Colección Teorema
Serie mayor

Susan Haack

1978 *Filosofía de las lógicas*

SEGUNDA EDICION

CÁTEDRA
TEOREMA

Traducción: Amador Antón, con la colaboración de Teresa Orduña

[redacted]

Cubierta: Diego Lara

Ediciones Cátedra, S. A., 1991
Telémaco, 43. 28027 Madrid
Depósito legal: M. 4.095/1991
ISBN: 84-476-0319-6
Printed in Spain
Impreso en Fernández Ciudad, S. L.
Catalina Suárez, 19. 28007 Madrid

Índice

Agradecimientos

Este libro se basa, en su mayor parte, en las lecciones de filosofía de la lógica explicadas en la Universidad de Warwick desde 1971.

Quiero expresar mi agradecimiento a todos los colegas y amigos con quienes he discutido las cuestiones planteadas aquí; especialmente a Nuel Belnap, Robin Haack, Peter Hemsworth, Paul Gochet, Dorothy Grover, Graham Priest y Timothy Smiley por sus detallados comentarios sobre el material de trabajo. Estoy en deuda también con mis alumnos, de quienes he aprendido mucho, y con Jeremy Mynott por sus consejos y apoyo editorial.

Prefacio

El siglo que ha seguido a la publicación de *Begriffsschrift* de Frege ha sido testimonio de un tremendo crecimiento en el desarrollo y estudio de los sistemas lógicos. La variedad de este crecimiento es tan impresionante como su propia escala. Se pueden distinguir cuatro áreas principales de desarrollo; dos en los estudios formales y dos en los filosóficos: (i) el desarrollo del aparato lógico estándar, iniciado con la presentación de la sintaxis del cálculo de enunciados y de predicados de Frege y de Russell y Whitehead, enriquecido más tarde con una semántica por la obra de, por ejemplo, Post, Wittgenstein, Löwenheim y Henkin, y estudiado metalógicamente en la obra de Church y Gödel; (ii) el desarrollo de los cálculos no estándar tales como las lógicas modales iniciadas por C. I. Lewis, las lógicas plurivalentes iniciadas por Lukasiewicz y Post, las lógicas intuicionistas iniciadas por Brouwer. Junto a estas áreas tenemos (iii) el estudio filosófico de la aplicación de estos sistemas al argumento informal, el estudio de la interpretación de las conectivas de oraciones y los cuantificadores, el de conceptos tales como verdad y verdad lógica; y (iv) el estudio de los fines y capacidades de la formalización por aquellos que, como Carnap y Quine, son optimistas sobre la significación filosófica de los lenguajes formales, por los que, como F. C. Schiller y Strawson, son escépticos respecto de las pretensiones de relevancia filosófica de la lógica simbólica, y por los que, como Dewey, propugnan una concepción de la lógica más psicológica que la reinante.

Veo cierta significación filosófica en el hecho de que estos desarrollos hayan tenido lugar en paralelo más que en serie; pues es saludable recordar que las lógicas "no estándar" se han desarrollado al lado de los sistemas estándar, y también que siempre ha habido críticos no sólo de los sistemas formales específicos, sino de las aspiraciones de la formalización misma.

Los desarrollos de las cuatro áreas que he distinguido no fueron, por supuesto, independientes unos de otros; y yo veo también significación filosófica en su interrelación. Por ejemplo, aunque algunas de las ideas clave tanto de la lógica modal como de la plurivalente ya fueron anticipadas por MacColl en 1880, su desarrollo formal sistemático se produjo, respectivamente, en 1918 tras la formalización canónica de los cálculos no modales de *Principia Mathematica*, y en 1920 tras la provisión de la semántica de tablas de verdad para la lógica bivalente. Sin embargo, la motivación para el desarrollo de los cálculos no estándar provino no sólo del interés matemático por las posibilidades de las extensiones y modificaciones de la lógica clásica, sino también de la crítica filosófica; en el caso de las lógicas modales, del establecimiento del condicional material para representar la implicación y, en el caso de las lógicas plurivalentes, de la suposición de que toda proposición es verdadera o falsa. Y un desarrollo de la lógica no estándar promovió otro: las dudas sobre el éxito de las lógicas modales al formalizar la idea intuitiva de entrañamiento condujo al desarrollo de las lógicas de la relevancia, mientras que el interés matemático de los sistemas modales fomentó el desarrollo, por analogía, de las lógicas epistémica, deóntica y temporal; o, de nuevo, la reflexión sobre la motivación filosófica de las lógicas plurivalentes llevó a la idea de supervaloraciones. Las innovaciones formales, a su vez, han dado una nueva dimensión a las cuestiones filosóficas originariamente generadas por los cálculos estándar: así, por ejemplo, se plantearon de forma nueva y aguda temas sobre la interpretación de los cuantificadores y su relación con los términos singulares cuando se puso en duda la inteligibilidad de la lógica modal de predicados; o, como las antiguas preocupaciones sobre si la lógica se ocupa de oraciones, enunciados o proposiciones que resultaron estar implicadas en el desafío que los sistemas plurivalentes lanzaron a la bivalencia. A veces, nuevos sistemas formales incluso han llegado a desafiar, explícita o implícitamente y más o menos radicalmente, los supuestos aceptados respecto a los fines y aspiraciones de las lógicas formales: la lógica de la relevancia, por ejemplo, cuestiona no solamente la adecuación de los condicionales material y estricto, sino también la concepción clásica de validez; el carácter distintivo de la lógica intuicionista procede en parte de un desafío a la presunción 'logicista' de la prioridad de la lógica sobre la matemática; y la lógica vaga rompe con el principio tradicional de que la formalización debería corregir o evitar la vaguedad, pero no comprometerse con ella. Y, como el último ejemplo recuerda, los nuevos desarrollos formales han aspirado a veces a superar lo que tanto defensores como críticos de la lógica formal habían considerado que eran sus limitaciones inherentes —tales como su supuesta incapacidad, acentuada por Schiller y Strawson, para ocuparse de las características pragmáticas que afectan a la aceptabilidad del razonamiento informal, quizás superada,

al menos en parte, por las pragmáticas formales iniciadas por Montague.

En este libro, mi preocupación se dirige más hacia la filosofía de la lógica que a su historia. Pero mi estrategia ha sido ideada con la mira puesta en la historia de la interacción de los temas formales y filosóficos que acabo de esbozar. Comienzo con la consideración de algunos problemas planteados por el aparato lógico estándar —la interpretación de las conectivas de oraciones, letras de oraciones, cuantificadores, variables, constantes individuales, los conceptos de validez, verdad y verdad lógica—; del capítulo 9 en adelante me dedico a considerar el modo en que algunos de estos problemas motivan innovaciones formales, lógicas 'divergentes' y 'extendidas', los modos en que estos nuevos formalismos conducen, a su vez, a una reevaluación de los temas filosóficos; y concluyo, en el último capítulo, con algunas preguntas —y más bien pocas respuestas— sobre el status metafísico y epistemológico de la lógica, las relaciones entre los lenguajes formales y naturales, y la relevancia de la lógica para el razonamiento.

Y dos temas que recurren en el libro reflejan también esta perspectiva histórica. A mi entender, los temas filosóficos vitales de la lógica se centran en la consideración (i) de la pluralidad de los sistemas lógicos y (ii) de los modos en que los cálculos tienen que ver con la valoración del argumento informal. Más específicamente, insistiré en que, en vista de la existencia de lógicas alternativas, la prudencia exige una postura razonablemente radical respecto de la cuestión del status epistemológico de la lógica, y en que la interpretación de los resultados formales es una tarea delicada a la que es altamente deseable prestar una atención juiciosa para los propósitos de la formalización.

He procurado escribir un libro que sea útil como introducción a los problemas filosóficos que plantea la lógica e inteligible para estudiantes con nociones de lógica formal elemental y con algunos conocimientos de los temas filosóficos, aunque no posean ningún conocimiento previo de la filosofía de la lógica. Pero no he ofrecido respuestas simples o ni siquiera preguntas simples; pues los temas interesantes de filosofía de la lógica son complejos y difíciles. He intentado, en cambio, empezar por el principio, explicar tecnicismos e ilustrar problemas muy generales con estudios de casos específicos. A este fin, y para los neófitos en la materia, he suministrado un glosario de términos posiblemente poco familiares usados en el texto y algún consejo de orientación en cuanto a la literatura específica; mientras que para los deseosos de ir más allá, he incluido una generosa (aunque espero que no intimidatoria) bibliografía. La respuesta de mis estudiantes me ha animado a creer que es innecesario, al mismo tiempo que no deseable, simplificar en exceso. He aspirado —aunque me temo que el resultado se quede corto con relación a la aspiración— a escribir un libro que pueda ser de alguna utilidad al estudiante y al mismo tiempo de algún interés para el profesor.

Resulta molesto, a mi entender, el estar poco seguros acerca de si, o de cómo, un autor ha modificado opiniones que antes había mantenido; pero, por otra parte, resulta pesado estar sometidos a frecuentes discusiones sobre los anteriores errores de un autor. A modo de compromiso, por tanto, indico aquí resumidamente dónde y cómo he modificado las ideas que mantuve en *Lógica Divergente*. Primero: he establecido, espero, la distinción entre cuestiones metafísicas y epistemológicas acerca del status de la lógica con mayor claridad; y esto me ha llevado a distinguir más cuidadosamente entre la cuestión del monismo versos pluralismo y la cuestión de la revisabilidad, y a mantener un pluralismo cualificado más bien que el monismo asumido un tanto confusamente en *Lógica Divergente*. Segundo: he llegado a apreciar que las consecuencias para la ontología de la interpretación sustitucional de los cuantificadores son algo menos simples de lo que suponía; y ello me ha llevado a una consideración más sutil, o en cualquier modo más compleja, de los papeles respectivos de los cuantificadores y términos singulares. Me atrevo a decir, sin embargo, que quizás habré olvidado algunos antiguos errores, además de cometer otros nuevos.

Notación y abreviaturas

A, B ...	metavariables, que fluctúan sobre las letras de oraciones
p, q ...	letras de oraciones
$-$	negación ('no es el caso que')
$\vee$	disyunción ('o'); a veces denominada 'vel'
&	conjunción ('y'); 'ampersand'
$\rightarrow$	implicación material ('si')
$\equiv$	equivalencia material ('si y sólo si')
x, y ...	variables individuales
$(\exists)$	cuantificador existencial ('al menos uno')
$(\ )$	cuantificador universal ('para todo')
(ιx) ...	descripción definida ('el x tal que...')
F, G ...	letras predicativas (R, ... para predicados poliádicos)
a, b ...	términos singulares
$=$	identidad
L	necesariamente
M	posiblemente
$\prec$	implicación estricta
$\Rightarrow$	implicación relevante
$\Rrightarrow$	entrañamiento
$\neg$	negación intuicionista
$\{\ \}$	conjunto
$\{x \mid \ldots x\}$	el conjunto de los x que son...
$\langle\ \rangle$	secuencia (par, triplo ... n-tuplo ordenados)
$\in$	pertenencia de conjuntos
$\lvert\ldots\rvert$	el valor de ...
$<$	menor que
$>$	mayor que
$\leqslant$	menor o igual que
$\geqslant$	mayor o igual que
sii	si y sólo si
fbf	fórmula bien formada
P.C.V.	principio de círculo vicioso
$\vdash$	consecuencia sintáctica
$\models$	consecuencia semántica
MPP	*modus ponens* (de A y $A \rightarrow B$ se infiere B)
RAA	*reductio ad absurdum*

1
"Filosofía de las lógicas"

No hay ningún sustituto matemático para la filosofía.

KRIPKE, 1976

1 LÓGICA, FILOSOFÍA DE LA LÓGICA, METALÓGICA

La tarea de la filosofía de la lógica, tal como yo la entiendo, es investigar los problemas filosóficos suscitados por la lógica —lo mismo que la tarea de la filosofía de la ciencia es investigar los problemas filosóficos suscitados por la ciencia— y la de la filosofía de la matemática investigar los problemas filosóficos suscitados por la matemática.

Una de las preocupaciones centrales de la lógica consiste en discriminar los argumentos válidos de los no válidos; y los sistemas lógicos formales, tales como los conocidos cálculos de oraciones y de predicados, han pretendido suministrar cánones precisos, estándares puramente formales, de validez. Así, entre las cuestiones propiamente filosóficas generadas por el quehacer de la lógica están éstas: ¿Qué quiere decir que un argumento es válido?, ¿que un enunciado se sigue de otro?, ¿que un enunciado es lógicamente verdadero? ¿Ha de ser explicada la validez como relativa a algún sistema formal? ¿O hay una idea extrasistemática que los sistemas formales aspiran a representar? ¿Qué tiene que ver el ser válido con ser un argumento bueno? ¿Cómo ayudan los sistemas lógicos formales a valorar los argumentos informales? Así como, por ejemplo, "*y*" es "&" ¿qué se debería pensar que representan "*p*" y "*q*"? ¿Hay una lógica formal correcta?, ¿y qué podría significar "correcto" aquí? ¿Cómo se reconoce un argumento válido o una verdad lógica? ¿Qué sistemas formales se consideran sistemas de lógica, y por qué? Ciertos temas son recurrentes: el asunto del ámbito y los fines de la lógica, las relaciones entre la lógica formal y el argumento informal, y las relaciones entre los diferentes sistemas formales.

La esfera de la filosofía de la lógica está relacionada con la de la metalógica, pero es distinta de ella. La metalógica estudia las propiedades formales de los sistemas lógicos formales; ello incluiría, por ejemplo, pruebas (o refutaciones) de su consistencia, completud o decidibilidad. La filosofía de la lógica trata asimismo de cuestiones sobre los sistemas lógicos formales —pero de cuestiones filosóficas más bien que puramente formales—. Tomemos, como ejemplo, las relaciones entre el cálculo estándar o bivalente de oraciones y el cálculo plurivalente: el filósofo querrá saber en qué sentido, si es que lo hay, las lógicas plurivalentes son alternativas a la lógica bivalente; si está uno obligado a elegir entre el cálculo plurivalente y el bivalente, y, si es así, por qué razones; cuáles serían las consecuencias para el concepto de verdad si se adoptase un sistema plurivalente, etc. Los resultados metalógicos pueden ayudar mucho a dar respuesta a cuestiones de este tipo: por ejemplo, es presumiblemente una condición necesaria, aunque no suficiente, para que una lógica plurivalente sea una alternativa seria, el que sea consistente; y puede ser relevante para las cuestiones de su status relativo el que (la mayoría de) las lógicas plurivalentes están contenidas en la lógica bivalente (es decir, que todos sus teoremas son teoremas de la lógica bivalente, pero no viceversa). Una segunda diferencia es que la filosofía de la lógica no trata exclusivamente de cuestiones de lógica formal; el argumento informal y las relaciones entre el sistema formal y el argumento informal caen también dentro de su esfera. El desarrollo de los sistemas formales, verdaderamente, aumenta en gran manera la profundidad y el rigor de los estudios lógicos; pero el estudio del argumento informal es con frecuencia un preliminar indispensable para tales desarrollos, y el éxito de la sistematización de los argumentos informales una confirmación de su utilidad. Es pertinente que Frege, uno de los pioneros de la lógica formal moderna, se viera obligado a desarrollar su *Begriffsschrift* (1879) porque necesitaba un medio menos ambiguo e incómodo que el lenguaje alemán en el cual presentar propiamente las pruebas aritméticas rigurosas.

La expresión "filosofía de la lógica" debe preferirse, pienso yo, a la de "lógica filosófica", la cual es propensa a transmitir la desafortunada impresión de que hay una forma filosófica peculiar de hacer lógica más bien que problemas peculiarmente filosóficos acerca de la lógica. (Observo que, a diferencia de "lógica filosófica", "ciencia filosófica" y "matemática filosófica" nunca han conseguido uso corriente.) Sin embargo, mis ejemplos han mostrado ya que el interés filosófico va ligado al hecho de que no hay sólo una, sino una pluralidad de lógicas formales; y, por tanto, "filosofía de las lógicas" es, creo yo, mejor aún.

2 El ámbito de la lógica

Entre los problemas de la filosofía de la ciencia están las cuestiones sobre el ámbito de la ciencia: ¿qué dominios del conocer (o "conocimiento") son considerados como ciencia? —por ejemplo, ¿deberán ser consideradas la alquimia, astrología, sociología o psicología como ciencias *bona fide?*—. ¿Y qué razones se deberían dar para incluir o excluir a un dominio dado del campo de la investigación? Similarmente, entre los problemas de la filosofía de la lógica hay cuestiones sobre el ámbito de la lógica y, en consecuencia, sobre el ámbito de la filosofía de la lógica: ¿qué es lógica?, ¿qué sistemas formales son sistemas de la lógica? ¿y qué los hace ser así? [1].

Como tengo que empezar por alguna parte, daré por supuesto la idea intuitiva de lo que debe ser un sistema formal. Pero indicaré qué rango de sistemas formales tengo en la mente cuando hablo de *lógica* formal.

Es relevante distinguir desde el principio entre sistemas formales *interpretados* y *no interpretados*: un sistema formal no interpretado es precisamente una colección de señales, y, por tanto, no puede ser identificado como una lógica formal más bien que como una formalización de una teoría matemática o física. La pretensión de que un sistema formal sea un sistema de lógica depende, pienso, de que posea una interpretación según la cual pueda considerarse que aspira a incorporar cánones del argumento válido: por ejemplo, considero a las lógicas plurivalentes como lógica porque poseen interpretaciones según las cuales sus valores son "valores de verdad", sus variables son oraciones, sus operadores son la negación, la conjunción, etc. (También poseen *otras* interpretaciones —por ejemplo, en términos de circuitos eléctricos; el isomorfismo entre las interpretaciones lógica y eléctrica es relevante para la forma de comportamiento de los computadores—. Véase Rescher, 1969, página 61, para referencias.) Por tanto, al hablar de varios formalismos como lógica estoy haciendo una apelación implícita a sus interpretaciones usuales.

Al decidir qué formalismos se consideran como lógica he adoptado, por ahora, la hospitalaria actitud de conceder el beneficio de toda duda — si bien, más adelante prestaré alguna atención a los argumentos de por qué sistemas que yo he *incluido* deberían ser *excluidos*—. Una razón en favor de esta actitud es que reduce el peligro de rechazar un sistema formal como "no realmente un sistema de lógica", cuando uno debe preguntarse seriamente si es un sistema bueno o útil. Temo, por ejemplo, que Quine (1970, cap. 5),

[1] La significación de cuestiones tales como éstas, espero, se pondrá más de manifiesto en el transcurso del libro. Los lectores que encuentren esta sección difícil es preferible que vuelvan a ella al terminar el libro.

que excluye el cálculo de predicados de segundo orden debido a lo que él considera que es su compromiso con una ontología de objetos —propiedades— abstractos e intensionales, pueda haber sucumbido a este peligro. (Similarmente, yo desconfiaría de las definiciones acerca de lo que hace que una obra sea artística que incitasen a evadir las cuestiones sobre obras *malas* de arte.) De todas formas, como lógica formal incluiré:

lógica "tradicional"	— silogística aristotélica
lógica "clásica"	— cálculo bivalente de oraciones cálculo de predicados [2]
lógicas "extendidas"	— lógicas modales lógicas temporales lógicas deónticas lógicas epistémicas lógicas de la preferencia lógicas imperativas lógicas erotéticas (interrogativas)
lógicas "divergentes"	— lógicas plurivalentes lógicas intuicionistas lógicas cuánticas lógicas libres
lógicas "inductivas"	

La intención es distinguir entre la lógica formal y los sistemas de aritmética, geometría o las axiomatizaciones de biología, física, etc. La demarcación no se basa en ninguna idea verdaderamente profunda sobre "la naturaleza esencial de la lógica" —en realidad, dudo de que haya una tal "naturaleza esencial"—. Pero esto no es completamente arbitrario; se corresponde razonablemente bien, pienso, con lo que los autores de filosofía de la lógica tienen en la mente, por lo general, cuando hablan de "lógica"; y ello tiene, al menos, la siguiente base racional pragmática.

