

Jesús F. Baceta V.



**SOBRE LA FUNCIÓN ANÓMALA
DE LA MENTE**

RESUMEN

Se expone la teoría de los estados funcionales de Putnam y se muestra en qué consiste su neutralidad ontológica. Se consideran las críticas del propio Putnam al funcionalismo y su falta de adecuación empírica para una explicación satisfactoria de la relación entre mente y cuerpo. Luego, se argumenta a favor del funcionalismo gracias a algunas pistas que proporciona el propio Putnam, indicios que retrotraen al monismo ontológico de Davidson.

Palabras clave: funcionalismo, ontología, monismo.

ABSTRACT

It describes the theory of functional states of Putnam and clarifies why it is neutral from the ontological perspective. Are considered critical to functionalism Putnam himself and his lack of empirical adequacy for a satisfactory explanation of the relationship between mind and body. Then, we argue in favor of functionalism through clues provided by Putnam himself, and reminiscent of Davidson's ontological monism.

Keywords: functionalism, ontology, monism.

Un estado funcional es una configuración única de información en un programa o máquina que produce una determinada salida (*output*) ante una determinada entrada (*input*) y que pasa a otro estado funcional. Por ejemplo, consideremos un videojuego que acepta como entrada cierto tipo de monedas y, específicamente, dos monedas para comenzar el juego. Podemos distinguir dos estados internos a la máquina. El estado en que la máquina pide dos monedas y el estado en que la máquina pide sólo una moneda. En la siguiente tabla ilustramos los estados funcionales:

Estado inicial	Entrada	Salida	Estado final
Pide 2	2 monedas	0, No pide, Juego	Pide 2
Pide 2	1 moneda	1, Pide una	Pide 1
Pide 2	0 monedas	2, Pide 2	Pide 2
Pide 1	2 monedas	1, Retiene 1	Pide 1
Pide 1	1 moneda	0, Juego, no pide	Pide 2
Pide 1	0 monedas	1, pide 1	Pide 1

Esta tabla define los estados funcionales del juego: hay dos estados diferentes, "Pide 1" y "Pide 2", que pasan a otros estados o mantienen el mismo estado dependiendo de las diferentes entradas. A cada estado inicial corresponde uno, y sólo un, estado final y las entradas y salidas caracterizan la función. La definición de los estados funcionales de un sistema es independiente de cómo opera físicamente la máquina y los materiales que la componen; pudo, fácilmente, como veremos enseguida, definir una máquina abstracta, una máquina ideal de Turing.

Las máquinas de Turing pueden entenderse como una cinta infinita dividida en cuadros con símbolos, donde un dispositivo (que representamos con el símbolo 'V') lee cada uno de los símbolos en los cuadros (entradas) y, de acuerdo a la definición, escribe uno por uno cada cuadro (salidas), luego, se mueve en una dirección determinada

a un cuadro anterior o posterior. El artificio tiene un funcionamiento totalmente mecánico y secuencial.

Nuestra definición de la máquina de videojuegos puede presentarse abstractamente así: se cambia '0' por, digamos, 's0'; '1' por 's1' y '2' por 's2', y llamamos a E1 y E2 los estados de la máquina abstracta en clara correspondencia con los estados de la máquina de videojuego:

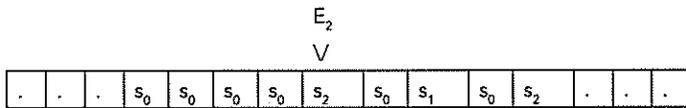
Estado inicial	Símbolo leído	Nuevo símbolo	Nuevo estado
E_2	s_2	s_0	E_2
E_2	s_1	s_1	E_1
E_2	s_0	s_2	E_2
E_1	s_2	s_1	E_1
E_1	s_1	s_0	E_2
E_1	s_0	s_1	E_1

Lo cual se simplifica en la siguiente tabla que proporciona las posibles transiciones de estados, (nuevo símbolo que escribe, dirección de la transición en la cinta y estado final), por un cruce de los estados posibles con las posibles entradas:

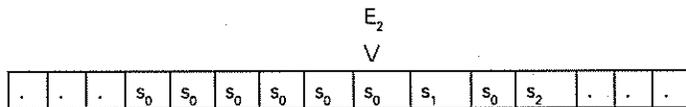
Entrada Estados	s_0	s_1	s_2
E_1	$s_1 \rightarrow E_1$	$s_0 \rightarrow E_2$	$s_1 \leftarrow E_1$
E_2	$s_2 \rightarrow E_2$	$s_1 \rightarrow E_1$	$s_0 \rightarrow E_2$

Veamos cómo funciona la máquina (los comentarios entre corchetes son una analogía con la máquina de videojuego):

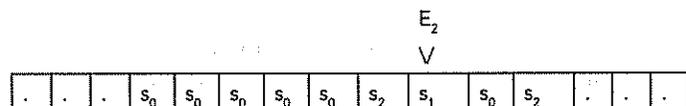
[Aquí pide 2 y entran dos. Juega.]



