero de objeciones se alzan en contra de esta equivalencia. De entre todas

159 Sobre el contenido estrecho véase el capítulo octavo.

ellas emerge una muy singular: la idea de que al funcionalizar a las propiedades psicológicas emerge la impresión de que se escapa algo esencial, intrínseco y de que por lo tanto, fracasa la funcionalización y ya no se pueden desarrollar los otros dos pasos del programa computacional, a saber, descubrir el algoritmo con el que se computa la función y descubrir la forma de realización de la propiedad psicológica. Sobre estas objeciones habremos de volver en los dos capítulos finales. En el siguiente capítulo me ocuparé de una tesis que incorpora al funcionalismo computacional, la tesis del pensamiento de Turing, la TRM y una tesis de psicología racionalista.