Los sistemas formales conocidos como lógica "clásica" o "estándar" (y enseñados en los cursos de lógica formal elemental) debien ciertamente considerarse como lógica, si hay algo que deba considerarse como lógica. Parece entonces apropiado admitir tam-

[2] De acuerdo con la actitud del "beneficio de la duda", hago esto para incluir la teoría de la identidad (es decir, los axiomas o reglas para "=") y el cálculo de predicados de segundo orden (i.e., la cuantificación que liga "F"..., etc., así como "x"..., etc.) además del cálculo de predicados de primer orden.

bién como lógica a aquellos sistemas que son análogos a éstos. Entre tales sistemas "análogos" incluyo: las extensiones de la lógica clásica, esto es, sistemas que añaden nuevo vocabulario lógico ("necesariamente" y "posiblemente" en las lógicas modales, "solía ser el caso que" y "será el caso que" en las lógicas temporales, "debe" y "puede" en las lógicas deónticas, "sabe" y "cree" en las lógicas epistémicas, "prefiere" en las lógicas de la preferencia) junto con nuevos axiomas o reglas para el nuevo vocabulario, o que aplican operaciones lógicas familiares a nuevos ítem (oraciones imperativas o interrogativas); las divergencias de la lógica clásica, i. e., sistemas con el mismo vocabulario, pero con diferentes (más restringidos usualmente) axiomas o reglas; las lógicas inductivas que pretenden formalizar una noción de soporte análoga a, pero más débil que, la consecuencia lógica. Su semejanza con la lógica clásica —no solamente semejanza formal, sino también semejanza en los propósitos y en la interpretación pretendida— hace que sea natural considerar estos sistemas como lógica. (Alternativamente, yo podía haber comenzado por la lógica tradicional aristotélica, de la cual la lógica clásica moderna es una extensión, y haber procedido desde allí mediante un proceso semejante de analogía.)

Sin embargo, la idea de que un sistema sea suficientemente similar a la lógica clásica es evidentemente bastante vaga; y uno podría preguntarse de modo razonado si el ámbito de la lógica podría ser delimitado de una manera menos pragmática y más precisa.

FORMA ▷ CONTENIDO

Podría pensarse que la idea tradicional de que la lógica se ocupa de la validez de los argumentos en cuanto tales, esto es, independientemente de su contenido —que la lógica es, como claramente afirma Ryle, "neutral respecto al tópico"—, proporciona un principio para, según el cual, delimitar el ámbito de la lógica. En este sentido, aquellos sistemas que son *aplicables al razonamiento independientemente de su contenido* se considerarían como lógica. Es ésta una idea con la que simpatizo; pero dudo de que sea en realidad apreciablemente más precisa que la noción de analogía con la lógica clásica con la que empecé. ¿Qué quiere decir, primero, que un sistema formal es "aplicable" al razonamiento de tal y cual contenido? Posiblemente que se ha pretendido que sus principios sean verdaderos para tal razonamiento. Y, luego, ¿qué se entiende por "independientemente de su contenido"? Podría sugerirse que, mientras los cálculos de oraciones y de predicados son indiferentes al contenido, la aritmética, por ejemplo, no es "neutral respecto al tópico" puesto que versa específicamente sobre números; pero esto da origen a cuestiones difíciles sobre "acerca de" (¿trata el cálculo de predicados de primer orden "acerca de individuos"?). Se sugiere de nuevo que la lógica se aplica al razonamiento independientemente de su contenido porque se ocupa de la *forma* de los argumentos más bien que de su *contenido*. Además la idea, pienso, es útil, aunque todavía

DEONTOLOGÍA: CIENCIA QUE ESTUDIA EL DEBER.

imprecisa. ¿Cómo debe uno distinguir entre la forma de un argumento y su contenido? La lógica temporal se aplica a oraciones temporales, la lógica imperativa a oraciones imperativas, y el tiempo o modo de una oración podía plausiblemente considerarse como asunto de su forma más bien que de su contenido; pero otros casos son más complicados —la idea de forma necesitaría de refinamiento antes de poner en claro, por ejemplo, que una oración sobre creencias es cuestión de la forma y una oración sobre números es cuestión del contenido.

Sin embargo, la vaguedad de la idea de neutralidad respecto al tópico así como la distinción establecida entre forma y contenido no son necesariamente censurables; como dije, dudo que la lógica posea un "carácter esencial" exactamente específico. Cuando manifesté que las lógicas modales, por ejemplo, son lo bastante semejantes a la lógica clásica para incluirlas en el ámbito de la lógica, confiaba implícitamente en la idea de que los adverbios "necesariamente" y "posiblemente" son suficientemente neutrales respecto al tópico para considerarlos como "nuevo vocabulario *lógico*". En consecuencia, la idea de neutralidad respecto al tópico puede ayudarnos desde luego a fortalecer nuestras intuiciones sobre qué sistemas formales son de manera relevante análogos a la lógica clásica. Es también significativo el que señalar dónde hay que trazar la línea divisoria entre la lógica y otros sistemas formales resulte más dudoso y discutible en unos casos que en otros. Por ejemplo: algunas teorías matemáticas, especialmente la teoría de conjuntos, son muy generales en su aplicación y parecen tener grandes afinidades con la lógica; mientras que las lógicas epistémicas y de la preferencia parecen ser más específicas en cuanto al contenido que los formalismos lógicos estándar y no tener una pretensión tan fuerte de ser incluidas. En pocas palabras, tanto más dudoso se encuentra uno acerca de la exclusión de un formalismo "matemático" cuanto más general es su aplicación, y más dudoso acerca de la inclusión de un formalismo "lógico" cuanto menos general es su aplicación; esto sugiere que la neutralidad tópica es vaga *en el sentido correcto*.

Estas ideas resultarán importantes posteriormente. La distinción entre la forma y contenido será objeto de un examen más minucioso cuando, en el próximo capítulo, trate la tesis de que la validez de un argumento depende de su forma; y la idea de que la lógica es neutral respecto al tópico adquirirá relevancia cuando, en el cap. 12, aborde la cuestión del monismo versus pluralismo en lógica, i. e., si hay, por así decir, una lógica correcta, o si podrían asignarse diferentes lógicas para cada una de las diferentes áreas del discurso.

Se sugiere a veces un criterio metalógico puramente formal para demarcar la lógica de otros sistemas formales. Kneale, por ejemplo, pide que solamente se permitan los sistemas *completos* dentro del ámbito de la lógica. El resultado de adoptar dicho criterio sería

restringir mi hospitalaria lista; el cálculo de predicados, dado que no es completo en el sentido usual, sería excluido según este criterio. Esta propuesta tiene la ventaja de la precisión; sin embargo, uno tiene derecho a preguntar qué fundamento racional tendría esto —¿por qué habría de ser la completud el criterio de que un sistema sea una lógica? Kneale (1956, págs. 258-9) argumenta de esta manera: el hecho de que una teoría sea incompleta muestra que sus conceptos básicos no pueden ser completamente formalizados, y esto, en virtud del carácter esencialmente formal de la lógica, justifica el excluir tales teorías del ámbito de la lógica. Así, de forma interesante, Kneale está proponiendo la completud como prueba de que un sistema es "puramente formal"; él conecta la idea precisa de completud con la noción más vaga de neutralidad tópica. Pero me temo que el argumento de Kneale pueda estar dependiendo de una equivocación demasiado "formal": el sentido en el que la incompletud de la teoría de conjuntos muestra que su concepto básico, la pertenencia, no es puramente "formal", es simplemente que ese concepto no puede ser completamente caracterizado por medio de un conjunto de axiomas y reglas, las cuales generan todas las verdades que lo implican esencialmente; no son obvias las razones que llevan a pensar que un concepto tal no es "formal" en el sentido de que pertenece al contenido más bien que a la forma de los argumentos.

Presiento que las perspectivas de un criterio formal bien motivado no son muy prometedoras (pero cfr. pág. 39n.). Un ejemplo apoya esta impresión: si se hace hincapié en el papel de la lógica como guía del razonamiento, como medio de valoración de los argumentos informales, se podría ver algún motivo al exigir que los sistemas lógicos sean decidibles, que hay un procedimiento mecánico para decidir si una fórmula es o no teorema. Pero esto restringiría el ámbito de la lógica muy severamente por cierto, pues aunque el cálculo de oraciones es decidible, el cálculo de predicados no lo es.

Es notable que prácticamente toda "lógica" no estándar ha sido, en algún momento, objeto de críticas sobre la base de que no es realmente lógica en modo alguno; lo cual hace sospechar que un punto de vista restrictivo acerca del ámbito de la lógica puede ocultar un conservadurismo que, si se propugnase abiertamente, sería puesto en tela de juicio.

No obstante, puede resultar instructivo considerar algunos argumentos aducidos para excluir sistemas que yo, en consonancia con la actitud del "beneficio de la duda", he incluido. Dummett ha insistido (1973, págs. 285-8; y cfr. Kneale y Kneale, 1962, página 610) en que las "lógicas" epistémicas no son realmente lógica porque creencia y conocimiento son nociones ineludiblemente vagas. Es verdad que el aumento de la precisión ha sido un elemento importante en la motivación de la formalización de la lógica y, en consecuencia, el lógico, al elegir las constantes, debe evitar normalmente la vaguedad, si bien, es más discutible si la vaguedad prohíbe de modo absoluto el empleo lógico de un concepto. Por supuesto, el tratamiento que el lógico hace de "no", "y", "o" o "si" involucra ya una considerable ordenación de la negación informal, de la conjunción informal, etc. (cfr. cap. 3, § 2); la cuestión no es, pienso, simplemente si "sabe" y "cree" son vagos, sino si su vaguedad es ineliminable, esto es, si se oponen *necesariamente* a la reglamentación. Y se debe admitir que las lógicas epistémicas que se encuentran en la literatura (cfr. Hintikka, 1962) son un tanto desilusionantes y por una razón sobre la que Dummett llamó la atención: que se tiende a descubrir un axioma con efecto de que si s cree que p, y q se sigue de p, entonces s cree que q. En otras palabras, el concepto vago corriente de creencia se reemplaza por un suplente lógico, quizás llamado "creencia racional", que permite la construcción de un sistema formalmente interesante, pero que limita bastante severamente su relevancia a argumentos informales sobre la creencia.

Además, otros, Leśniewski por ejemplo, han sugerido que los sistemas plurivalentes no deberían en realidad considerarse como lógica (véase Rescher, 1969, pág. 25). Es cierto que algunos sistemas plurivalentes fueron ideados y estudiados lejos del interés puramente formal o con fines tecnológicos de computadores; pero también es cierto e importante que pioneros tales como Lukasiewicz y Bochvar indicaron bastante claramente que ellos presentaban sistemas lógicos como alternativos al aparato clásico. Con todo, la afirmación de que un sistema formal es lógica depende, como manifesté, de que posea un cierto tipo de interpretación; y una razón que se podría aducir para excluir los sistemas plurivalentes es que éstos exigen un cambio demasiado radical en la teoría de la verdad, o tal vez de los portadores de verdad, para ser suficientemente análogos a la lógica clásica bivalente. La importancia que uno debe dar a este tipo de argumento depende obviamente de lo radical que crea que es el efecto de la plurivalencia sobre el concepto de verdad (cfr. Haack, 1974, cap. 3, para una discusión interesante).

Concedí a los sistemas epistémicos y plurivalentes el beneficio de la duda acerca de su status como lógica. En cada caso, sin embargo, las dudas que surgen se basan en consideraciones cuya relevancia acato: en el caso de las lógicas epistémicas, de la dificultad de eliminar la vaguedad de los nuevos operadores; en el caso de las lógicas plurivalentes, de la dificultad en proporcionar una interpretación apropiada de los nuevos valores. La importancia de estas consideraciones radica en que cuestionan la fuerza de la analogía de las "lógicas" epistémicas o plurivalentes con la lógica clásica respecto a los fines y a la interpretación. Sin embargo, yo prefiero

admitir estos sistemas como lógica al mismo tiempo que, por supuesto, someter a un examen riguroso sus credenciales de alternativas a la lógica clásica. Esta tolerancia ayudará a contrarrestar cualquier conservadurismo inherente al procedimiento de demarcación de la lógica mediante la analogía con los sistemas clásicos.

Se podría preguntar uno con todo derecho: ¿qué importancia tiene exactamente el delimitar el ámbito de la lógica? Algunas veces se ha considerado este asunto como crucial para una tesis filosófica; el caso del *logicismo* nos proporciona un ejemplo interesante.

El logicismo es la tesis (sugerida por Leibniz y desarrollada en detalle por Frege) de que la aritmética es reducible a la lógica: esto es, que los enunciados aritméticos pueden expresarse en términos puramente lógicos y, por tanto, los teoremas aritméticos pueden derivarse de los axiomas puramente lógicos[3]. Dado que cierto conjunto de fórmulas puede, en el sentido explicado, ser reducido a un cierto otro conjunto, el que esto cuente como "reducción de aritmética a lógica" dependerá de si se ha permitido que el primer conjunto represente adecuadamente la aritmética y de si el segundo conjunto es propiamente descrito como "puramente lógico". En el caso del logicismo, cabría plantear dudas en ambos sentidos. Pudiera alegarse que el teorema de incompletud de Gödel muestra que no es posible derivar todas las verdades de la aritmética a partir de conjunto alguno de axiomas, y que, por tanto, *a fortiori*, no es posible derivarlas de conjunto alguno de axiomas lógicos. O, lo que vendría más a propósito con nuestra presente consideración, pudiera alegarse que los axiomas a los que redujo Frege los postulados de Peano para la aritmética no son "puramente lógicos", sino matemáticos, puesto que incluyen principios de teoría de conjuntos. Quine, por ejemplo, alega (1970, págs. 64 y ss.) que la teoría de conjuntos no debiera ser contada como parte de la lógica. Pero sus razones son menos que conclusivas: advierte que hay teorías de conjuntos alternativas, pero también hay lógicas alternativas (cfr., más abajo, caps. 9-12), y pone objeciones a los graves compromisos ontológicos que comporta la teoría de conjuntos, pero el criterio de compromiso ontológico que él emplea está abierto a discusión (véase cap. 4, § 2).

He aquí, pues, un caso en el que el destino de una teoría filosófica parece depender de la demarcación de la lógica. ¿Y no es más

[3] Frege inventó el primer sistema lógico formal completamente desarrollado, como él esperaba, para establecer la verdad del logicismo derivando realmente los postulados de Peano para la aritmética a partir de sus axiomas lógicos. Desarrolló el aparato lógico en 1879, suministró las definiciones lógicas apropiadas de los términos aritméticos en 1884 y las derivaciones en 1893 y 1903; véase Carnap, 1931, para una introducción sencilla a la filosofía logicista de la matemática.

bien descorazonador pensar que la verdad del logicismo dependiera de una cuestión tan pragmática como he considerado que es el ámbito de la lógica? Pienso que no, después que uno profundiza un poco y se pregunta por qué se juzgaría importante que la aritmética sea en realidad puramente lógica. El tema de verdad importante queda, o así me lo parece a mí, oscurecido por plantear la cuestión como si el ámbito de la lógica fuese el punto clave. ¿Por qué pensó Frege que era importante mostrar que la aritmética es reducible a la lógica? La motivación del logicismo fue, al menos en parte, epistemológica; los principios de la lógica, pensaba Frege, son autoevidentes, de modo que si se puede mostrar que las leyes de la aritmética son derivables a partir de ellos, se muestra de ese modo que son epistemológicamente seguras —adquieren la inocencia por asociación, por decirlo así—. Pero resultó que la lógica de Frege (o "la lógica") era inconsistente —la paradoja de Russell (cfr. cap. 8) es derivable en ella. La respuesta de Frege al descubrimiento de la inconsistencia fue que él nunca había pensado realmente que el axioma relevante fuese tan autoevidente como los demás —comentario éste que bien puede originar un saludable escepticismo ante el concepto de autoevidencia. Si bien, la relevancia de esta historia para nuestros actuales intereses es la siguiente: que, como la base de Frege —lógica o no— no posee la importancia epistemológica que él pensaba, el aspecto epistemológico de su programa está perdido sin hacer caso de la decisión sobre la demarcación de la lógica.

Una cosa, al menos, quedará ya totalmente clara: el que si un sistema se considera o no como lógica es una cuestión que involucra en sí misma problemas filosóficos verdaderamente profundos y difíciles. Lo mejor es que la omnipresencia de los problemas filosóficos en la lógica sea evidente desde el principio. Pues el mismo rigor, que es la virtud principal de la lógica formal, tiende a darle a la lógica un aire de autoridad como si estuviese por encima de la reflexión filosófica. Y es esa también la razón por la que acentúo la pluralidad de sistemas lógicos; ya que al tener que decidir entre alternativos se ve uno obligado con frecuencia a reconocer previas concepciones metafísicas o epistemológicas que, de otro modo, habrían permanecido implícitas.

2
Validez

1 Valoración de argumentos

Los argumentos se valoran de muchas maneras; por ejemplo, se considera que unos son más persuasivos o convincentes que otros, que algunos son más interesantes o provechosos que otros, etc. Los tipos de valoración que cabe hacer se pueden clasificar, en líneas generales, de esta manera:

(i) lógica: ¿hay una conexión del tipo adecuado entre las premisas y la conclusión?
(ii) material: ¿son verdaderas las premisas y la conclusión?
(iii) retórica: ¿es el argumento persuasivo, atractivo e interesante para la audiencia?

Sólo he dado una indicación de lo más imprecisa sobre los tipos de pregunta característica de cada dimensión de valoración, pero una tosca indicación podría ser adecuada para los propósitos del momento. La categoría aparte que se ha dado a las consideraciones retóricas no intenta sugerir que la validez de un argumento, o la verdad de sus premisas, es completamente irrelevante respecto a su persuasión; se intenta, más bien, tener en cuenta el hecho de que, aunque si los hombres fueran completamente racionales serían persuadidos sólo por argumentos válidos con premisas verdaderas, de hecho, con bastante frecuencia *son* persuadidos por argumentos no válidos o argumentos con premisas falsas y *no* son persuadidos por argumentos correctos (cfr. pág. 34, abajo) (para discusión de tales fallos de racionalidad y consejos para evitarlos, véase, por ejemplo, Thouless, 1930; Stebbing, 1939; Flew, 1975, y Geach, 1976).