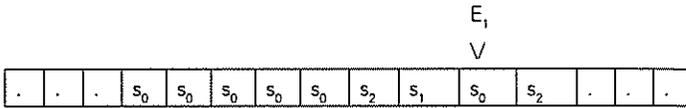
[La salida es un estado donde pide 2 y, en este caso, nada recibe.]



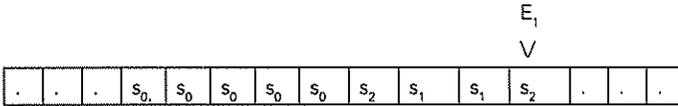
[Como nada recibe, la salida es otro estado donde sigue pidiendo 2, pero le dan 1.]



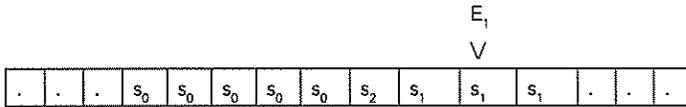
[Como necesita otra moneda para el juego, la salida es un estado donde pide 1, pero en este caso nada se le dio].



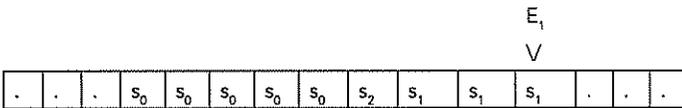
[Por supuesto, sigue pidiendo 1, pero ahora le dan 2]



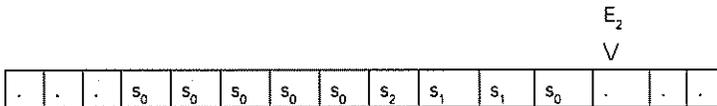
[Retiene 1 y juega]



[Así pide 1 y le dan 1...]



[... juega. La salida es un estado donde pide 2...]



Esta es una máquina con cinta infinita que nunca se detiene y que llena la cinta de los símbolos s_0 , s_1 y s_2 según la definición.

La idea de presentar los dos ejemplos análogos, la máquina "concreta" y la "abstracta", es la de hacer notar que la definición de máquina de estado funcional es neutra desde el punto de vista ontológico: no requiere que la máquina sea concreta o abstracta, es decir, las máquinas funcionales no necesariamente tienen que ser reales. Dijo Putnam:

Esta concepción [el funcionalismo] es completamente neutral con respecto a la cuestión de si esa "forma de funcionar" es explicable en términos físicos (...). Yo siempre rechacé la tesis de que nuestros predicados mentales son reducibles a predicados físicos (...). Así que, en ese sentido, soy neutral con

respecto a la cuestión del materialismo (...). Mi posición se podría describir, quizá, como computacionismo, pero no como materialismo.¹

Si los estados mentales son análogamente funcionales, entonces son independientes de los estados cerebrales concretos que los producen. En este sentido presentamos el ejemplo con una máquina abstracta y no con supuestos estados mentales; los estados mentales son estados neurales, epifenoménicos o espirituales, en todo caso, no son observables en forma directa, son internos. Dijo Putnam:

Lo que el funcionalismo añade a la teoría de Aristóteles es la idea de que nuestra forma de funcionar a nivel psicológico, nuestros estados mentales, pueden describirse con el lenguaje que se usa para describir el funcionamiento de un computador².

Como la tesis funcionalista no tiene compromisos ontológicos y proporciona una plena autonomía de las investigaciones científicas, particularmente, de la investigación neurofisiológica, se le ha considerado como un marco lingüístico adecuado para la fundamentación, desde el punto de vista filosófico, de la investigación en las ciencias cognitivas³.

El conductismo de antaño no pudo explicar, en sus términos, la falta de correspondencia entre los estados mentales, imágenes, algún tipo de sentimientos, o lo que sea, con algunas conductas definidas; tampoco pudo aclarar la posibilidad de que dos individuos tengan el mismo patrón de comportamiento con estados mentales diferentes. Como respuesta al conductismo, surgió la propuesta de Place y Smart⁴, la teoría de la identidad mente-cerebro, o *teoría de la identidad mente-cuerpo*, o *teoría de la identidad de la mente*. En términos muy sencillos, la teoría de la identidad postula que los estados mentales son internos (en contra del conductismo) y que son de un tipo idéntico a los estados de naturaleza física-neurofisiológica, es decir, a cada estado mental (por ejemplo, el dolor) le corresponde un tipo de suceso idéntico en el entorno neurofisiológico (por ejemplo, la estimulación de las fibras cerebrales en ciertas localizaciones del cerebro). El dolor y los estados mentales en general, son fenómenos internos y externos (físicos) a la vez. Y aquí es donde aparece el funcionalismo de Putnam. Putnam rechaza la teoría de la identidad. ¿Por qué la rechaza? Porque,

1 Putnam, H. (1989). "An Interview with Professor Hilary Putnam: The Vision and Arguments of a Famous Harvard Philosopher". *Cogito*, 3, 85-91; p. 86. (De aquí en adelante, todas las traducciones son del autor).

2 Idem.

3 Ver artículo: "Functionalism" en Wilson R. A. y Frank C. Keil (Ed.). (1999). *The MIT Encyclopedia of the Cognitive Sciences*. Cambridge, Massachusetts: The MIT Press.

4 Vid. Lowe, E. J.: *An Introduction to the philosophy of mind*. Cambridge, Massachusetts: Cambridge University Press. 2004. Pp. 44 y ss.

simplemente, desde su perspectiva la teoría de la identidad no puede explicar cómo, por ejemplo, el estado mental de "sentir dolor" se pueda producir en especies muy distintas que tienen una naturaleza físico-neurológica diferente. De ahí la postulación, por parte de Putnam, de la tesis funcionalista que explica, como hemos visto, que distintos estados físicos pueden producir los mismos estados mentales. En el ejemplo vimos que nuestras máquinas pueden cambiar al mismo estado mediante distintas entradas y estados iniciales. Más de un estado físico puede producir el mismo estado mental, pero no tiene porque existir una identidad entre los distintos conjuntos de estados físicos que produce el mismo estado mental.

Ahora bien, nuestro buen filósofo nos mostró, con empeño, que nuestras representaciones mentales no dependen exclusivamente de factores internos; sus famosos ejemplos del agua como H₂O en un mundo y como XYZ en otro, o del aluminio y el molibdeno, donde se mantienen los estereotipos de las comunidades lingüísticas⁵, evidencian las influencias externas en nuestro mundo interno. No obstante, nos dice que nuestra comprensión y la noción de verdad ideal racional depende de la identificación de nuestros estados mentales con cierto tipo de estados funcionales pero, justamente, el ejemplo del agua en aquellos dos mundos o el ejemplo del aluminio y el molibdeno, nos muestra que dos personas con la mismas entradas (los mismos estereotipos y uso de los términos) y la lectura de los mismos símbolos (la percepción de los objetos) pueden tener ideas diferentes, esto es, que las ideas no son idénticas a los estados funcionales. Puede ser que los datos que obtengo de la comunidad y de mis percepciones para la apreciación del aluminio y el molibdeno sean los mismos y, sin embargo, puedo concluir en algunos casos que "Ese pedazo de metal es aluminio" y en otros que "Ese pedazo de metal es molibdeno" que son, claramente, ideas diferentes. Es decir, nuestros estados mentales no son meramente estados internos, sino que se conforman por la influencia de la comunidad y el medio ambiente./

Putnam rechaza su tesis funcionalista de los estados de la mente, que no es más que la metáfora computacional que dice que nuestros estados son similares a máquinas de Turing, independientes de los sistemas concretos físicos o neurofisiológicos, porque nuestros estados mentales no son meramente estados internos, sino que se conforman por la influencia de la comunidad y el medio ambiente y, dado que los

5 Putnam, H. (1975). "The Meaning of 'Meaning'". En *Mind, Language and Reality*. Op. Cit.; p. 225.

estados funcionales son plenamente estados internos, no es posible identificar estados mentales y funcionales. Dijo Putnam:

...el funcionalismo sostenía que los estados mentales no pueden reducirse a simples estados físico-químicos (...); ahora arguyo que los estados mentales tampoco pueden reducirse a estados computacionales⁶.