En lo que sigue trataré casi exclusivamente de la primera dimensión de valoración, la lógica. En esta dimensión, a su vez, necesito distinguir diferentes estándars de valoración que se pueden emplear: un argumento se puede considerar que es *deductivamente válido*, o *deductivamente inválido* pero *inductivamente fuerte*, o ninguna de las dos cosas. Los niveles deductivos, como esto indica, y como veremos con más detalle después, son más rigurosos que los inductivos —la conexión entre premisas y conclusión tiene que ser, por decirlo así, más estrecha para la validez deductiva que para la fuerza inductiva[1].

A veces se sugiere (por ejemplo, Barker, 1965; Salmon, 1967) que hay *dos tipos de argumentos;* por una parte, los argumentos deductivos, y los argumentos inductivos, por otra. Esta "distinción", al menos tal como se explica normalmente, sólo confunde materias. Se dice que los "argumentos deductivos" son "explicativos" o "no-ampliadores", es decir, que "no contienen nada en la conclusión que no estuviera ya contenido en las premisas". Si esto se pretende que sea, tal como lo parece, una explicación de aquello por lo que un argumento es *deductivamente válido*, es susceptible de resultar falso, si "no contienen nada en la conclusión que no estuviera ya contenido en las premisas" se toma literalmente (pues mientras "A y B, por tanto, A" satisface esta condición, "A, por tanto, $A \vee B$", que también es deductivamente válido, no lo hace), o trivial, si "no contienen nada en la conclusión que no estuviera ya contenido en las premisas" se toma metafóricamente (¿para qué sirve la prueba de que "$A \vee B$" está *implícitamente* "contenido en" "A", si no para demostrar que "$A \vee B$" se sigue deductivamente de "A"?). Los "argumentos inductivos", por contraste, se dice que son "ampliadores" o "no-explicativos"; esto quiere decir que "sus conclusiones van más allá de lo que está contenido en sus premisas". Esto pone las cosas peor, porque no se puede tomar, simétricamente con la explicación de "argumento deductivo", como una explicación de para qué sirve que un argumento sea inductivamente fuerte. Pues todo lo que se dice sobre los argumentos inductivos es que no son deductivamente válidos; pero no todos los argumentos deductivamente inválidos son inductivamente fuertes.

Así pues, yo prefiero (con Skyrms, por ejemplo, 1966, cap. 1) presentar el asunto de esta manera: no es que haya dos tipos de argumento, sino que los argumentos pueden ser valorados lógicamente según estándars diferentes, deductivo o inductivo; pueden ser deductivamente válidos, inductivamente fuertes o ninguna de las dos cosas. Y esto aclara lo que proponen las siguientes preguntas: ¿Qué es un argumento? ¿Qué condiciones debe reunir un argumento para considerarlo como deductivamente válido o inductivamente fuerte?

¿Qué es un argumento? Bien, se reconoce que algunos fragmentos del discurso tratan de sustentar una conclusión mediante premisas, razonando la conclusión a partir de las premisas; en el discurso informal de los lenguajes naturales esta intención puede señalarse indicando el paso de un enunciado a otro por medio de locuciones como "así pues", "por tanto", "se sigue que", "porque", etc.; en lógica formal mediante la presentación de una serie de fórmulas con una indicación en cada línea en la que se afirma que se sigue por tales y tales reglas de inferencia de tales y tales línea o líneas anteriores. Lo que se juzga, si es válido o inválido, puede pensarse simplemente como un fragmento de discurso: si se está considerando un argumento formal, una secuencia de fbf de un lenguaje formal, o, si se está considerando un argumento informal, una secuencia de oraciones (o quizás enunciados o proposiciones; cfr. cap. 6) del lenguaje natural. (Del mismo modo, algunas de las cosas que la gente dice, se quieren presentar aseverativamente —el hablante tiene la intención de afirmar su verdad— y otras no; pero esto es *lo que se dice* que es verdadero o falso.)

[1] Algunos autores, Peirce notablemente y, más recientemente, Hanson, piensan que hay otros estandars lógicos, estándars "abductivos" también. Cfr. Haack, 1977b, para una discusión relevante.

2 VALIDEZ DEDUCTIVA: CON BREVES COMENTARIOS SOBRE LA FUERZA INDUCTIVA

Validez en un sistema

En un sistema lógico formal, la validez puede definirse de dos formas, sintáctica y semánticamente, es decir, en términos de los axiomas o reglas del sistema, y en términos de su interpretación. Representaré un argumento formal como una secuencia de fórmulas bien formadas (es decir, oraciones gramaticales de un lenguaje formal; en adelante, "fbfs"). $A_1, \ldots, A_{n-1}, A_n$ ($n \geqslant 1$), en la cual $A_1, \ldots, A_{n-1}$ son premisas y A_n la conclusión. *La validez sintáctica* puede explicarse entonces según las siguientes líneas:

> $A_1, \ldots, A_{n-1}, A_n$ es válido-en-L en el caso de que A_n sea derivable de $A_1, \ldots, A_{n-1}$, y los axiomas de L, si los hay, mediante las reglas de inferencia de L.

Esto normalmente se representa por $A_1, \ldots, A_{n-1} \vdash_L A_n$.

La validez semántica puede explicarse según las siguientes líneas:

> $A_1, \ldots, A_{n-1}, A_n$ es válido-en-L en el caso de que A_n sea verdadero para todas las interpretaciones en las que $A_1, \ldots, A_{n-1}$ son verdaderas.

Esto normalmente se representa por $A_1, \ldots, A_{n-1} \vDash_L A_n$.

La "L" en "$\vdash_L$" y "$\vDash_L$" sirve para recordar que estas dos concepciones de validez son *relativas-al-sistema*.

En correspondencia con las ideas sintáctica y semántica de validez de secuencias de fbfs están respectivamente las ideas de teorematicidad y verdad lógica de las fbfs. Se puede observar que admití la posibilidad de argumentos que consten tan sólo de una fbf (a veces se les llama "conclusiones de cero-premisas"). Si las ideas de validez esbozadas hace un momento se aplican a este caso especial, el resultado es:

> A es válido-en-L (es un *teorema* de L) en el caso de que A se siga de los axiomas de L, si los hay, mediante las reglas de inferencia de L ($\vdash_L A$)

y

> A es válido-en-L (es una *verdad lógica* de L) en el caso de que A sea verdadero para todas las interpretaciones de L ($\vDash_L A$).

He representado teorematicidad y verdad lógica como, por así decirlo, casos especiales de la validez sintáctica y semántica, respectivamente. También habría sido posible enfocar la cuestión al revés y explicar la validez como la teorematicidad del condicional correspondiente. El anterior enfoque tiene la ventaja de insistir en la preocupación de la lógica por la conexión entre premisas y conclusión, que es por lo que lo elegí.

¿Cómo se acoplan las ideas sintácticas y semánticas? Bien, naturalmente se aspira a tener un sistema formal en el que aquellas fbfs que son sintácticamente válidas sean semánticamente válidas (los resultados de corrección[2] y completud muestran que teorematicidad y verdad lógica coinciden).

Validez extra-sistemática

Las concepciones de validez, sintáctica y semántica, consideradas hasta aquí, son relativas-al-sistema y aplicables sólo a argumentos formales. ¿Qué sucede, sin embargo, cuando se considera que un argumento informal es válido? Supongo que se afirma que su conclusión *se sigue de sus premisas, que sus premisas no podrían ser verdaderas y su conclusión falsa*. (Si, además de ser válido, un argumento tiene premisas verdaderas —y, por tanto, al ser válido, conclusión verdadera también— se dice que es *correcto*.) Cuando intuitivamente consideramos buenos algunos argumentos informales corrientes, y otros malos, probablemente se está desplegando algo parecido a esta concepción de validez. Por supuesto, considerar un argumento "bueno" es susceptible de incluir *más* que considerarlo válido; pero reconocemos que la validez es una virtud importante, aunque no la única, de un argumento.

[2] Más adelante se definirá un sentido diferente de "correcto" aplicable no a los sistemas lógicos, sino a los argumentos.

Surge la pregunta de si hay también una concepción informal y extrasistemática que se corresponda con las nociones relativas-al-sistema de teorematicidad y verdad lógica. Yo pienso que la hay, aunque sospecho que es algo menos desarrollado y central que la idea extrasistemática de validez (otra razón para tratar la verdad lógica como un caso especial de validez y no a la inversa). La idea extrasistemática de argumento válido como aquel en el cual sus premisas no pueden ser verdaderas y su conclusión falsa, adaptada al caso de un enunciado único (como las definiciones formales se adaptaron al caso de "conclusiones de cero premisas") proporciona la noción de un enunciado que no puede ser falso —en otras palabras, la noción de una verdad necesaria—. Y algo parecido a esta idea se encuentra en efecto en el nivel informal. Por ejemplo, se considera que algunos enunciados son "tautológicos"; esto significa, en sentido no técnico, que esos enunciados son trivialmente verdaderos, que dicen (como sugiere la etimología de "tautológico") la misma cosa dos veces, y en consecuencia no pueden ser falsos. La noción informal de tautología, por supuesto, es más amplia que el uso técnico, el cual incluye sólo verdades lógicas de la lógica *veritativo-funcional*. Y la idea informal de verdad necesaria es también más amplia que la idea formal de verdad lógica (cfr. cap. 10, § 1). No debería producir gran sorpresa el que estas concepciones informales se hayan depurado con el desarrollo y estudio de los sistemas lógicos formales.

Pero ¿qué decir de la conexión entre las concepciones de validez relativas-al-sistema, aplicables a los argumentos formales y la concepción extrasistemática, aplicable a los argumentos informales? Algo parecido a esto: los sistemas lógicos formales pretenden formalizar los argumentos informales, para representarlos en términos precisos, rigurosos y generalizables; y un sistema lógico formal aceptable debería ser tal que, si un argumento informal dado está representado en él por cierto argumento formal, entonces ese argumento formal debería ser válido en el sistema en el caso de que el argumento informal fuera válido en el sentido extrasistemático.

Logica utens y logica docens

De hecho, probablemente ha de haber un proceso de ajuste bastante complejo. Se puede comenzar por desarrollar un sistema formal sobre la base de juicios intuitivos de la validez extrasistemática de los argumentos informales, representando esos argumentos en una notación simbólica, e inventar reglas de inferencia de tal modo que las representaciones formales de los argumentos informales considerados (in) válidos sean (in) válidas en el sistema. Dadas estas reglas, sin embargo, otros argumentos formales resultarán ser válidos en el sistema, quizás argumentos formales que representan argumentos informales considerados inválidos intuitivamente; y entonces se pueden revisar las reglas del sistema, o, en su lugar, especialmente si la regla es adecuadamente sencilla y plausible y la intuición de la invalidez informal no es fuerte, se puede revisar la propia opinión sobre la validez del argumento informal, o si no revisar la propia opinión sobre la conveniencia de representar ese argumento informal de esa forma concreta.

Y una vez que un sistema lógico formal se convierte en bien-establecido, es, por supuesto, probablemente el que guiará cada vez las propias intuiciones sobre la validez o invalidez de los argumentos informales. Siguiendo a Peirce (quien a su vez toma prestada la terminología de los lógicos medievales) se puede llamar al propio juicio irreflexivo sobre la validez de los argumentos informales, la *logica utens*, y considerar a los juicios más rigurosos y precisos desarrollados como sistemas formales mediante la reflexión sobre esos juicios, la *lógica docens*. La imagen es más o menos así (fig. 1).

logica utens
argumentos informales

logica docens
argumentos formales

representación
simbólica
del argumento
informal

validez
extrasistemática

validez
relativa al sistema

Fig. 1

Algunos autores tienen dudas sobre la adecuación de la concepción extrasistemática de validez tal como la he expuesto anteriormente. A lo que se oponen específicamente en la idea de que un argumento es válido si es imposible que sus premisas sean verdaderas y su conclusión falsa, es a la "y". Según esta consideración, si las premisas de un argumento son imposibles, o si su conclusión es necesaria, entonces, ya que *a fortiori* es imposible que sus premisas puedan ser verdaderas y su conclusión falsa, ese argumento es válido; y, por supuesto, esto es así incluso si las premisas son bastante irrelevantes para la conclusion. Los defensores de la "relevancia lógica", por consiguiente, objetan esta concepción de

validez; y debido a esta objeción, instan a la adopción de una lógica formal no-clásica que exige la relevancia de las premisas para la conclusión (véase Anderson y Belnap, 1975, § 22.2.1, y cfr. cap. 10, § 7); así pues, su descontento con la habitual concepción informal de validez está ligado íntimamente a su reto a la lógica clásica. (Convencionalmente, las consideraciones sobre la relevancia son apropiadas para ser relegadas más a la dimensión retórica de la valoración de argumentos que a su dimensión lógica.)

Fuerza inductiva

La fuerza inductiva podría caracterizarse, sintáctica o semánticamente, relativamente a los sistemas formales de la lógica inductiva. Sin embargo, ya que no hay ningún sistema formal de lógica inductiva que posea nada semejante a la especie de atrincheramiento de que disfruta la lógica deductiva clásica, la idea extra-sistemática tiene, en el caso de la fuerza inductiva, un papel particularmente central. La idea es que un argumento es inductivamente fuerte si sus premisas proporcionan un cierto grado de apoyo, aun cuando este no sea concluyente del todo, a su conclusión: es decir, si es *improbable* que sus prémisas sean verdaderas y su conclusión falsa. (Obsérvese que, así establecido, todos los argumentos deductivamente válidos se considerarían como inductivamente fuertes; la validez deductiva será un caso límite de la fuerza inductiva, en que la probabilidad de que las premisas sean verdaderas y la conclusión falsa es cero.)

Hay que señalar, sin embargo, que en su caracterización de la idea extrasistemática de la fuerza inductiva, Skyrms (1966, págs. 9-11) insiste en la formulación: "es improbable, *dado que* las premisas sean verdaderas, que la conclusión sea falsa", porque no quiere admitir que la alta probabilidad de su conclusión o la baja probabilidad de sus premisas sean suficientes, por sí mismas, para la fuerza inductiva de un argumento. Entonces, de modo significativo, esta opinión acerca de la fuerza inductiva es muy semejante a la de relevancia en la concepción de validez deductiva de los lógicos[3].

[3] Es notorio, por supuesto, que hay un problema acerca de la justificación de la inducción. Nada de lo que he dicho muestra que *haya* argumentos que son (deductivamente inválidos pero) inductivamente fuertes. De hecho, creo que la deducción y la inducción son más simétricas de lo que generalmente se supone; cfr. Haack, 1976a.

3 Sistemas lógicos formales: la "L" de "válido-en-L"

Antes distinguí entre concepciones de validez relativa-al-sistema aplicables a los argumentos formales y una concepción extrasistemática aplicable a los argumentos informales. Una exposición adecuada de la primera —de validez-en-L— requerirá evidentemente una explicación de cómo se identifican e individualizan los sistemas formales. El problema puede ilustrarse teniendo en cuenta la lógica de oraciones que se encuentra en los *Principia Mathematica* (Russell y Whitehead, 1910) y en *Beginning Logic* (Lemmon, 1965): si nos interesamos por la diferencia entre lógicas bivalentes y plurivalentes las consideraríamos, naturalmente, como formulaciones alternativas del *mismo* sistema (bivalente), mientras que si estuviéramos interesados en el contraste entre técnicas axiomáticas y de deducción natural (véase, más adelante, pág. 39) podríamos tomarlas como ejemplos de sistemas *diferentes*.

A fin de tener una adecuada terminología neutral, llamaré a la presentación específica de un sistema una "formulación" de un sistema lógico. Ahora bien, las diferencias entre formulaciones son de dos tipos: diferencias en el vocabulario, y diferencias en los axiomas y/o las reglas de inferencia. Esbozaré primero algunas diferencias significativas entre formulaciones, y luego propondré dos consideraciones del "mismo sistema", una más amplia y otra más restringida.

Variantes notacionales

Expresiones tipográficamente diferentes pueden usarse para las mismas operaciones (por ejemplo, para las mismas funciones de verdad). Entre las más comunes de las variantes notacionales corrientes encontramos:

para la negación:	$-p$, $\sim p$, $\bar{p}$, Np
para la disyunción:	$p \vee q$, Apq
para la conjunción:	$p \,\&\, q$, $p \cdot q$, $p \wedge q$, Kpq
para la implicación material:	$p \rightarrow q$, $p \supset q$, Cpq
para la equivalencia material:	$p \equiv q$, $p \leftrightarrow q$, $p \sim q$, Epq
para la cuantificación universal:	(x), $(\forall x)$, Λx, Πx
para la cuantificación existencial:	$(\exists x)$, $(\mathrm{E}x)$, $\mathrm{V}x$, Σx

La última notación, en cada caso, es notación polaca, que tiene la ventaja de no utilizar paréntesis; los operadores preceden a las fórmulas a las que rigen, y el alcance está determinado sin paréntesis.

Constantes primitivas alternativas

Diferentes conjuntos de constantes son equivalentes en poder expresivo; por ejemplo, "&" y "–" o "∨" y "–" para expresar funciones de verdad bivalentes, "$(\exists x)$" y "–" o "(x)" y "–" para la cuantificación existencial y universal. Algunas formulaciones toman por ejemplo "&" y "–" como primitivas y definen "∨" y "→"; otras toman "∨" y "–" como primitivas y definen "&" y "→", y así sucesivamente. Los *Principia Mathematica*, por ejemplo, tienen sólo como primitivas la negación y la disyunción, mientras que *Beginning Logic* tiene la negación, la disyunción, la conjunción y la implicación material.

Formulaciones de deducción axiomática y natural

Un sistema axiomático de lógica (por ejemplo, *Principia Mathematica*) incluye, además de una o más reglas de inferencia, un conjunto privilegiado de fbf, los axiomas, que pueden usarse en cualquier lugar de un argumento, y la verdad de los cuales es incuestionable en el sistema. Los axiomas se incluyen entre los teoremas del sistema, ya que, trivialmente, son derivables de sí mismos. (Un sistema axiomático debe tener al menos una regla de inferencia, ya que no sería posible ninguna derivación o prueba sin tener el medio para pasar de una fbf a otra.)

Una formulación de deducción natural (por ejemplo, *Beginning Logic*), por contraste, descansa sobre reglas de inferencia. (Una regla de supuestos nos permitirá comenzar sin la necesidad de axiomas de los que partir.) Vale la pena observar que las reglas de deducción natural tienen un carácter indirecto, incluso cuasi metalógico; consideremos la regla de eliminación de la disyunción: si se ha derivado C a partir del supuesto A (posiblemente más otros supuestos) y derivado C a partir del supuesto B (posiblemente más otros supuestos), podemos derivar C a partir del supuesto de que $A \vee B$ (más cualquier otro supuesto utilizado para derivar C de A y de B)[4].

A veces axiomas de los que no se sabe si son verdaderos, o incluso que se sabe que son falsos, se adoptan simplemente con el propósito de investigar sus consecuencias. La historia de la geometría presenta un famoso ejemplo. Saccheri tomó como un axio-

[4] Cfr. Blanché, 1962, y Prawitz, 1965, para una discusión detallada de las técnicas axiomáticas y de deducción natural respectivamente. La pionera presentación de la deducción natural por Gentzen, en 1934, incluye un axioma. Gentzen inventó también un cálculo metalógico, el cálculo de secuentes; véase Hacking, 1979, para un interesante intento de demarcación formal del ámbito de la lógica mediante la referencia al cálculo de secuentes.

ma la contradictoria del postulado de las paralelas de Euclides, esperando mostrar que el resultado era un sistema inconsistente, y, por tanto, que el postulado de las paralelas era deducible de los otros axiomas de Euclides. Ya que este postulado es en realidad independiente de los otros, no tuvo éxito en su empeño (cfr. la discusión de las reglas de inferencia de Prior para "tonk", cap. 3, § 2).

Los mismos argumentos válidos y teoremas pueden generarse bien axiomáticamente o bien mediante reglas de deducción natural: por medio de los axiomas de los *Principia Mathematica* o por las reglas de *Beginning Logic*, por ejemplo. Pero, por supuesto, esto no quiere decir que la diferencia entre la deducción natural y las técnicas axiomáticas no sea significativa. Kneale, por ejemplo, argumenta (1956, § 4) que las formulaciones de la deducción natural reflejan mejor la preocupación central de la lógica por la validez de los argumentos. Un desafortunado efecto de la formulación axiomática de *Begriffsschrift* y los *Principia*, sugiere Kneale, fue un traslado de la atención desde la validez de los argumentos a la verdad lógica de las fbfs. Y sugiere Blumberg (1967, pág. 24) que las formulaciones de la deducción natural ponen de relieve la diferencia entre la lógica formal y otras teorías formales, tales como la geometría o la biología, que requieren axiomas especiales relacionados con sus materias específicas, además de una base común de reglas de inferencia. Estoy de acuerdo en subrayar la preocupación de la lógica por los argumentos, y estoy de acuerdo en que con las presentaciones de deducción natural de la lógica formal esta preocupación se pone de relieve. Sin embargo, ya que la validez de los argumentos y la verdad lógica de las fórmulas están íntimamente relacionadas, una formulación axiomática no requiere necesariamente deformar nuestra perspectiva. (Carnap, en 1934, hace notar que se puede pensar en los axiomas como reglas de inferencia bastante peculiares, en el sentido de que se puede inferir una fórmula bien formada dada a partir de algunas premisas o de ninguna.) Y la distinción entre sistemas lógicos y otros sistemas formales no necesita perderse en un planteamiento axiomático de la lógica, ya que queda lugar para una distinción entre axiomas *lógicos* y *propios* (es decir, geométricos, biológicos o cualquier otro). Resulta pertinente que algunos filósofos instrumentalistas de la ciencia hayan instado a considerar las leyes científicas más como reglas que como axiomas.

Axiomas y/o reglas alternativos

Si dos formulaciones tienen variantes notacionales, sus axiomas y/o reglas se diferenciarán al menos tipográficamente; si toman constantes diferentes como primitivas, normalmente, cada uno utilizará sus constantes primitivas en sus axiomas y/o reglas. (A veces, sin embargo, un sistema se formula de un modo tal que las constantes definidas aparecen en los axiomas y/o reglas; en los *Principia* sólo "$-$" y "$\vee$" son primitivas, pero "$\rightarrow$" también aparece en los axiomas.)

Algunas formulaciones utilizan *esquemas de axiomas* en lugar de axiomas y una regla de sustitución. La diferencia radica en poseer, digamos, el axioma:

$$(p \rightarrow q) \rightarrow ((q \rightarrow r) \rightarrow (p \rightarrow r))$$

y la regla de que cualquier instancia de sustitución de un axioma es un teorema, o poseer el esquema:

$$(A \rightarrow B) \rightarrow ((B \rightarrow C) \rightarrow (A \rightarrow C))$$

donde el uso de las "metavariables" "A", "B", "C" indica que independientemente de qué fbf del lenguaje se coloque en el lugar de esas letras la fbf que resulta es un axioma.

Al margen de las divergencias de notación y presentación ya mencionadas, formulaciones diferentes pueden simplemente tener conjuntos diferentes de axiomas/reglas, incluso cuando se tienen en cuenta las diferencias notacionales: sus conjuntos de axiomas-reglas pueden coincidir parcialmente o incluso ser completamente distintos. Como ejemplo compárense los esquemas de axioma de Mendelson y Meredith, ambos con "$-$" y "$\rightarrow$", para el cálculo bivalente de oraciones:

Conjunto de Mendelson:
1. $(A \rightarrow (B \rightarrow A))$
2. $((A \rightarrow (B \rightarrow C)) \rightarrow ((A \rightarrow B) \rightarrow (A \rightarrow C)))$
3. $((-B \rightarrow -A) \rightarrow ((-B \rightarrow A) \rightarrow B))$

Conjunto de Meredith:
1. $((((A \rightarrow B) \rightarrow (-C \rightarrow -D)) \rightarrow C) \rightarrow E) \rightarrow ((E \rightarrow A) \rightarrow (D \rightarrow A))$

(Y véase Prior, 1955, págs. 301 y ss.; Mendelson, 1964, págs. 40-1, para conjuntos alternativos de axiomas.)

El ejemplo dado es para conjuntos alternativos de axiomas del cálculo bivalente de oraciones; las formulaciones alternativas producen los mismos conjuntos de teoremas e inferencias válidas. Otro modo en el cual pueden diferir las formulaciones es en que pueden tener como resultado diferentes teoremas o inferencias válidas; por ejemplo, la lógica intuicionista de oraciones carece de algunos teoremas clásicos, entre ellos la doble negación y el tercio excluso.

En este momento tengo material suficiente para volver a mi problema inicial, tratar las formulaciones alternativas como formulaciones *del mismo sistema*. Sugeriré dos consideraciones del "mismo sistema", una más amplia y otra más restringida, cada una adecuada para ciertos propósitos.

Sentido restringido: L_1 y L_2 son formulaciones alternativas del mismo sistema si tienen *los mismos axiomas y/o reglas de inferencia*; una vez se ha establecido la tolerancia para diferencias de notación (por ejemplo, reemplazando "&" por ".") y constantes primitivas (por ejemplo, reemplazando "$p \& q$" por "$-(-p \vee -q)$").

Sentido amplio: L_1 y L_2 son formulaciones alternativas del mismo sistema si tienen *los mismos teoremas e inferencias válidas* una vez se ha establecido la tolerancia para diferencias de notación y de constantes primitivas.

Un ejemplo: las formulaciones de los *Principia Mathematica* y de *Beginning Logic* son formulaciones de sistemas diferentes en el sentido restringido (uno tiene axiomas, más *modus ponens*, el otro sólo reglas de inferencia), pero del mismo sistema en el sentido amplio (producen los mismos teoremas e inferencias).

Estos dos sentidos del "mismo sistema" ayudarán, espero, a reconciliar algunas intuiciones en conflicto. El más restringido de estos sentidos parece adecuado para utilizarlo en las definiciones de validez-en-L, mientras que el más amplio será más útil, por ejemplo, para contrastar lógicas bivalentes y plurivalentes. Una ventaja del sentido más restringido para la consideración de la validez es que evita un círculo que de otro modo nos amenaza, en el que "teorema" e "inferencia válida" se definen relativamente al sistema, y "sistema" relativamente a los conjuntos de teoremas e inferencias válidas.

El sentido más restringido será también útil para la discusión de formulaciones inconsistentes. Ya que, excepto en algunos sistemas no-convencionales, de una contradicción se sigue cualquier cosa, en virtud del teorema "$A \rightarrow (-A \rightarrow B)$", todos los sistemas inconsistentes se considerarán como el mismo sistema en el sentido más amplio. El sentido más restringido permite respetar la intuición de que algunas aunque no todas las formulaciones inconsistentes son, no obstante, de considerable interés filosófico; un ejemplo sería la de Frege, en la que la paradoja de Russell es un teorema.

4 Validez y forma lógica

No se puede decir si un argumento informal es válido (en el sentido extrasistemático) simplemente investigando los valores de verdad de sus premisas y su conclusión. Si el argumento tiene premisas verdaderas y conclusión falsa, esto nos muestra que es *in*válido;

pero si tiene premisas verdaderas y conclusión verdadera, o falsas premisas y conclusión verdadera, o falsas premisas y conclusión falsa, esto no nos muestra que es válido. Pues es válido sólo si *no puede tener*, no si *no las tiene*, premisas verdaderas y conclusión falsa. Una técnica que con frecuencia se utiliza para mostrar que un argumento es inválido, aunque, casualmente, no tenga premisas verdaderas y conclusión falsa, es encontrar otro argumento *que sea de la misma forma* y que tenga premisas verdaderas y conclusión falsa. Por ejemplo, para mostrar que: "o la prueba de Gödel es inválida, o la aritmética es incompleta, por tanto la aritmética es incompleta", aunque tiene premisa verdadera y conclusión verdadera, es, sin embargo, inválido, se podría señalar que el argumento estructuralmente similar: "O 7 + 5 = 12 o los perros maullan, por tanto los perros maullan" tiene premisa verdadera y conclusión falsa. Esto por supuesto es un método mejor para mostrar la invalidez que la validez; si no se puede encontrar un argumento de la misma forma con premisas verdaderas y conclusión falsa, no es esto una prueba concluyente de que un argumento es válido (cfr. Massey, 1974).

Para mostrar que un argumento es inválido, lo que se busca es un argumento *estructuralmente similar* con premisas verdaderas y conclusión falsa; y esto sugiere que hay algo de verdad en la afirmación de que los argumentos son válidos o inválidos "en virtud de su forma". Y los sistemas lógicos formales están concebidos para representar de una forma generalizada y esquemática la estructura que consideramos es compartida por un grupo de argumentos, y que es la base de su validez o invalidez. Esto es susceptible de sugerir, a su vez, una imagen de los argumentos informales en la que tendrían una estructura única y reconocible, compuesta, por así decirlo, de un esqueleto: las expresiones que constituyen su forma, revestido de carne: las expresiones que constituyen su contenido; y del lógico formal como la de un simple inventor de símbolos para representar las "constantes lógicas", es decir, los componentes estructurales. Pero esto simplifica demasiado. Una mejor imagen pienso que es la siguiente: se reconocen similaridades estructurales entre los argumentos informales, similaridades indicadas especialmente por la ocurrencia de ciertas expresiones, tales como "y" o "a menos que" o "todo". (Pero no deberíamos esperar que cada argumento informal tenga necesariamente un lugar único en este modelo.) El lógico formal selecciona, entre las expresiones cuya ocurrencia indica similaridades estructurales, aquellas que son (por varias razones, de funcionalidad veritativa, por ejemplo, cfr. cap. 3, § 2) prometedores candidatos para un tratamiento formal.

Esta imagen —aun siendo superficial— ya empieza a explicar por qué intentos de precisar qué expresiones del lenguaje natural deberían considerarse como "constantes lógicas" tuvieron la tendencia a concluir con la admisión, en cierto modo incómoda, de que no todas las expresiones adecuadamente "neutrales respecto al tópico" (Ryle, 1954), no todas las expresiones que parecen ser esenciales para la validez de los argumentos informales (von Wright, 1957), están representadas en el simbolismo de la lógica formal; por ejemplo, "varios" es tan neutral respecto al tópico y puede ser tan esencial para un argumento como "todo"; el material de los lógicos formales incluye un análogo para el último, pero no para el primero. Compárese con la enumeración de Quine de las constantes lógicas: "... partículas básicas tales como 'es', 'no', 'y', 'o', 'a menos que', 'si', 'entonces', 'ni...ni', 'alguno', 'todos', etc." (1940, pág. 1); hay que hacer notar que la lista comprende exclusivamente aquellas expresiones del lenguaje natural que pueden representarse cómodamente en el cálculo clásico de oraciones y predicados, y que excluye "necesariamente" y "posiblemente", por ejemplo, sin duda debido al esceptismo de Quine acerca de la inteligibilidad de la lógica modal. El "etc.", por supuesto, no ayuda en absoluto, ya que no se nos da ninguna indicación sobre qué se consideraría como una permisible adición a la lista.

La relación entre los argumentos informales y sus representaciones formales, como nos hace suponer lo anterior, no es claramente de uno a uno. Un argumento informal puede representarse apropiadamente de varias maneras en diferentes formalismos; por ejemplo:

> Todo número natural es mayor que o igual a cero, y todo número natural es par o impar, por tanto, todo número natural es mayor que o igual a cero y par o impar

podría representarse correctamente en el cálculo de oraciones como:

$$\frac{p}{q}$$

y en el cálculo de predicados como:

$$\frac{(x)Fx \;\&\; (x)Gx}{(x)(Fx \;\&\; Gx)}.$$

(Obsérvese que el que se disponga de representaciones alternativas no tiene por qué depender de ninguna ambigüedad en el original, aunque si un argumento informal *es* ambiguo, esto, naturalmente, significará que tiene más de una representación formal; cfr. el argumento espléndidamente ambiguo de Anscombe "si puedes comer pescado, puedes comer pescado".)

"*p*, por tanto *q*" es inválido, pero "$(x)Fx$ & $(x)Gx$, por tanto $(x)(Fx$ & $Gx)$" es válido; y ya que el último revela más la estructura del argumento informal original que el primero, se puede tener

la tentación de pensar que la mejor representación formal será aquella que ponga de manifiesto al máximo la estructura. Pero mi argumento informal puede representarse, una vez más, en el simbolismo del cálculo de predicados, con otra estructura manifestada por:

$$\frac{(x)(Fx \vee Gx) \;\&\; (x)(Hx \vee Ix)}{(x)((Fx \vee Gx) \;\&\; (Hx \vee Ix))}$$

Es evidente que hay un sentido en el cual ésta pone de manifiesto más estructura de la necesaria; es preferible pensar que la representación formal óptima es aquella que revela la mínima estructura coherente para proporcionar un argumento formal que es válido en el sistema si el argumento informal se considera extrasistemáticamente válido. Ésta es la *máxima del análisis de superficie* de Quine (1960a, pág. 160): "donde no te pique, no te rasques". IK

En la interacción entre *logica utens* y *logica docens*, sugerí (página 36) que se puede considerar que vale la pena sacrificar los juicios pre-formales de validez en favor de la uniformidad de la teoría formal, o modificar nuestra teoría formal para adaptar valoraciones de los argumentos informales, o —y ésta es la cuestión que quiero seguir aquí— revisar nuestra opinión sobre el modo adecuado de representar un argumento informal en lógica formal. Un criterio mediante el cual se juzga si un argumento informal está correctamente representado por un argumento formal dado, es aquel en el que se respetan los juicios intuitivos de validez. Por ejemplo, nuestra seguridad de que "Alguien es Primer Ministro y alguien es Reina, por tanto, el Primer Ministro es Reina" es inválido nos llevaría a resistirnos a representarlo mediante un argumento formal válido en el cálculo de predicados, como:

$$a = b$$
$$\frac{a = c}{b = c}$$

y exigir algo parecido al argumento inválido

$$\frac{(\exists x)Fx \;\&\; (\exists x)Gx}{(\iota x)Fx = (\iota x)Gx}$$

Si, por el contrario, se considera que un argumento informal es válido, buscaremos una representación por medio de un argumento formal válido. Por ejemplo, dentro de los límites del cálculo de predicados estándar, los predicados modificados adverbialmente se representan normalmente mediante nuevas letras predicativas, así pues, un argumento como

> El Presidente firmó el tratado con una pluma roja.
> Por tanto, el Presidente firmó el tratado.

se representaría por:

$$\frac{Fa}{Ga}$$

donde "a" representa "el Presidente", "F" representa "firmó el tratado con una pluma roja" y "G" "firmó el tratado". Por supuesto, este es un argumento inválido en el cálculo de predicados, y, por tanto, dada la supuesta validez del argumento informal original, se ha instado a concebir un medio más claro de representar la modificación adverbial, que no sea el de eliminar la conexión lógica entre un predicado modificado adverbialmente y su forma sin modificar. Davidson, por ejemplo (1968a), propone una representación de la siguiente forma:

$$\frac{(\exists x)(x \text{ era una firma del tratado por el Presidente y } x \text{ estaba hecha con una pluma roja})}{(\exists x)(x \text{ era una firma del tratado por el Presidente})}$$

la cual, como el original, es válida. Obsérvese que esto proporciona al argumento original una representación dentro del cálculo de predicados estándar al cuantificar sucesos y tratar los adverbios como predicados de sucesos; otra posibilidad sería ampliar el formalismo estándar, por ejemplo, mediante la adición de operadores de predicado para representar adverbios. En el caso de los adverbios modales "necesariamente" y "posiblemente" esta especie de ampliación del vocabulario de la lógica formal ya ha tenido lugar.

Algunos filósofos de la lógica han instado a la demanda de una imagen más limpia, según la cual cada argumento informal tiene una forma lógica única —quizás no reconocible de inmediato— que la representación simbólica correcta pondrá de manifiesto. A esta opinión se mantuvieron fieles Wittgenstein y Russell durante sus períodos de atomismo lógico (véase, por ejemplo, Russell, 1918; Wittgenstein, 1922; y cfr. los comentarios a la teoría de las descripciones de Russell en el cap. 4, § 3); pues aspiraban a inventar un lenguaje único, idealmente perspicuo, en el que la forma lógica se pondría perfectamente de manifiesto. Más recientemente, Davidson ha adoptado una postura similar: para Davidson, la forma lógica de un argumento es su representación en un lenguaje formal

para el que la verdad puede definirse de acuerdo con los imperativos impuestos por la teoría de Tarski (cap. 7, § 5). Russell pensó que la forma gramatical de una oración era susceptible de resultar engañosa al tomarla como su forma lógica; algunos autores recientes, impresionados por la propuesta de Chomsky de una estructura gramatical profunda que subyace a la estructura gramatical superficial, pero que quizás es bastante diferente de ella (véase Chomsky, 1957), han sugerido que la forma lógica de un argumento debería identificarse con su estructura gramatical profunda (véase, por ejemplo, Harman, 1970). La estructura gramatical/lógica profunda relevante tendría que ser, probablemente, universal para todos los lenguajes, ya que, de otro modo, correríamos el riesgo de permitir que un argumento fuera, por ejemplo, válido en hebreo, pero inválido en hindú; y, a mi entender, es dudoso si se tiene derecho a suponer que los lingüistas descubrirán con el tiempo una estructura gramatical universal y suficientemente rica. Así pues, no puedo ser completamente optimista acerca de las perspectivas de este panorama —es cierto que satisfactoriamente ordenado—. Pero, sin embargo, no veo razón para desalentarse ante la interdependencia entre los juicios de validez informales e intuitivos, presentimientos en cuanto a los rasgos estructurales esenciales de los argumentos informales, y el desarrollo de los sistemas lógicos formales. Más bien, incluso se podría sentir cierta satisfacción en el sentido de que esta interdependencia explica por qué las cuestiones centrales de la filosofía de la lógica se agrupan en torno al tema de la adecuación entre los argumentos informales y sus representaciones formales: un tema que, en lo que respecta a conectivas, cuantificadores y términos singulares, investigarán más a fondo los tres capítulos siguientes.