Esa concepción [el funcionalismo], al ver el lenguaje como algo dentro de la cabeza (o del ordenador) inevitablemente hace totalmente misterioso cómo cualquier ítem del lenguaje pueda referirse a lo que está fuera de la cabeza, o fuera del ordenador⁷.

El funcionalismo es una apasionante tesis propuesta por Putnam; es excitante porque es neutra desde el punto de vista ontológico; esto quiere decir que puede ser aplicada tanto a cierto tipo de estados internos de objetos abstractos como concretos. Es un claro contraejemplo de la teoría de la identidad de la mente y, también, resulta fascinante porque sostiene que nuestras mentes, de alguna forma, trabajan como máquinas de Turing. No obstante, la tesis, pese a ser muy fructífera para los computistas, falla en explicar, según Putnam, las relaciones de los estados internos de la mente con el mundo exterior al no posibilitar ninguna conexión con este último. Trataré, en lo que sigue, de reivindicar al funcionalismo gracias algunas pistas que nos proporciona el propio Putnam, indicios que, a nuestro parecer, nos retrotraen al monismo ontológico de Davidson.

La mente, ese maravilloso mundo interno, se creyó trascendente a nuestro efímero cuerpo. Así lo conjeturó Platón, al sostener la transmigración de las almas; Aristóteles, al postular la trascendencia del hálito que sostiene toda la vida; los medievales, al institucionalizar la discriminación del cuerpo a favor de la eterna inmanencia del alma; Descartes, con su fascinante interacción entre la mente y el cuerpo; Leibniz, con el rescate del pneuma aristotélico, sus mónadas y su armonía preestablecida. El perenne contraste es entre un inseguro universo percedero, ante el mejor de los mundos posibles, un mundo que no encaja en la horma del espacio y el tiempo. La oposición entre el miedo propio de quien se enfrenta a los contundentes objetos de nuestro universo físico y la ilusión del temple que proporciona el mundo propio de las entidades abstractas. Bajo esta hipótesis, ha de ser la mente, el alma, la consciencia, el mundo interno, lo inmaterial, el espíritu, el vaho sublime y etéreo que se distingue de la percedera y finita materia. Una dualidad ontológica entre el mundo ideal y el

6 Putnam, H. (1989). "An Interview... Op. Cit.; p. 87.

7 Putnam, H. (1992). "Replies." *Philosophical Topics* 20.1. pp. 347-408; p. 356

universo material; entre el inmutable mundo real de las formas y el aparente e inestable mundo de la multiplicidad sensible. La dualidad por años desvió la investigación y la llenó de razonamientos basados en premisas de carácter *a priori*; una de ellas fue la aceptación de la propia dicotomía. La hipótesis analítica propuesta por la distinción constituyó un gran escollo para la investigación empírica sobre la mente. ¿Cómo tantear la experiencia de lo mental, si la mente la trasciende? ¿Si la existencia del alma es independiente del universo, cómo hurgar en nuestro mortal y transitorio cuerpo sobre su naturaleza? Solo la razón puede aproximarnos a la mente, si hemos de partir del supuesto analítico según el cual ésta es ideal. La ciencia puede dar cuenta sobre la diferencia entre el peso de una lágrima y el de una gota de sangre; pero solo la mente puede proporcionar el verdadero significado de su diferencia. Se originó, entonces, una peculiar forma de especulación: la mente se relaciona con nuestro cuerpo, pero su naturaleza es ajena a él, porque es inescrutable solo por métodos empíricos.

La mente, desde los albores de la humanidad, es algo que pretende diferenciarnos de otras especies; se asume que existe de alguna manera y, aunque no sabemos qué es, se sospechó que sin ella no se puede asir el contenido objetivo de nuestras especulaciones éticas, políticas y epistémicas; incluso, nuestro lenguaje, la semántica, la manera de entendernos los unos a los otros, pareciera, según algunos⁸, depender de un significado que es una función de nuestras representaciones mentales. Pero, ¿cómo poner en duda la existencia de un mundo interno? Ud. y yo no negamos nuestro mundo interno, nuestro lenguaje e intenciones; nuestras pasiones y creatividad; nuestros sueños, temores, deseos y sentimientos; nuestra razón y memoria. ¿No se estará planteando mal el problema sobre lo qué es la mente? No; desde este punto de vista, ni siquiera se plantea el problema. No se pone en duda su existencia y tal asunción ontológica pretende ser el basamento de la objetividad de nuestras preocupaciones por los otros, de nuestra racionalidad e incluso de nuestra forma de explicar el mundo externo. El entendimiento de nuestro mundo interno es un gran problema; mayor, en grado superlativo, es el problema que apunta a la comprensión del mundo interno de los otros como parte de nuestro mundo externo. Putnam, con la negación de su tesis funcionalista como una explicación satisfactoria sobre los fenómenos mentales, nos dio una pista sobre cómo abordar los diversos problemas de la mente; dijo, como apuntamos anteriormente, que él no creía que los estados

8 Vid. e.g. Grice, P.: *Studies in the Way of Words*, Cambridge, MA: Harvard University Press, 1989.

mentales pudieran reducirse a simples estados físico-químicos o a alguna metáfora computacional que nos acercara a su entendimiento. Sostengo, como se verá de inmediato, que las objeciones de Putnam son valederas, solo si se sostiene la tajante división empirista entre mundo interno y mundo externo. Argumento que si se cuestiona la tal división, la tesis funcionalista se erige, nuevamente, como una estupenda opción que nos permite una aproximación al funcionamiento de nuestros estados mentales.

La división exhaustiva entre el mundo interno y el mundo externo, el llamado por Davidson *tercer dogma*⁹, pareciera volverse difusa cuando se consideran los fenómenos de la mente. Nuestro *Conocimiento del mundo exterior*, como diría Russell, incluye el estudio de los acontecimientos mentales, porque, en principio, todo evento mental es el resultado de alguna actividad en nuestro cerebro. Podemos modificar nuestros estados mentales mediante una alteración en nuestro cerebro; ésta puede ser debida a alguna percepción o a la supresión o consumo de alguna sustancia. Esto apoya la premisa según la cual nuestros estados mentales son el resultado de un acontecimiento físico. Pero, ¿se sigue que podemos explicar nuestros estados mentales como cualquier otro fenómeno físico del llamado *mundo externo*?; ¿las teorías físicas que estudian los fenómenos externos pueden, entonces, indagar, sin más, en nuestra mente? Si es así, podemos sostener, sin duda alguna, la división entre aquello que es interno y aquello otro que es externo; el mundo interno sería epifenoménico al ser el resultado de algún tipo de interacción debida al mundo externo y nada más. Desde el punto de vista del empirismo, nuestros estados mentales, incluyendo nuestra comprensión del mundo (output) son el resultado del modo en que nuestro cerebro procesa los datos sensoriales (inputs). De no ser así, los fenómenos mentales se ubican en una zona difusa, ya que parte de su naturaleza sería físico-química y parte irreductiblemente interna, es decir, serían fenómenos ontológicamente duales, cuyas entradas no son necesariamente externas en el sentido empirista, ya que no podrían ser tratadas como otros fenómenos físicos, lo cual aniquila la exhaustividad de la división entre nuestros esquemas conceptuales y el contenido empírico. Davidson cree en la última alternativa:

...cuando otros piensan de manera distinta a la nuestra, no hay principio general, ni apelación a la evidencia, que nos obligue a decidir que las diferencias están en nuestras creencias en lugar de en nuestros conceptos. (...)

9 Davidson, D. "The Very Idea of a Conceptual Scheme", *Proceedings and Addresses of the American Philosophical Association*, Vol. 47, (1973 -1974), pp. 5-20, p. 11.

Sería un error (...) anunciar la admirada noticia de que toda la humanidad— todos los hablantes de una lengua, al menos— comparte un esquema y una ontología común. Es así porque si no podemos decir inteligiblemente que los esquemas son diferentes, tampoco podemos decir inteligiblemente que son uno.

Al renunciar a la dependencia del concepto de una realidad sin interpretar, algo fuera de todo esquema y toda ciencia, no renunciamos a la noción de la verdad objetiva— más bien al contrario. Dado el dogma de un dualismo de esquema y realidad, obtenemos la relatividad conceptual, y la verdad relativa a un esquema. Sin el dogma esta clase de relatividad se va por la borda. Por supuesto, la verdad de las oraciones sigue siendo relativa al lenguaje, pero esto es todo lo objetivo que se puede ser. Al renunciar al dualismo de esquema y mundo no renunciamos al mundo, sino que reestablecemos el contacto inmediato con los objetos familiares cuyas diabluras hacen que nuestras oraciones y opiniones sean verdaderas o falsas¹⁰.