3
Conectivas de oraciones

1 CONSIDERACIONES FORMALES

Comenzaré esbozando algunas características importantes de las conectivas de oraciones, y continuaré con la consideración de algunas cuestiones filosóficas acerca de los significados de las conectivas.

Conjuntos adecuados de conectivas: completud funcional

Las conectivas —"$-$", "&", "$\vee$", "$\rightarrow$" y "$\equiv$"— de un cálculo clásico son veritativo-funcionales: el valor de verdad de una oración compuesta formada por medio de ellas depende sólo de los valores de verdad de sus componentes. Un conjunto de conectivas es *adecuado* si puede expresar todas las funciones de verdad. Hay 16 (2^{2^2}) funciones de verdad bivalentes de dos argumentos[1]. Cada uno de los conjuntos $\{-, \rightarrow\}$, $\{-, \vee\}$, $\{-, \&\}$, $\{|\}$ y $\{\downarrow\}$ ("$A|B$" es "no a la vez A y B" y "$A \downarrow B$" es "ni A ni B") es adecuado para expresarlas todas. Un sistema formal es *funcionalmente completo* si tiene un conjunto adecuado de conectivas. Por ejemplo, el de *Principia*, con "$-$" y "$\vee$" como primitivas, es funcionalmente completo, mientras que el fragmento del cálculo de oraciones referido a la implicación, con tan sólo "$\rightarrow$", no lo es. Muchas formulaciones —por ejemplo, la de Lemmon, 1965— tienen más conectivas de las que se necesitan para la completud funcional. Esto es así porque hay conjuntos alternativos adecuados de conectivas de los que se obtienen formulaciones del cálculo de oraciones con diferentes conjuntos de conectivas primitivas. Dado cualquier conjunto adecuado, las otras conectivas se pueden definir. Por ejemplo, con "$-$" y "$\rightarrow$" como primitivas, "$A \vee B$" puede definirse como "$-A \rightarrow B$", y entonces "A % B" como "$-(-A \vee -B)$"; con "$|$" o "$\downarrow$" como primitivas "$-A$" puede definirse como "$A|A$" o "$A \downarrow A$". Algunas formulaciones utilizan una constante "F", que siempre tiene el valor *f*, y define "$-A$" como "$A \rightarrow F$". En cada caso la corrección de las definiciones se puede comprobar comparando las tablas de verdad del *definiens* y el *definiendum* y observando que corresponden a la misma función de verdad.

Matrices características: decidibilidad

Una matriz, o conjunto de tablas de verdad, M, es *característica* de un sistema S sii todos y solamente los teoremas de S están designados en M, y todas y solamente las inferencias válidas de S preservan la designación en M. Puede designarse cualquier caso, pero normalmente lo que interesa es designar el valor "como verdad", o, quizás, los valores, en el caso de las lógicas plurivalentes; en la lógica bivalente, por supuesto, "V" está designada. Una fbf está designada en M sii toma un valor designado cualquiera que sea la asignación que se haga a sus partes atómicas; una regla, de $A \ldots A_n$ se infiere B, es preservadora de designación sii B toma un valor designado cada vez que lo hacen $A \ldots A_n$. Por ejemplo, las tablas de verdad bivalentes son características del cálculo de oraciones clásico.

Las tablas de verdad finitas proporcionan un *procedimiento de decisión*, es decir, un método mecánico para determinar, si es un teorema cualquiera de las fbfs del sistema.

Lógica plurivalente

Por supuesto, sería posible imaginar las matrices plurivalentes características para el cálculo bivalente de oraciones. Le llamo lógica bivalente de oraciones, "bivalente" más bien que "plurivalente", porque es el número más pequeño de valores que puede proporcionar una matriz característica. Por una "lógica *n*-valente" entenderé un sistema que tiene una matriz característica con n valores y ninguna matriz característica con m valores, siendo $m < n$. Algunos de los sistemas a los que he calificado como "desviados" tienen matrices características finitas; una de las motivaciones no

---

[1] A saber:

A	*B*	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
v	*v*	*v*	*v*	*v*	*v*	*v*	*v*	*v*	*v*	*f*	*f*	*f*	*f*	*f*	*f*	*f*	*f*
v	*f*	*v*	*v*	*v*	*v*	*f*	*f*	*f*	*f*	*v*	*v*	*v*	*v*	*f*	*f*	*f*	*f*
f	*v*	*v*	*v*	*f*	*f*	*v*	*v*	*f*	*f*	*v*	*v*	*f*	*f*	*v*	*v*	*f*	*f*
f	*f*	*v*	*f*	*v*	*f*	*v*	*f*	*v*	*f*	*v*	*f*	*v*	*f*	*v*	*f*	*v*	*f*

improvisada, para inventar tales sistemas, ha sido la creencia de que algunas oraciones del ámbito de la lógica no son ni verdaderas ni falsas, sino que o no tienen valor de verdad, o, quizás, tienen un valor de verdad intermedio: creencia que recibirá mayor atención en el cap. 11. Otros sistemas desviados, tales como el intuicionista y algunas lógicas cuánticas, no tienen matrices características finitas, sino tan sólo infinitas. En lo que sigue, "lógica plurivalente" significará "lógica n-valente, siendo $2 < n < \infty$", excepto cuando hable especialmente de sistemas *infinitamente* polivalentes.

En una lógica n-valente, cada lugar dado en una tabla de verdad puede ser ocupado por cualquiera de los n valores; así pues, ya que la tabla de verdad para una conectiva de k posiciones tiene n^k entradas, el número de funciones de verdad de k argumentos en una lógica n-valente será n^{n^k} —número que se incrementa enormemente para pequeños incrementos de n—. La lógica trivalente de Lukasiewicz, con "$-$", "&", "$\vee$", "$\rightarrow$" y "$\equiv$", es funcionalmente incompleta; Slupecki mostró que se convierte en funcionalmente completa con la adición de una nueva conectiva composicional T ("*tertium*") tal que "TA" toma el valor intermedio cualquiera que sea el valor de A. Como era de esperar, las habituales relaciones de interdefinibilidad puede que lleguen a venirse abajo en la lógica plurivalente; por ejemplo, en la lógica trivalente de Lukasiewicz "$A \vee B$" no tiene, como en la lógica bivalente, la misma tabla de verdad que "$-A \rightarrow B$"; en su lugar, puede definirse como "$(A \rightarrow B) \rightarrow B$".

2 LOS SIGNIFICADOS DE LAS CONECTIVAS

Lenguajes formales y lecturas informales

El cálculo de oraciones puede, por decirlo así, enfocarse desde cuatro niveles:

(i) los axiomas/reglas de referencia
(ii) la interpretación formal (matrices)
(iii) las lecturas en lenguaje ordinario de (i)
(iv) la explicación informal de (ii)

(i) es el nivel de la sintaxis; a los niveles (ii) y (iv) Plantinga (1974, págs. 126-8) los ha apodado, respectivamente, semántica "pura" y "depravada". Los niveles (i) y (ii), al ser formales, son agradablemente manejables; pero los niveles (iii) y (iv), aunque más complicados, no son menos importantes. Observé en el cap. 1 que la identificación de un sistema como un sistema de lógica exige recurrir a su (¿pretendida?) interpretación. Para identificar un sistema

como un cálculo de oraciones no sólo es necesario conocer los axiomas/reglas y su interpretación formal mediante matrices, sino que también es necesario conocer que los valores representan verdad y falsedad, las letras "p", "q", etc., oraciones, "$-$" negación, "&" conjunción, "$\vee$" disyunción, y así sucesivamente. El conocimiento de las conectivas probablemente tiene que derivar de alguna manera de alguno o de todos estos niveles.

La opinión acerca de cómo las conectivas obtienen su significado afectará a la actitud frente a varios temas. Por ejemplo, se ha mantenido que en las lógicas desviadas las conectivas difieren en significado de las conectivas tipográficamente idénticas de la lógica clásica; así pues, cuando un lógico desviado niega, "$A \vee -A$", lo que niega no es, contrariamente a las apariencias, lo que el lógico clásico asevera cuando dice "$A \vee -A$". El argumento para esta tesis de "varianza del significado" (cfr. cap. 12, § 1) sería el de que el significado de las conectivas viene dado simplemente por los axiomas/reglas y/o matrices del sistema (niveles (i) y (ii)), de lo que se sigue que las conectivas de la lógica plurivalente tienen que diferir en significado de las de la lógica bivalente, ya que difiere en los axiomas/reglas y las matrices. Otra disputa se refiere a la adecuación de las lecturas castellanas de las conectivas, con qué precisión, por ejemplo, "y" representa "&", o "si", "$\rightarrow$". Lo que parece que aquí está en discusión es si los axiomas/reglas son verdaderos/preservadores de verdad y las matrices correctas, cuando se piensa en caracterizar como lecturas las expresiones castellanas utilizadas: pues si, ya que "A & B" es V sii "A" es V y "B" es V, "A y B" es verdadero sii "A" es verdadero y "B" es verdadero. Esto, a su vez, plantea otra pregunta: ¿importa si hay una discrepancia?

"*tonk*"

Prior ha argumentado que los significados de las conectivas no pueden derivar de los axiomas/reglas del sistema en el que aparecen, ni de sus tablas de verdad, sino que debe ser dado por sus lecturas castellanas. (Si estuviera acertado, por supuesto, la opinión de la "varianza del significado" de las lógicas plurivalentes, mencionada anteriormente, sería refutada.) Prior (1960, 1964) presenta una pretendida *reductio ad absurdum* de la tesis de que hay "inferencias analíticamente válidas", es decir, inferencias cuya validez es únicamente resultado de los significados de las constantes lógicas contenidas en ellas. De acuerdo con esta tesis, la inferencia de "A" a partir de "A & B" es analíticamente válida, pues el significado de "&" viene enteramente dado por las reglas de inferencia de la introducción de & y la eliminación de &. Prior argumenta que "según este sentido de 'analíticamente válido' cualquier enunciado puede inferirse, de un modo analíticamente válido, de cualquier otro enunciado" (1960, pág. 130). Supongamos que el significado de "*tonk*" está dado por las reglas de inferencia:

(T_1) de "*A*" se infiere "*A tonk B*" (introducción de *tonk*)
(T_2) de "*A tonk B*" se infiere "*B*" (eliminación de *tonk*)

Utilizando estas reglas, $A \vdash B$, para cualesquiera *A* y *B*:

(1)	*A*	supuesto
(2)	*A tonk B*	(1), (T_1)
(3)	*B*	(2), (T_2)

De este modo, por supuesto, un sistema con (T_1) y (T_2) sería inconsistente. Nada vital, excepto los axiomas, depende de la utilización de las reglas de inferencia de Prior; los axiomas "$A \rightarrow (A$ *tonk* $B)$" y "$(A$ *tonk* $B) \rightarrow B$", junto con la regla para inferir *B* de $A \rightarrow B$ y *A* (*modus ponens*, en adelante MPP), nos llevarían a consecuencias igualmente alarmantes; véase Prior, 1964, pág. 192.

Prior cree haber mostrado que la noción de una inferencia analíticamente válida es una confusión, y que "una expresión debe tener algún significado independientemente determinado antes de que podamos descubrir si las inferencias que le conciernen son válidas o inválidas" (1960, págs. 129-30).

Argumenta Prior que así como las reglas (T_1) y (T_2) no pueden dar el significado de "*tonk*", los significados de las conectivas, en general, no pueden ser dados por los axiomas/reglas en los que ocurren. Sin embargo, se podría replicar que (T_1) y (T_2) no logran precisar el significado de "*tonk*" por la razón suficiente de que son reglas *defectuosas*. Permiten que $A \vdash B$ para cualesquiera *A* y *B*; y no hay ningún sistema en el que algo es derivable de algo que no tenga alguna posibilidad de distinguir las inferencias aceptables de las inaceptables (cfr. Belnap, 1961; Stevenson, 1961). Prior no ha mostrado que las reglas de inferencia aceptables no puedan dar el significado de las conectivas que ocurren en ellas.

Sugerí antes que el objetivo principal de la construcción de los sistemas formales de lógica es dar axiomas/reglas tales que las inferencias informales posibles de expresar en el lenguaje del formalismo, que se consideran intuitivamente válidas en sentido extrasistemático, sean válidas en el sistema. L_T sería tan defectuoso que no tiene ninguna posibilidad de éxito en esta empresa.

Propósitos de la formalización

Es necesario decir algo más, sin embargo, acerca de la forma en que los sistemas lógicos formales aspiran a representar las inferencias intuitivamente válidas. Se podría pensar que un sistema lógico formal ha sido concebido más o menos del siguiente modo. Algunos argumentos informales se consideran como intuitivamente válidos, otros como inválidos. Entonces se construye un lenguaje formal en el que las características estructurales relevantes de esos argumentos pueden representarse esquemáticamente, y los axiomas/reglas que permiten los argumentos aprobados intuitivamente y prohiben los intuitivamente desaprobados. Esta es, por supuesto, en el mejor de los casos, una "reconstrucción racional" *muy* esquemática y sin ánimo de considerarla como historia seria y detallada. Sin embargo, aunque concedo que las lógicas formales a veces se han concebido simplemente por curiosidad matemática, pienso que algo parecido al proceso que he descrito estaba funcionando cuando, por ejemplo, Frege concibió su *Begriffsschrift*. Por supuesto, los lenguajes lógicos estándar son ahora tan familiares que no se es muy consciente de cómo y por qué se construyeron por primera vez. Pero el mismo proceso se puede ver en los recientes intentos de idear nuevos formalismos para los tipos de argumento desatendidos hasta el momento; véase, por ejemplo, el procedimiento adoptado por D. K. Lewis en 1973 para elaborar su análisis de los contrafácticos.

Bien, suponiendo que lo anterior es rigurosamente exacto, ¿cuál es su importancia en las cuestiones acerca de los significados de las conectivas? Pienso que algo así: primero, *ambas*, la sintaxis y la semántica pura (niveles (i) y (ii)) *y* las lecturas informales y la semántica depravada (niveles (iii) y (iv)) se puede suponer que contribuyen a los significados de las conectivas; pues parte del objetivo de la empresa es tener los niveles (i) y (ii) representados adecuadamente en (iii) y (iv).

Sin embargo, si la lógica formal siguiera fielmente a los argumentos informales en toda su complejidad y vaguedad habría una pequeña cuestión en la formalización: el propósito, al formalizar, de generalizar, simplificar y aumentar la precisión y el rigor. Esto significa, pienso, que ni se espera ni se desea una representación formal directa de todos los argumentos informales considerados, extrasistemáticamente, válidos. Sino que, más bien, los juicios presistemáticos de validez proporcionarán datos para la construcción de una lógica formal, pero las consideraciones de simplicidad, precisión y rigor es de suponer que lleven a discrepancias entre los argumentos informales y sus representaciones formales, e incluso en algunos casos quizás a una nueva valoración de los juicios intuitivos. Se utilizan los juicios intuitivos de algunos argumentos para construir una teoría formal que da veredictos, quizás veredictos bastante inesperados, sobre otros argumentos; y con el tiempo podrían sacrificarse algunos de los juicios originales a las consideraciones de la simplicidad y la generalidad. Este punto se relaciona, por supuesto, con la interdependencia entre los juicios sobre la corrección en la traducción de un argumento informal a un lenguaje formal y la opinión presistemática acerca de su validez, comentada en el cap. 2. (Un ejemplo sería la interpretación estándar de "Todos los F son G" como "$(x)(Fx \rightarrow Gx)$", que es verdadero si su antecedente es falso, es decir, si no hay ningún F. Es bastante dudoso si presistemáticamente se estaría de acuerdo en que todos los unicornios son violeta, y bastante cierto que no se estaría de acuerdo en que todos los unicornios son violeta *y* todos los unicornios son naranja.)

Se admitiría, entonces, que un fracaso por parte de un sistema formal en representar *todos* los entresijos de los argumentos informales que sistematiza no es necesariamente objetable. Por otra parte, se debe tener cuidado con suponer que *todos* los ajustes son aceptables; es necesario preguntarse si los aumentos de simplicidad y generalidad compensan la discrepancia. Algunos de los entresijos del castellano pueden ser importantes. Estas observaciones pueden parecer desagradablemente vagas; trataré de hacerlas más específicas considerando unos ejemplos.

¿Por qué las lógicas formales corrientes tienen, por ejemplo, "&" que se lee "y", pero ningún término formal análogo para "porque" o "pero"; y "($\exists$ ...)" que se lee "al menos uno", pero ningún término formal para "varios" o "unos pocos"? Dos características de las expresiones favorecidas sugieren: funcionalidad de verdad y precisión.

"&" es veritativo-funcional; y las funciones de verdad son en especial fácilmente sensibles al tratamiento formal —es de notar que permiten la posibilidad de un procedimiento de decisión mecánica. Por esto es, sin duda, por lo que el lógico formal dispone de un análogo para "y", pero ninguno para "porque" o "pero"; "y", al menos en una amplia clase de usos, es veritativo-funcional, mientras que el valor de verdad de "A porque B" depende no sólo de los valores de verdad de "A" y "B", sino también de si B es razón para A, y el valor de verdad de "A pero B" depende también de si la combinación de A y B es sorprendente. "Al menos uno" y "todos" no son funciones de verdad (aunque en el caso particular de un universo finito son equivalentes a "$Fa \vee Fb \vee \ldots \vee Fn$" y "$Fa$ & Fb & ... & Fn", respectivamente). Pero son precisas —a diferencia de "varios" y "algunos pocos"—. Hay que hacer notar que la habitual lectura de "($\exists$ ...)" como "algún" es más vaga que "($\exists$ ...)"; "al menos uno" es una lectura más exacta. (Otras ciencias comparten la tendencia de la lógica a precisar e idealizar; compárense los puntos inextensos de la geometría o las superficies sin fricción de la mecánica.)

Sin embargo, mientras que está claro que las expresiones precisas y veritativo-funcionales son preferibles desde el punto de vista de la simplicidad y el rigor, a las expresiones vagas o no veritativo-

-funcionales, no está tan claro que esta preferencia sea primordial. Pues operadores no veritativo-funcionales —por ejemplo, "L" y "M" para "necesariamente" y "posiblemente"— son usados por el lógico formal. Von Wright ha sugerido (1963) un sistema con una conectiva de oraciones "T" que se lee "y entonces" que preserva el sentido temporal que a veces tiene "y" en castellano[2]. Y, mientras el cálculo de predicados estándar se reduce asimismo a "todos" y "al menos uno", Altham ha inventado una lógica con cuantificadores para "muchos" y "pocos".

La conveniencia de la funcionalidad de verdad es bastante indiscutible; pero igualmente está claro que una lógica reducida a las funciones de verdad estaría inaceptablemente limitada. Hasta qué punto es esencial la precisión en la empresa de la lógica formal es más controvertido. La objeción de Dummett a reconocer las "lógicas" epistémicas como lógicas genuinas, se recordará, era que "sabe" y "cree" son inherentemente vagos. Otros lógicos, sin embargo, han hecho uso deliberadamente de las ideas vagas. Por ejemplo, en un análisis de condicionales contrafácticos "Si fuera el caso de que A, sería el caso de que B", D. K. Lewis (1973, especialmente cap. 4) propone emplear la ciertamente vaga idea de similitud entre mundos posibles (aproximadamente, "en todos aquellos mundos posibles más similares al mundo actual, pero en los que A, B"); defiende su compromiso con la vaguedad al observar que la vaguedad del *analysans* no es objetable, ya que el *analysandum* es él mismo vago. Zadeh, con su "lógica vaga" (véase, por ejemplo, 1975) propone una salida incluso más radical para la preocupación tradicional de la lógica por la precisión. Dudo de si tales salidas están justificadas por sus resultados hasta el momento, pero serían necesarios muchos más argumentos para mostrar que esta duda está bien fundada; véase cap. 9, § 4.

"&" *e* "*y*", "$\vee$" *y* "*o*", *etc.*

Sobre las lecturas "no" (de "$-$"), "y" (de "&"), "o" (de "$\vee$") y "si ... entonces ---" (de "$\rightarrow$"), Strawson ha observado (1952, página 79) que "los dos primeros son los menos engañosos" y los restantes "definitivamente erróneos". Naturalmente, hay discrepancias.

Mientras en el cálculo de oraciones "$-$" es un operador de formación de oraciones, "no" en castellano puede negar o una oración

[2] Una cuestión relacionada con esto es que el aparato lógico estándar es insensible a las consideraciones temporales; normalmente se nos aconseja entender los "p" y "q" atemporalmente. Algunas propuestas de lógicas temporales se discuten en el cap. 9, § 3.

entera o su predicado. Esta distinción (entre negación "externa" e "interna") se ha considerado importante para la comprensión de oraciones supuestamente carentes de sentido; por ejemplo, se ha sugerido que "Virtud no es triangular", así como "Virtud es triangular" es carente de sentido, mientras "No es el caso de que la virtud es triangular" es verdadero. También se ha observado que en el habla coloquial las dobles negaciones no siempre se "anulan", sino que pueden usarse como negativas con énfasis. "Y", como ya he observado, a veces se usa en el sentido de "y entonces", mientras "&" es indiferente al orden temporal.

Algunos han argumentado que "o" tiene dos sentidos, uno inclusivo y el otro exclusivo; pero esto no sería una divergencia demasiado seria con el " $\vee$ " del cálculo de oraciones, ya que una disyunción exclusiva podría definirse como "$(A \vee B) \& -(A \& B)$". Un segundo argumento a favor de una discrepancia entre "o" y " $\vee$ " recurre al hecho de que, en el habla corriente, bien podría ser seriamente engañoso aseverar "John tiene el libro o lo tiene Mary" si se estuviera en posición de aseverar "John tiene el libro". Sin embargo, se podría mantener que la rareza del análogo de la regla de introducción de $\vee$ (de "A" inferir "$A \vee B$") en el discurso ordinario es más una cuestión de lo que Grice llamó *implicación conversacional* que de validez. De acuerdo con la postura de Grice, un hablante implica conversacionalmente que B si su *aseveración de que* A da a su oyente razón para creer que él cree que B. Ya que aseverar "A o B" cuando se tiene derecho a aseverar "A" (o "B") contraviene uno de los principios del candor conversacional de Grice: que no se puede hacer una afirmación más débil cuando se tiene derecho a hacer una más fuerte —un hablante que asevera "A o B" implica conversacionalmente que no sabe si es A o B lo que es verdadero—. Esta explicación, ya que no se refiere a los valores de verdad de las aseveraciones, permite aceptar que "A o B", así como "$A \vee B$", es verdadero justamente en el caso de que "A" es verdadero o "B" es verdadero, y de este modo se justificaría la aparente discrepancia.

Las discrepancias entre "$\rightarrow$" y "si ... entonces" se han considerado generalmente como las más serias. Parece ser bastante más aceptado que si "Si A entonces B" es verdadero, entonces "$A \rightarrow B$" es verdadero, pero es altamente polémico que si "$A \rightarrow B$" es verdadero, "Si A entonces B" sea verdadero. Faris argumenta que "Si A entonces B" es derivable de "$A \rightarrow B$", de modo que "$A \rightarrow B$" y "Si A entonces B" son interderivables, si no sinónimos. Supone que una condición necesaria y suficiente para la verdad de "Si A entonces B" es la *condición E*: hay un conjunto S de proposiciones verdaderas tal que B es derivable de A junto con S. Si "$A \rightarrow B$" es verdadero, continua Faris, hay un conjunto de proposiciones verdaderas, a saber, el conjunto del que "$A \rightarrow B$" es el único miembro, del cual, junto con A, es derivable B; de este modo se satisface E, y "Si A

entonces B" es verdadero. El argumento de Faris ha sido atacado en varios puntos; comprensiblemente, los objetores parecen convencidos de que la conclusión es errónea, pero menos seguros sobre dónde está el defecto del argumento (véase, por ejemplo, Bajer, 1967; Clark, 1971; L. J. Russell, 1970). Puede ser interesante observar que el argumento de Faris depende profundamente de una noción de "derivabilidad" que comprende a los lenguajes naturales y formales de una manera en cierto modo irregular. Otros autores han argumentado que las aparentes discrepancias entre "$\rightarrow$" y "si" son más una cuestión de implicación conversacional que de condiciones de verdad. Su explicación sería algo parecido a esto: no es que "Si A entonces B" sea falso si "A" es falso o "B" verdadero, sino más bien que cuando no hay ninguna conexión entre "A" y "B" sería inútil y engañoso aseverar "Si A entonces B" si se tuviera el derecho de aseverar "$-A$" o "B" (véase, por ejemplo, Johnson, 1921; Moore, 1952). Otros, además, han sugerido que "si" tiene varios usos en lenguaje ordinario, uno de los cuales puede corresponder exactamente a "$\rightarrow$", pero los otros requieren una representación diferente (véase, por ejemplo, Mackie, 1973, donde se distinguen nueve usos y seis interpretaciones del condicional en inglés).

La lógica moderna ofrece, de hecho, más de una clase de condicional. El condicional material, que he discutido hasta ahora, es veritativo-funcional; y "$A \rightarrow B$" es verdadero si o "A" es falso o "B" es verdadero. De este modo tiene los teoremas:

$$A \rightarrow (B \rightarrow A)$$
$$-A \rightarrow (A \rightarrow B)$$
$$(A \rightarrow B) \vee (B \rightarrow A)$$

Estas son las "paradojas de la implicación material". Las "paradojas" resultan si se lee "$\rightarrow$" como "si" o "implica"; C. I. Lewis comenta que el tercero de los teoremas anteriores dice que si se toman de un periódico dos oraciones cualesquiera al azar, o la primera implicará la segunda o la segunda a la primera. La reflexión sobre estas "paradojas" llevó a Lewis a proponer un condicional más fuerte, "$A \dashv B$", donde "$\dashv$" es la implicación *estricta*, definida como "Necesariamente $(A \rightarrow B)$". "Necesariamente $(A \rightarrow B)$", dada la semántica estándar para la lógica modal, se supone que es verdadero si B es verdadero en todos los mundos posibles en los que A es verdadero. Otras relaciones de implicación, inspiradas en la implicación estricta, se han presentado en el análisis de contrafácticos (véase Stalnaker, 1968; D. K. Lewis, 1973).

Sin embargo, la implicación estricta tiene sus propias paradojas; en resumen, así como una proposición falsa implica materialmente cualquier proposición, del mismo modo una proposición imposible implica estrictamente cualquier cosa, y cualquier cosa implica estrictamente una proposición necesaria. Lógicos de relieve proponen, en consecuencia, un condicional más estricto todavía, que exige una relación de relevancia entre el antecedente y el consecuente (véase Anderson y Belnap, 1975, § 1). Estos lógicos se oponen a llamar a "→" "implicación material", así como a leerlo "Si ... entonces ---"; "negación inmaterial", sugieren, no sería más inadecuado. También extienden su crítica de la lógica veritativo-funcional a la disyunción —recuérdese que en el sistema estándar "$A \rightarrow B$" es equivalente a "$-A \vee B$"— argumentando que el "o" informal es, como el "si", intensional.

La cuestión aquí es qué condicional corresponde mejor al "si"; a lo cual, por supuesto, la respuesta puede ser que diferentes condicionales formales corresponden mejor a diferentes usos o sentidos de "si". Otra cuestión es, admitiendo que la implicación material, al ser veritativo-funcional, es el más simple de los condicionales formales, si recurrir a los condicionales estrictos o subjuntivos o relevantes proporciona ventajas que compensen la pérdida de simplicidad. Y aquí, pienso, los propósitos para los que se emprende la formalización pueden ser cruciales. Si sólo se está interesado en representar formalmente los argumentos válidos que se usan en matemáticas, por ejemplo, podría ser que una implicación veritativo-funcional fuera adecuada; aunque incluso esto es discutible (véase Anderson y Belnap, 1975, § 3). Si, por otra parte, se está también interesado en representar argumentos de la ciencia empírica, puede ser que, ya que la ciencia está en apariencia profundamente confiada a las predisposiciones, y, por tanto, a los condicionales subjuntivos ("x es soluble" o "Si x se pusiera en agua, se disolvería"), sea apropiado necesitar algo más fuerte; pero esto también es discutible (véase, por ejemplo, Goodman, 1955, o Quine, 1973, páginas 8-16). Así pues, la significatividad de las discrepancias entre "si" y "→" dependerá de las respuestas por lo menos a dos preguntas más: ¿con qué propósito(s) se quiere hacer la formalización? y ¿requiere este propósito algo más fuerte que el condicional material? Ambas —tal como las veremos en el curso de un examen más cercano de los condicionales estrictos y relevantes en el capítulo 10— son preguntas difíciles y profundas.

4
Cuantificadores

1 Los cuantificadores y su interpretación

"$(x)Fx$" se lee usualmente "Para todo x, Fx" y "$(\exists x)Fx$" se lee "Para algún x, Fx" o, más exactamente, "Para, al menos, un x, Fx"; "$(...)$" es conocido generalmente como el cuantificador *universal*, "$(\exists\ ...)$" como el cuantificador *existencial*. De una variable que se encuentra dentro del alcance de un cuantificador, tal como "x" en "$(\exists x)Fx$", se dice que está *ligada*, de una variable que no está ligada por ningún cuantificador, tal como "x" en "Fx", o "y" en "$(\exists x)Rxy$", se dice que está *libre*. Una fórmula con una o más variables libres se denomina *oración abierta* (de 1, 2... n posiciones), una fórmula sin variables libres se denomina *oración cerrada* (u "oración abierta de 0-posiciones"). Así, prefijando el cuantificador, "(x)" o "$(\exists x)$", a una oración abierta tal como "Fx", precisamente con "x" libre, produce una oración cerrada, "$(x)Fx$" o "$(\exists x)Fx$"; en general, prefijando un cuantificador que ligue una de sus variables libres a una oración abierta con n variables libres produce una oración abierta con $n - 1$ variables libres.

Algunas formulaciones del cálculo de predicados poseen *términos singulares*, "a", "b", "c" ... etc., y también variables; éstas son constantes individuales que denotan cada una de ellas algún individuo específico. Eliminando un cuantificador y reemplazando la(s) variable(s) que éste ligaba por términos singulares, se obtiene un caso de la fórmula cuantificada, como, por ejemplo, "$Fa \rightarrow Ga$" es un caso de "$(x)(Fx \rightarrow Gx)$". Se podría pensar que las variables ligadas desempeñan un papel análogo al de los pronombres que, en los lenguajes naturales, aseguran la referencia cruzada, y que los términos singulares desempeñan un papel análogo al de los nombres propios que, en los lenguajes naturales, se refieren a individuos (pero cfr. el cap. 5).

En la lógica moderna, como acabo de indicar, cuantificadores y términos singulares pertenecen a categorías sintácticas completamente diferentes. Frege, que inventó la teoría de la cuantificación (Frege, 1879; los cuantificadores fueron ideados también, independientemente, por Peirce y Mitchell; véase Peirce, 1885), acentuó en gran manera la importancia de trasladar la atención desde la distinción sujeto-predicado a la distinción función-argumento. Una consecuencia de ello, esencial para la adecuación del formalismo en la representación del argumento matemático, es tener en cuenta las relaciones, ya que puede haber funciones de más de un argumento. Otra, más relevante para nuestros propósitos actuales, es tener en cuenta las funciones de segundo nivel, la categoría de los cuantificadores. Por ejemplo, decir que existen perros de tres patas, según Frege, es decir que el concepto *perro de tres patas* no es vacío; el cuantificador "($\exists$ ...)" es un concepto que se aplica a conceptos, una función de segundo nivel (véase Frege, 1891, 1892). Sin embargo, algunos autores han pensado que los cuantificadores del lenguaje natural, "alguno", "todo", "cada", etc., se comportan de un modo muy similar al de los nombres. Russell, por ejemplo, intentó en un tiempo tratar a estos "cuantificadores" como "expresiones denotativas"; "algún chico" era igual que "Juan" excepto en que denota un individuo "ambiguo"; pero posteriormente se decidió por una explicación al estilo fregeano (Russell, 1903; y cfr. las críticas de Geach, 1962). Los autores recientes, especialmente Montague, 1973, han perseguido la idea de tratar a los cuantificadores como nombres (y cfr. la defensa que hace Hintikka de este enfoque en 1976 y los comentarios de Fogelin y Potts). Sin embargo, limitaré mi discusión a los cuantificadores "fregeanos" usuales.

En el cálculo de predicados de *primer orden* sólo las variables "individuales" "x", "y", ... etc., pueden estar ligadas por cuantificadores: en los cálculos de *segundo orden*, "F", "G", ... etc., también pueden estar ligadas, como en "$(x)(F)Fx$". Una letra oracional, "p", "q" ... etc., puede considerarse como un caso límite de una letra predicativa; "R" en "Rxy" es un predicado biposicional, "F" en "Fx" es un predicado uniposicional y "p" en "p", un predicado de 0 posiciones. Por tanto, un cálculo de oraciones cuantificado, que permita cuantificadores que ligan "p", "q" ... etc., como en "$(p)(p \vee -p)$", es un tipo de cálculo de segundo orden. Los cálculos con diferentes clases de variables, variando cada una según las diferentes clases de cosas, tales como un formalismo con una clase de variables para los números naturales y otra para los números reales, se denominan teorías *multi-surtidas*.

Los enunciados numéricos —"Hay n x que son F"— pueden ser formulados con ayuda de los cuantificadores. "Hay *al menos un* x que es F" es:

$$(\exists x)Fx$$

y "Hay *a lo sumo* un x que es F" es:

$$(x)(y)(Fx \,\&\, Fy \rightarrow x = y)$$

(si esto no es obvio, obsérvese que la fórmula anterior se puede leer tal como "Si hay dos F, son el mismo"); por tanto, "Hay *exactamente un* x que es F" es:

$$(\exists x)(Fx \,\&\, (y))(Fy \rightarrow x = y)$$

y "Hay *exactamente dos* x que son F" es:

$$(\exists x)(\exists y)(Fx \,\&\, Fy \,\&\, x \neq y \,\&\, (z)(Fz \rightarrow z = x \vee z = y)$$

y así sucesivamente[1]. Los cuantificadores numéricos menos específicos, tales como "muchos" y "pocos", también han recibido tratamiento formal (Altham, 1971), al modo de "al menos n" y "a lo sumo n" para la variable n.

Las distinciones hechas en el capítulo anterior entre las *lecturas informales* de los símbolos de un lenguaje formal (nivel (iii), su *interpretación formal* (nivel (ii)) y la explicación informal ofrecida de la semántica formal (nivel (iv)), se aplican, por supuesto, tanto a los cuantificadores como a las conectivas oracionales. Mientras que en el caso de las conectivas la controversia principal se centra en torno a la cuestión de cómo las conectivas veritativo-funcionales representan adecuadamente a sus análogos en castellano, en el caso de los cuantificadores el tema clave atañe a su interpretación formal apropiada. Se ha observado con frecuencia que el cuantificador universal es análogo a la conjunción:

$$(\exists x)(Fx \,\&\, (y)(Fy \rightarrow x = y))$$

y "Hay *exactamente dos* x que son F" es:

$$(\exists x)(\exists y)(Fx \,\&\, Fy \,\&\, x \neq y \,\&\, (z)(Fz \rightarrow z = x \vee z = y))$$

En realidad, en una teoría cuyo dominio sea finito (por ejemplo, donde las variables fluctúan sobre los miembros del gobierno bri-

[1] Parte del programa logicista consistía en la definición de los números naturales como ciertos conjuntos; por ejemplo, 0 como el conjunto de conjuntos de 0 miembros, 1 como el conjunto de conjuntos de 1 miembro, ... n como el conjunto de conjuntos de n miembros. Obsérvese cómo esto define el uso *nominal* de los números (como en "9 > 7") en términos del uso *adjetivado* (como en "Hay 9 planetas") que, como ya he explicado, se puede expresar en términos de cuantificadores y de identidad.

tánico) una fórmula universalmente cuantificada es equivalente a una conjunción finita y una fórmula existencialmente cuantificada a una disyunción finita. Sin embargo, en una teoría de dominio infinito (por ejemplo, cuando las variables fluctúan sobre los números naturales) las fórmulas cuantificadas sólo pueden ser representadas por conjunciones o disyunciones de longitud infinita —el "... etc." es ineliminable—. Por tanto, una interpretación aceptable tendrá que proporcionar la generalidad indispensable. Y de hecho, se han ofrecido dos tipos distintos de interpretación para los cuantificadores. La *interpretación objetual* apela a los *valores* de las variables, los objetos sobre los que fluctúan las variables:

"$(x)Fx$" se interpreta como "Para todos los objetos, x, en el dominio, D, Fx"
"$(\exists x)Fx$" se interpreta como "Para al menos un objeto, x, en el dominio, D, Fx".

El dominio puede ser restringido, esto es, D puede ser especificado como un conjunto de objetos asignados como el rango de las variables —como podrían ser los números naturales, personas, caracteres ficticios, o cualquier otra cosa; o puede ser no restringido, esto es, se requiere que D sea "el universo", esto es, todos los objetos que hay. Sin embargo, los dominios restringidos asignados en el enfoque de teoría de modelos no son necesariamente subconjuntos "del universo"; el conjunto de caracteres ficticios, por ejemplo, no lo sería (cap. 5, § 4). La *interpretación sustitucional* apela, no a los valores, sino a los *sustituyentes* de las variables, esto es, a las expresiones por las que pueden ser sustituidas las variables:

"$(x)Fx$" se interpreta como "Todas las instancias de sustitución de 'F...' son verdaderas"
"$(\exists x)Fx$" se interpreta como "Al menos una instancia de sustitución de 'F...' es verdadera".

La interpretación objetual ha sido defendida —entre otros— por Quine y Davidson; y la interpretación sustitucional, por Mates y Marcus —entre otros—. Ambas interpretaciones tienen una historia bastante larga; por ejemplo, las explicaciones que da Russell de los cuantificadores están a veces en favor de una y a veces en favor de la otra. Pienso, sin embargo, que sería justo decir que la interpretación objetual es considerada generalmente como estándar y la sustitucional como aspirante cuyas credenciales están pendientes de escrutinio. Hay, como esto sugiere, dos posibles puntos de vista sobre el status de los dos tipos de interpretación: que son rivales, y sólo uno de ellos puede ser "correcto"; o que los dos tienen sus usos. Yo, junto con, por ejemplo, Belnap y Dunn, 1968; Linsky,

1972; Kripke, 1976, me inclinaré por el segundo, que es más tolerante.