La premisa adicional de Davidson, a la negación del carácter dual de la ontología propio del empirismo, es la simple y sorprendente tesis según la cual *las razones de una acción son sus causas*¹¹. Si tengo sed y tomo agua, las razones por las que tomo agua, a saber, que tengo sed, es la causa de mi acción, en este caso, tomar agua. Si deseo jugar, las razones por las que quiero jugar son las causas de que juegue, si es que juego. Si hago un regalo, las razones por las que hago el regalo, o deseo, o tengo la intención de hacer el regalo, y la satisfacción que me produciría hacer el regalo, es lo que causa mi acción. Esta tesis, de grandiosa simplicidad, es asombrosa. Es sorprendente porque evidencia una relación causal entre razón y acción distinta a la estipulada por las teorías físicas entre los objetos y los eventos. De ser así, como Putnam lo afirmó, los predicados mentales no son reducibles a predicados físicos. Los objetos son partes integrales de los eventos; no hay evento que no involucre un objeto, pero los objetos en sí mismos no son causa de los eventos; los eventos causan otros eventos; ciertas acciones, que también son eventos, no son causadas por otros eventos, con localizaciones espacio-temporales específicas, sino por razones, razones siempre internas; he ahí la gran diferencia. Hay un solo mundo, un monismo ontológico anómalo donde las generalizaciones psicológicas no son de la misma naturaleza que las leyes físicas porque, si bien todo acontecimiento mental es un acontecimiento físico, no es posible reducir lo mental exclusivamente al mundo de las teorías físicas.

10 Davidson, D. "The Very Idea of a..." Op. cit., p. 19-20.

11 Davidson, D. *Thinking causes in mental causation*. Nueva York: Oxford University Press, 1993.

Las causas de una acción, como en los estados funcionales de Putnam, pueden ser diversas y no tienen que ser, necesariamente las mismas. Las razones de las acciones en los estados funcionales son sus causas y pueden ser estados diversos. El funcionalismo pierde toda relevancia para la filosofía de la mente, si se sigue sosteniendo dogmáticamente la dicotomía entre mundo externo y mundo interno. El funcionalismo reivindica su neutralidad ontológica y su carácter fundamental como sustento de las ciencias cognitivas, si abandona el dogma empirista entre mundo externo y mundo interno, al menos para la consideración de la verdad de nuestras oraciones. La función se erige, así, en el más puro sentido de la filosofía de Frege, como la unidad básica de análisis con respecto a los problemas de la filosofía de la mente. Las máquinas podrán, desde el punto de vista funcionalista, ser buenos modelos de la actividad mental, pero nunca podrán emular a la mente, porque sus fenómenos pueden ser reducidos completamente a fenómenos físicos, mientras que el monismo anómalo niega tal posibilidad para nuestras modalidades mentales. Pareciera, bajo esta argumentación, que el funcionalismo computacional permitiría, de alguna forma, por demás plausible, modelar parte de lo que es nuestra mente, a saber, aquella parte que se puede reducir a funciones computables, mientras que la modulación siempre quedará incompleta. Modelos que siempre serán parciales porque, como mostró Gödel en 1933, la verdad trasciende a la demostración en el sentido de que hay más verdades que enunciados que podemos demostrar. Se imponen los modelos y, con ellos, funciones que involucran fascinantes interpretaciones. Nuestros estados internos no carecen de interpretación, de hecho, cuando logramos demostrar un teorema, en un lenguaje no-interpretado, la emoción y excitación que produce tal demostración y las motivaciones que nos permitieron desarrollar tal o cual estrategia demostrativa, jamás carecen de intencionalidad y significado, esto es, de contenido. No se desecha el funcionalismo, no obstante somos conscientes de sus limitaciones. De tal forma podremos, como de hecho lo hacemos, simular ciertos procesos cognitivos simples que tienen que ver con funciones computables que, por la hipótesis de Church, pueden ser traducidos en una máquina de Turing. La situación es similar a la de los modelos atómicos: modelamos parte de la realidad y, paulatinamente vamos afinando el modelo, pero siempre con la única seguridad de que nuestros modelos son imperfectos y de carácter progresivo. No es extraño que podamos modular aquella parte que nuestra mente creó referida a los algoritmos; las funciones son una de las útiles ficciones

que creó la mente. Creo que mi extrapolación, sustentada en la tesis de Davidson y gracias al propio Putnam, apoya mi conclusión, a saber, el funcionalismo modela aquella parte de la mente que se restringe a nuestro uso de las funciones recursivas.