Pero esto *no* quiere decir que carezca de importancia qué interpretación se elija. Por el contrario, la elección puede tener consecuencias filosóficas importantes. No podré considerar todas las ramificaciones detalladamente; pero esbozaré una explicación del papel crucial que desempeña la interpretación objetual en el punto de vista ontológico de Quine. Ello resultará de interés tanto con vistas a ilustrar los temas metafísicos que tienden a enredarse con las cuestiones acerca de la interpretación de los lenguajes formales, como también al objeto de mostrar más adelante (cap. 10) cómo las ideas de Quine sobre cuantificación y ontología determinan su actitud ante la inteligibilidad de la lógica modal.

2 INTERLUDIO METAFÍSICO: SOBRE CUANTIFICACIÓN Y ONTOLOGÍA EN QUINE

La ontología puede caracterizarse como aquella parte de la metafísica que se ocupa de la cuestión de qué tipos de cosas hay. Los puntos de vista de Quine sobre la ontología pueden considerarse como el producto de dos ideas clave, las ideas expresadas en dos de sus aforismos más conocidos: "*ser es ser el valor de una variable*" y "*ninguna entidad sin identidad*" (véase la fig. 2). El primer eslogan introduce *el criterio del compromiso ontológico* de Quine, que es un test acerca de qué tipos de cosas hay según una teoría; el segundo introduce sus *cánones de admisibilidad ontológica* —solamente pueden tolerarse aquellas entidades para las que se pueda proporcionar criterios adecuados de identidad. Me centraré en la primera de estas ideas, en el criterio del compromiso ontológico, pues es por esto por lo que resulta principalmente importante el apoyo que presta Quine a la interpretación objetual.

No obstante, serán de utilidad algunos breves comentarios sobre la segunda idea. Los criterios de identidad proporcionan condiciones para que cosas de un determinado tipo sean idénticas, así como: los conjuntos son idénticos si poseen los mismos miembros, o como: dos objetos físicos son idénticos si ocupan la misma posición espacio-temporal. Obsérvese que el requisito de que sólo sean permitidos aquellos tipos de entidades para los que pueda proporcionarse criterios de identidad es más bien fuerte (por ejemplo, nosotros tenemos plena confianza de que hay gente, pero existe un problema notorio acerca de proporcionar criterios de identidad personal). Quine sostiene que las nociones intensionales (de significado) son desesperadamente oscuras; en consecuencia, las condiciones de identidad establecidas en términos intensionales no son adecuadas según los cánones de Quine; y, por consiguiente, los tipos de entidades supuestas que sólo pueden ser individuadas apelando al significado —propiedades o proposiciones, por ejemplo— no son admisibles según dichos cánones.

El criterio del compromiso ontológico

¿Cuál es, entonces, el criterio del compromiso ontológico de Quine y cómo se conexiona con la interpretación objetual de los cuantificadores? El criterio se presenta de varias formas no siempre equivalentes:

> *las entidades de una clase dada son asumidas por una teoría si y sólo si alguna de ellas ha de ser incluida entre los valores de las variables para que los enunciados afirmados en la teoría sean verdaderos.* (1953a, pág. 103.)
>
> decir que una cuantificación existencial dada presupone objetos de un tipo dado es decir simplemente que la oración abierta que sigue al cuantificador es verdadera para algunos objetos de ese tipo y no lo es para ningún objeto que no sea de ese tipo. (1953a, pág. 131.)

La idea es, más o menos, que uno sabe lo que una teoría dice que hay poniéndola en notación del cálculo de predicados y preguntando qué tipos de cosas se requieren como valores de sus variables si los teoremas que comienzan por "$(\exists x)$..." han de ser verdaderos. (Así, una teoría en la que "$(\exists x)(x$ es primo y $x > 1.000.000)$" es un teorema, está comprometida con la existencia de números primos mayores que un millón y, *a fortiori*, con la existencia de números primos y con la existencia de números.) Es obvio que el criterio se aplica sólo a las teorías *interpretadas*. Es importante también que el criterio ha de aplicarse solamente cuando la teoría se encuentra expresada en *notación primitiva;* si, por ejemplo, la cuantificación de los números es sólo una abreviación de la cuantificación de las clases, entonces la teoría está comprometida con clases, pero no con números. El criterio de Quine es un test de lo que una teoría dice que hay, no de lo que hay. *Lo que hay* es lo que una teoría *verdadera* dice que hay. La negativa a admitir entidades intensionales actúa como una especie de filtro preliminar; las teorías que afirman que hay entidades intensionales no son, en opinión de Quine, realmente inteligibles y, por consiguiente, *a fortiori*, no son verdaderas.

La explicación que da Quine de su criterio deja bastante que desear. Como observa Cartwright en 1954, algunas formulaciones —por ejemplo, la primera anteriormente citada— emplean locuciones tales como "tiene que", "debe", "requiere" que son empero

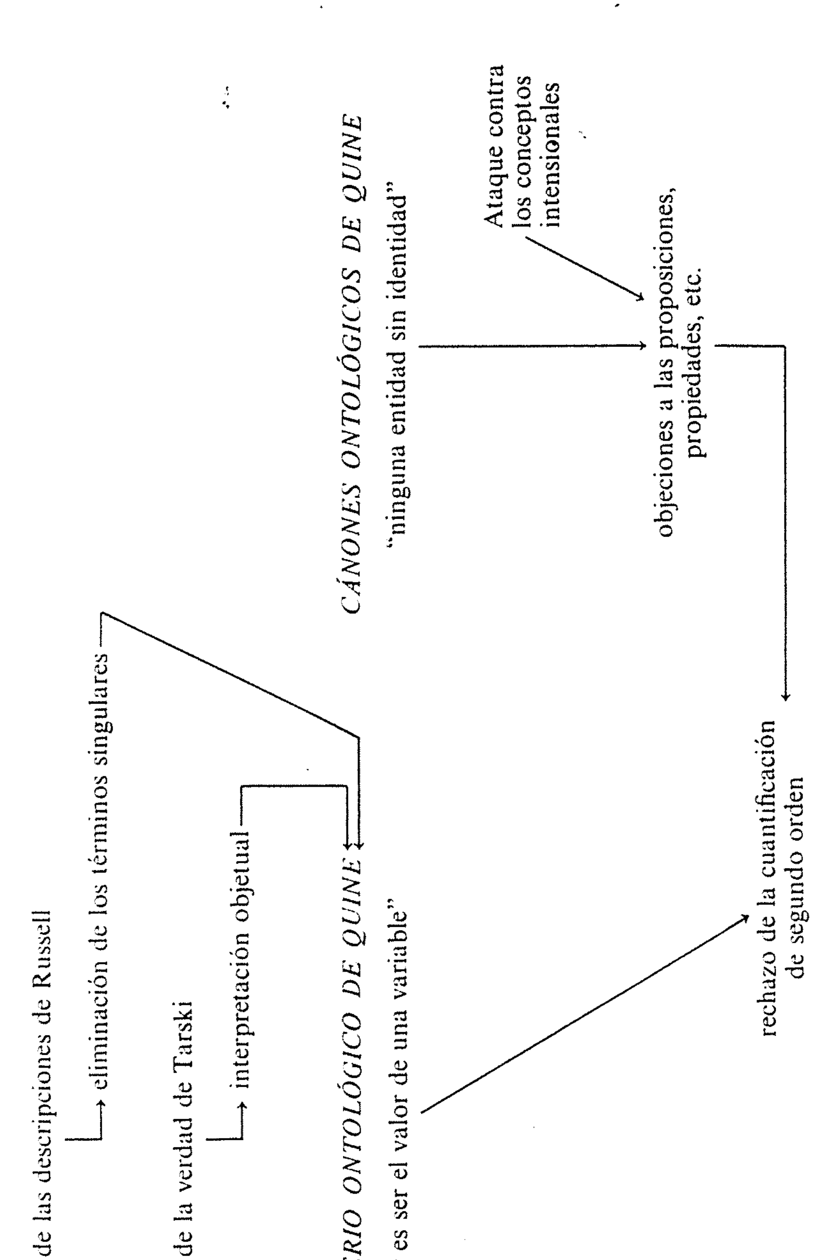


Fig. 2

giros idiomáticos intensionales y Quine insiste oficialmente en que deben evitarse. Quine ha mantenido explícitamente que su criterio es extensional (1953a, págs. 15, 131)[2]; y otras formulaciones —por ejemplo, la segunda anteriormente citada— se presentan en términos puramente extensionales. La cuestión es si las formulaciones extensionales son adecuadas. Scheffler y Chomsky (1958) argumentan más bien persuasivamente que no lo son. El problema está en cómo se ha de entender en la formulación extensional la condición "la oración abierta que sigue al cuantificador es verdadera para algunos objetos de cierto tipo, y no lo es para ningún objeto que no sea de ese tipo". Si se lee "∃ objetos del tipo k, tal que la oración abierta es verdadera para ellos y no lo es para ninguno que no sea de ese tipo", se sigue que es imposible afirmar que una teoría dice que hay objetos del tipo k, sin que *se* afirme asimismo que hay objetos del tipo k, puesto que "∃ objetos del tipo k..." conlleva en sí misma el compromiso ontológico. Pero, si se lee "Si la oración abierta es verdadera para cualquier objeto, es verdadera para algún objeto del tipo k, y no lo es para ningún objeto que no sea de ese tipo", se sigue que cualquier teoría que esté comprometida con algo que no exista, está, por tanto, comprometida con todo lo que no existe, puesto que si el antecedente es falso, el condicional es verdadero. Pero si el criterio no puede establecerse adecuadamente de forma extensional, falla por los propios cánones de Quine. Quien no haya compartido los escrúpulos de Quine, por supuesto, podría encontrar aceptable el criterio a pesar de su carácter intensional. Pero hay otras cuestiones que obligan a uno a preguntarse por las razones que ofrece Quine en favor del criterio.

Una razón importante de por qué Quine coloca el compromiso ontológico en las variables es que él piensa que la *eliminabilidad de los términos singulares* muestra que el compromiso ontológico de una teoría no puede residir en sus nombres. Esto da lugar a dos cuestiones: ¿tiene razón Quine al proclamar que los términos singulares son eliminables?; y, si son eliminables, ¿tiene razón para pensar que el compromiso ontológico ha de ser llevado a cabo por las variables ligadas? Consideraré estas cuestiones sucesivamente.

La propuesta de Quine para la eliminación de los términos singulares consta de dos momentos: primero, los términos singulares son reemplazados por descripciones definidas y, luego, se eliminan las descripciones definidas en favor de los cuantificadores y las variables.

[2] Él permite que, cuando el criterio se aplica a una teoría que no está en forma cuantificacional, se introduzca un elemento intensional en forma de apelación a una traducción correcta de dicha teoría en el cálculo de predicados (1953a, pág. 131). Esta concesión, junto con la tesis de la indeterminación de la traducción de Quine (1960a, cap. 2), conduce a la tesis de la *relatividad ontológica* (1968). Pero de lo que ahora me ocupo es de si el criterio en sí puede presentarse en forma extensional.

(i) En el caso de algunos nombres propios, al menos, se puede proporcionar una descripción definida que denota la misma cosa: por ejemplo, "el profesor de Platón" por "Sócrates". Para evitar las dificultades que pueden a veces aparecer al encontrar un predicado ordinario indudablemente verdadero del individuo denotado por un nombre, propone Quine la construcción de predicados artificiales y define "a" (por "Sócrates") como "$(\iota x)Ax$" (por "el x que socratea"). Se podría pensar, sugiere Quine, que el nuevo predicado, "A", significa "$= a$" (por consiguiente, "... socratea" significa "... es idéntico a Sócrates"). Pero este comentario extraoficial no se ha de considerar como *definición* de los nuevos predicados, pues el verdadero motivo de introducirlos es desembarazarse de los nombres totalmente; ésta es meramente una explicación intuitiva de los predicados, que debe considerarse como primitiva.

(ii) El segundo momento consiste en usar la teoría de las descripciones de Russell para eliminar las descripciones definidas que reemplazan a los términos singulares. Esto elimina las descripciones definidas en favor de los cuantificadores, las variables y la identidad (para más detalles, véase el cap. 5, § 3), así:

> El x que F es G = df. Hay exactamente un F y todo lo que es F es G

i.e., en símbolos:

$$G((\iota x)Fx) = \text{df. } (\exists x)((y)(Fy \equiv x = y) \,\&\, Gx)$$

De este modo, las oraciones que contienen nombres (como "Sócrates tomó veneno") pueden ser reemplazadas por oraciones que contienen descripciones ("El x que socratea tomó veneno") y, luego, por oraciones que contienen solamente cuantificadores y variables ("Hay exactamente un x que socratea y todo lo que socratea tomó veneno").

Quine sacó la conclusión (1953a, pág. 13) de que, puesto que "todo lo que decimos con la ayuda de los nombres se puede decir en un lenguaje que evite completamente los nombres", no pueden ser los nombres quienes realicen el compromiso ontológico, sino que han de ser las variables cuantificadas.

La tesis de la eliminabilidad de los términos singulares ha recibido críticas (véase, por ejemplo, Strawson, 1961). Pero la duda real concierne no tanto a la viabilidad formal de la propuesta de Quine como a su significación filosófica. El hecho de que Quine pueda suministrar una descripción definida apropiada para reemplazar un nombre solamente mediante el uso de predicados que, aunque inanalizables oficialmente, se explican extraoficialmente con la ayuda de los nombres ("A" significa "$= a$") nos garantiza poco que la eliminabilidad de los términos muestre en realidad que son ontológicamente irrelevantes.

Igualmente inseguro es el descubrimiento de que no sólo son eliminables los términos singulares, sino también los cuantificadores y las variables. En la *lógica combinatoria* debida a Schönfinkel y Curry —y, bastante irónicamente, explicada por el propio Quine en 1960b— las variables son suplantadas por operadores de predicados llamados "combinadores". El operador de predicados "*Des*", para "desrelativización", cambia un predicado n-posicional en un predicado $(n - 1)$-posicional. Si "F" es un predicado uniposicional, por ejemplo, "... es un perro", "*Des F*" es un predicado de *0* posiciones —una oración cerrada— "Algo es un perro"; si "... R ---" es un predicado biposicional, por ejemplo, "... muerde ---", "*Des R*" es un predicado uniposicional, "... muerde algo", y "*Des Des R*" es un predicado de 0 posiciones, "Algo muerde algo".

"*Inv*", para "*inversión*", invierte el orden de las posiciones de un predicado biposicional; así, "((*Inv R*)..., ---)" significa "--- R ...". "*Ref*", para "*reflexivo*", convierte un predicado biposicional en un predicado reflexivo uniposicional: así, "*Ref R*..." significa "... está en R consigo mismo". El procedimiento se generaliza a predicados poliádicos y a predicaciones compuestas; y el resultado es una traducción sin cuantificadores de las fórmulas de la teoría de la cuantificación, en la que la inversión se ocupa de la permutación del orden de las variables, la reflexión de la repetición de las variables y la desrelativización de la cuantificación.

Quine concede que su criterio no es aplicable directamente a la lógica combinatoria, pero observa que se puede aplicar indirectamente mediante la traducción de fórmulas combinatorias a fórmulas cuantificadas. Pero esto sólo sirve para oscurecer el problema, que consiste en que si la eliminabilidad de los términos singulares fuese una buena razón para negar que ellos realizan el compromiso ontológico, entonces la eliminabilidad de los cuantificadores sería presumiblemente tan buena razón para negar*les* significación ontológica.

Esto pone en claro, pienso, la importancia verdaderamente considerable que tiene la insistencia de Quine sobre la *interpretación objetual de los cuantificadores* para su criterio ontológico. Aunque la misma teoría podría expresarse usando términos singulares y cuantificadores, u operadores combinatorios en lugar de cuantificadores, su forma cuantificacional, piensa Quine, revela sus compromisos ontológicos más transparentemente, ya que *una oración de la forma* "$(\exists x)$..." *dice que hay algo que...*

> Insistir en la corrección del criterio... es, desde luego, simplemente decir que no se hace ninguna distinción entre el "hay" de "hay universales", "hay unicornios", "hay hipopótamos" y el "hay" de "$(\exists x)$", "hay entidades x tales que...".

Y una desviación de la interpretación objetual amenazaría el criterio:

> El no estar conforme con el criterio... es simplemente decir o que la notación cuantificacional familiar se utiliza en un sentido nuevo (en cuyo caso no tenemos que preocuparnos), o bien que el familiar "hay" de "hay universales", y otros, se utiliza en un sentido nuevo (en cuyo caso tampoco tenemos que preocuparnos). (1953a, pág. 105.)

En la interpretación objetual "$(\exists x)Fx$" significa que hay un objeto x en el dominio, D, que es F. Ahora bien, si se toma D como "el universo" —todo lo que hay— que, según parece, es lo que supone Quine, entonces, en efecto, "$(\exists x)Fx$" significa que hay un objeto (real existente) que es F; cfr. el uso que hace Quine de "entidad" en el párrafo recientemente citado.

Si "$(\exists x)Fx$" *significa* "Hay un objeto (existente) que es F", entonces, si es un teorema de una teoría que $(\exists x)Fx$, entonces esa teoría dice que hay un objeto que es F; y si uno dice que hay unos F, está comprometido con el ser de los F. La lectura objetual del cuantificador coloca ciertamente el compromiso ontológico en las variables ligadas de una teoría. Quizás se me permita refundir el aforismo de Quine de esta manera: que se diga que ser es ser el valor de una variable ligada por un cuantificador objetual; ¡esta forma es menos famosa, pero más verdadera! Sin embargo, hay que observar que el criterio de Quine empieza ya a tener un aire singularmente oblicuo: como si se descubriese que una teoría que dice que hay tales-y-tales se compromete ontológicamente con tales-y-tales traduciéndola primero a la notación del cálculo de predicados y apelando, después, a la interpretación objetual de los cuantificadores para mostrar que sus teoremas existenciales dicen que hay tales-y-tales.

Hay que hacer un esfuerzo serio al decidir qué aserciones existenciales *ostensibles* de una teoría necesitan permanecer en notación primitiva y cuáles son eliminables mediante la paráfrasis adecuada. Un ejemplo sería la propuesta de Morton White (1956) de reducir "Hay una posibilidad de que Jaime venga", que parece aseverar la existencia de posibilidades, a "Que Jaime venga no es ciertamente falso", que no parece aseverarla. (Quedan todavía muchas preguntas filosóficas espinosas acerca del significado de la paráfrasis aquí, pero de momento no me detendré en ellas; véase Alston, 1958, para una crítica de la idea de que la paráfrasis puede eliminar el compromiso ontológico, y cfr. Lewis, 1973, cap. 4, donde se supone que la paráfrasis *preserva* el compromiso ontológico.)

Cuantificación sustitucional y ontología

La interpretación sustitucional no da respuesta negativa a las preguntas ontológicas; sino que, más bien, las *pospone*. Según la explicación sustitucional, "$(\exists x)Fx$" significa "Alguna instancia de sustitución de 'F...' es verdadera"; las cuestiones de existencia dependen ahora de las condiciones de verdad de las instancias de sustitución. Si, por ejemplo, "Fa" es verdadera solamente si "a" es un término singular que denota un objeto (existente), entonces tendrá que haber un objeto que es F si "$(\exists x)Fx$" ha de ser verdadera; pero no es *inevitable* que las condiciones de verdad de las instancias apropiadas de sustitución tengan un compromiso ontológico. Un ejemplo: la presencia en el cálculo de predicados de teoremas tales como:

$$(\exists x)(Fx \vee -Fx)$$

que, según la interpretación objetual, dice que hay al menos un objeto que es o F o no F, i.e., que hay al menos un objeto, resulta embarazosa cuando se piensa que no debería ser asunto de la *lógica* el que *algo* exista. ¿Evitaría la interpretación sustitucional el compromiso existencial de los teoremas embarazosos? Según esta interpretación, el teorema significa que:

> Al menos una instancia de sustitución de "F ... $\vee$ $-$ F ..." es verdadera.

Si sólo se admiten como *sustituyentes* los nombres que denotan un objeto, entonces, de acuerdo con esta interpretación, también el cálculo de predicados precisará al menos de un objeto. Pero si se permiten como *sustituyentes* términos no denotativos como "*Pegaso*", entonces puede evitarse el compromiso ontológico. Esto ilustra el modo en que la cuantificación sustitucional *aplaza* las cuestiones ontológicas trasladándolas de los cuantificadores a los nombres; ¡Quine tiende a sugerir que esta ordenación de cuestiones existenciales es una evasión deplorable de la responsabilidad metafísica! Pero indicaré más adelante que ello puede tener ventajas.

Dado que en la lectura sustitucional "$(\exists x)Fx$" significa que al menos una instancia de sustitución de "F..." es verdadera, si este cuantificador metalingüístico se interpreta objetualmente, resultará comprometido con la existencia de las expresiones apropiadas, i.e., las instancias de sustitución. Pero no sucederá esto, si se interpreta también sustitucionalmente.

3 La elección de interpretación

¿Es solamente una la interpretación "correcta" de los cuantificadores? ¿O se puede elegir entre las dos según las intenciones que uno tenga? Y, si es así, ¿cuáles son los flancos fuertes o débiles de cada interpretación?

Cuantificadores sustitucionales y verdad

Cualquier interpretación de los cuantificadores que se adopte establecerá una diferencia en la definición de verdad de las oraciones cuantificadas. Seré breve ahora, ya que se tratará esto más detenidamente en el cap. 7, §§ 5 y 6. Si los cuantificadores se interpretan sustitucionalmente, entonces la verdad de las fórmulas cuantificadas puede definirse directamente en términos de la verdad de fórmulas atómicas (como " '$(\exists x)Fx$' es verdadera sii alguna instancia de sustitución de 'F...' es verdadera"); si los cuantificadores se interpretan objetualmente, la definición de verdad será menos directa. Ahora bien, Tarski propone, como una "condición de adecuación material" para las definiciones de verdad, que cualquier definición aceptable debe tener como consecuencia todas las instancias del "esquema (T)": "S es verdadera sii p", donde "S" nombra la oración "p"; y Wallace, 1971, teme que si se adopta la interpretación sustitucional, la definición de verdad no reunirá este requisito. Kripke, empero, ha argumentado (1976) que la condición de Tarski no es violada; y que, en cualquier caso, puede haber reservas acerca del requisito mismo. Por consiguiente, daré por sentado que la interpretación sustitucional no es objetable por *este* lado.

¿Demasiado pocos nombres?

Esto plantea la cuestión de si las interpretaciones sustitucional y objetual son siempre igualmente adecuadas. La respuesta es claramente que no. Es, desde luego, un requisito para una interpretación aceptable que los teoremas de la teoría que se está interpretando resulten verdaderos. La interpretación sustitucional hará obviamente verdaderas a las fbfs existencialmente cuantificadas, sólo si se dispone de *sustituyentes* apropiados. Por ejemplo, "$(\exists x)(Fx \vee -Fx)$" es un teorema del cálculo de predicados de primer orden; sin embargo, en una formulación con cuantificadores, pero sin términos singulares, la interpretación sustitucional no podría hacer, por falta de instancias de sustitución apropiadas, que una tal fbf resultase verdadera (de modo que la eliminación de términos singulares imposibilitará la interpretación sustitucional). Otra situación en la cual se imposibilitaría la interpretación sustitucional, sería un sistema formal en el que "$(\exists x)Fx$" fuese un teorema, a no ser que toda instancia de sustitución "$-Fa$" fuese demostrable; porque la interpretación sustitucional no podría hacer verdadero el teorema cuantificado sin hacer verdadera, al menos, a una de sus instancias (esta posibilidad es tratada por Quine, 1968; y cfr. Weston, 1974).

Tiempo

Sus defensores, sin embargo, argumentan que algunas veces la interpretación sustitucional ofrece ventajas sobre la objetual. Por ejemplo, Marcus sugiere que una lectura sustitucional evitará dificultades acerca del tiempo. Strawson se ha preguntado (1952, páginas 150-1) cómo representar "Había al menos una mujer entre los supervivientes": "¿Hay (¿había?) al menos un x tal que x es (¿era?) una mujer y...?" Veo poco probable que este problema se resuelva mediante una lectura sustitucional: "Al menos una instancia de sustitución de '...' es (¿era?) verdadera." Es cierto e interesante que el tiempo tiene importancia para la (in)validez de los argumentos informales, y que el aparato lógico usual es insensible al tiempo (cfr. cap. 9, § 3).

Modalidad

Marcus sugiere también, a mi entender con más razón, que la cuantificación sustitucional podría resolver algunos problemas de la interpretación de la lógica modal de predicados. De la oración presumiblemente verdadera:

> Necesariamente (la estrella de la tarde = la estrella de la tarde)

por un razonamiento presumiblemente válido del cálculo de predicados, se sigue:

> $(\exists x)$ Necesariamente (x = la estrella de la tarde)

que, según la interpretación objetual, es

> Hay al menos un objeto, x, tal que necesariamente x es idéntico a la estrella de la tarde.

Pero esto es difícil de entender; de hecho, Quine basa en ello el argumento de que toda la empresa de la lógica modal de predica-

dos está equivocada; pues, ¿cuál es ese objeto que es necesariamente idéntico a la estrella de la tarde? De ningún modo la estrella de la tarde; ese objeto es la estrella de la mañana, y la estrella de la mañana no es necesariamente, sino sólo contingente, idéntica a la estrella de la tarde. Sin embargo, las cuestiones difíciles de Quine se evitan mediante la lectura sustitucional de la oración problemática:

> Al menos una instancia de sustitución de "Necesariamente (... = la estrella de la tarde)" es verdadera

cuando parece (ya que "Necesariamente (la estrella de la tarde = la estrella de la tarde)" es verdadera) verdadera sin problemas.

La interpretación sustitucional también parece ofrecer ciertas ventajas cuando uno se dirige a la cuantificación de segundo orden.

Cuantificación de segundo orden

Si, según la interpretación objetual, "$(\exists x)$..." significa que hay un objeto tal que..., y "(x)..." significa que para todos los objetos, ..., entonces se debería esperar que los *sustituyentes* apropiados para las variables ligadas fuesen expresiones cuyo papel es denotar objetos, es decir, términos singulares. Quine, efectivamente, define a veces un término singular como una expresión que puede ocupar la posición de una variable ligada. Según la interpretación sustitucional, sin embargo, la cuantificación se relaciona directamente no con los objetos, sino con los *sustituyentes*; y, por tanto, no hay ninguna necesidad particular de insistir en que sólo las expresiones de la categoría de términos singulares puedan ser ligadas por cuantificadores. La clase de sustitución podría ser la clase de términos singulares, pero igualmente podría ser la clase de predicados o la clase de oraciones ... etc.

Por consiguiente, según la interpretación objetual, del mismo modo que una cuantificación de primer orden como:

> 1. $(\exists x)Fx$

dice que hay un objeto (individuo) que es F, una cuantificación de segundo orden como:

> 2. $(\exists F)Fx$

dice que *hay un objeto (propiedad) que x tiene*, y:

> 3. $(\exists p)(p \rightarrow -p)$

dice que hay *un objeto (proposición) que implica materialmente su propia negación*. La restricción natural de la clase sustitucional nos obliga a interpretar sintácticamente a "F" y "p" ligadas como términos singulares; se observará que esto violenta la lectura, ya que si "F" es un término singular, "Fx" debe leerse como "x TIENE F" para formar una oración gramatical. Y la interpretación en términos de objetos nos obliga a considerar que la cuantificación de segundo orden nos compromete con los objetos (abstractos). Quine no presta atención a tales supuestos objetos en cuanto propiedades o proposiciones, pero, viéndose obligado por la interpretación objetual a permitir que la cuantificación de segundo orden le comprometa con la existencia de los mismos, prefiere no admitir la cuantificación de segundo orden de ninguna manera, sino ceñirse a teorías de primer orden.

Sin embargo, con la interpretación sustitucional no queda uno limitado a términos singulares como *sustituyentes*; y mientras que en el caso de la cuantificación de primer orden los términos singulares serían adecuados, en el caso de la cuantificación de segundo orden, como en *2* y en *3*, los predicados o fórmulas, respectivamente, serían los *sustituyentes* apropiados. *2* significará que alguna instancia de sustitución de "...x" es verdadera, *3* que alguna instancia de sustitución de "... → —..." es verdadera. Ahora ya no es necesario violentar la lectura de las variables ligadas para hacerlas similares a los nombres, y naturalmente no hay compromiso alguno con los objetos intensionales, puesto que no hay ningún compromiso con los objetos (véase la fig. 3).

Estas consideraciones, pienso, tienen cierta relevancia para las cuestiones metafísicas. Los nominalistas admiten la existencia solamente de los particulares, mientras que los realistas, o platónicos, permiten también la realidad de los universales. C. S. Peirce pensaba que la influencia de los nominalistas en la historia de la filosofía había sido tan profunda desde Ockham que sólo el nominalismo y el "platonismo nominalista", el punto de vista de que los universales son una especie de particulares abstractos, parecían ser las únicas alternativas; Peirce, rechazando las dos, abogó por un realismo que, en lugar de asimilarlas, permitía diferenciar entre particulares y universales. Ahora bien, si solamente los nombres son sustituibles por las variables ligadas, se ve uno obligado, por decirlo así, a elegir entre un tipo de nominalismo (se permite solamente la cuantificación de primer orden, con variables sustituibles por nombres de particulares: la posición de Quine) y un tipo de platonismo nominalista (se permite la cuantificación de segundo orden, con variables reemplazables por nombres de objetos abstractos, propiedades o proposiciones: la posición de Church). Sin embargo, permitiendo *sustituyentes* de otras categorías sintácticas tenemos una tercera opción; que, sin duda, podría resultar atrayente desde un punto de vista metafísico.

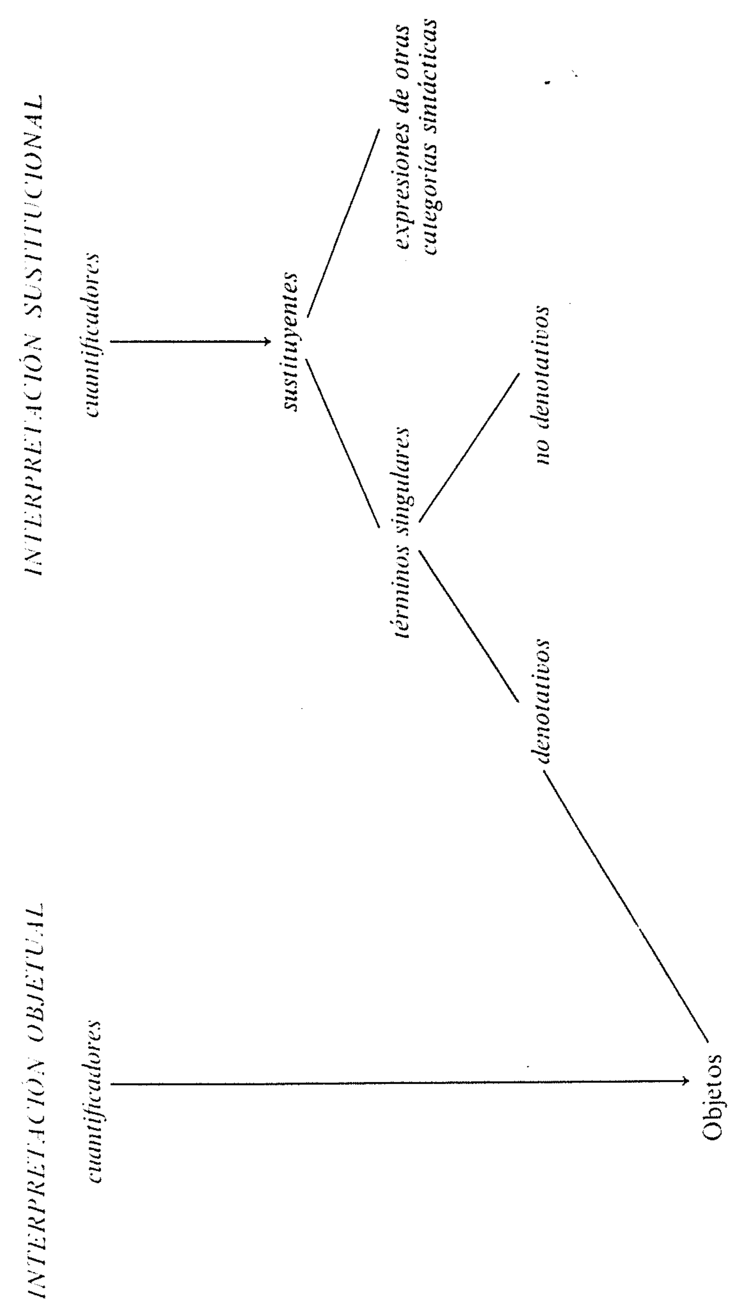


Fig. 3 Cuantificadores y ontología

Hay otros temas importantes que dependen también de la provisión de una interpretación aceptable de los cuantificadores de segundo orden. Uno de ellos es la viabilidad de la teoría de la redundancia de la verdad; y las consideraciones de la presente sección serán altamente pertinentes cuando, en el cap. 7, § 7, me ocupe detalladamente de dicho tema.

5
Términos singulares

1 Términos singulares y su interpretación

Algunas formulaciones del cálculo de predicados emplean términos singulares ("*a*", "*b*", ...etc.) además de variables. Si los cuantificadores se interpretan sustitucionalmente, es esencial, por supuesto, la presencia de términos singulares en el lenguaje para proporcionar las instancias de sustitución apropiadas. ¿Qué corresponde a los términos singulares de la lógica formal en los argumentos informales? Los términos singulares normalmente se consideran como los análogos a los nombres propios de los lenguajes naturales. (Cuando las variables fluctúan sobre números, los numerales corresponderían a los términos singulares.) La interpretación formal de los términos singulares asigna a cada uno un individuo específico del dominio que recorren las variables; y, en los lenguajes naturales, los nombres propios están pensados para funcionar de un modo parecido, al representar cada uno a una persona concreta (o lugar o cualquier otra cosa).

Así pues, mientras en el caso de los cuantificadores la polémica principal circunda la cuestión de las interpretaciones formales más convenientes, en el caso de los términos singulares los problemas se centran, más bien, en la comprensión de sus "análogos" en el lenguaje natural. La interpretación formal de los términos singulares en lenguajes extensionales sencillos no es polémica; sin embargo, se han usado opiniones rivales sobre cómo entender los nombres propios en los lenguajes naturales para apoyar propuestas alternativas sobre la interpretación formal de los términos singulares en los cálculos menos sencillos, por ejemplo, el cálculo modal. Entre las cuestiones discutidas acerca de cómo funcionan exactamente los nombres propios, están, por ejemplo: ¿exactamente qué expresiones son nombres propios *bona fide?* Por ejemplo, ¿se tienen en cuenta los "nombres" de entidades míticas o ficticias ("Pegaso", "Mr. Pickwick")? Y si es así, ¿qué hay que decir del valor de verdad de las oraciones que contienen tales nombres "que no denotan"? Especialmente, ¿cómo explicar la verdad intuitiva de los existenciales negativos como "Pegaso nunca existió"? Si el papel de los nombres es pura y simplemente denotar un individuo, ¿cómo puede ser informativo un enunciado de identidad verdadero (como "Cicerón = Tulio")? y ¿cómo puede la sustitución de un nombre por otro que denota al mismo individuo cambiar a veces el valor de verdad de una oración (como "No se necesita ningún conocimiento de la historia romana para saber que Tulio = Tulio" que es probablemente verdadera, y "No se necesita tener ningún conocimiento de la historia romana para saber que Tulio = Cicerón", que probablemente es falsa)? Una cuestión central, y la única en que me concentraré, es si los nombres propios tienen significado ("sentido", "connotación") así como denotación, y si es así, qué significado tienen. Aquellos que piensan que los nombres propios tienen significado por lo general asocian su significado más o menos estrechamente al significado de las descripciones definidas o designativas. La primera opinión, que los nombres tienen denotación, pero no significado, contribuye a crear una marcada diferenciación entre los nombres propios ("Sócrates", "Bismarck", etc.) y las descripciones ("El maestro de Platón", "El canciller responsable de la unificación de Alemania"), mientras que la segunda opinión ve a los nombres más como descripciones. Esto nos lleva a una segunda pregunta clave: ¿cómo funcionan las descripciones definidas?

2 Nombres

En primer lugar trataré la primera cuestión. Las alternativas están esbozadas en la tabla 1.

Pueden ir bien algunos comentarios preliminares sobre la distinción entre los nombres propios y las descripciones definidas. Normalmente la distinción no es difícil de hacer, pero hay algunas expresiones que son complicadas de clasificar: por ejemplo, aunque la Estrella de la Mañana no es una estrella, sino un planeta (Venus), se la sigue *llamando* "la Estrella de la Mañana"; así, "la Estrella de la Mañana" parece haber llegado a ser quizás más nombre y menos descripción de lo que era originalmente; las letras mayúsculas pueden ser significativas de su status intermedio. (La Universidad de Warwick tampoco está en Warwick.) Además, no todos los nombres son de la misma clase; los lógicos son propensos a tomar como ejemplos nombres de personas o con menos frecuencia de lugares, pero hay también títulos de libros, nombres de marcas, nombres comerciales... (y obsérvese cómo los nombres comerciales tienen la costumbre de convertirse en nombres comunes, e incluso verbos, por ejemplo, "fagorízate"). También es digno de atención el que los

lógicos sean muy aficionados a los nombres de gente famosa ("Aristóteles", "Napoleón", etc.); y saludable recordar que sin duda hay muchos Aristóteles y Napoleones, y sólo un fondo de información compartida nos hace a todos pensar en el mismo. Prestar atención a la variedad de clases de nombre propio puede inducir a adoptar cierta cautela sobre la suposición de que haya algo tal como *el* modo en que funcionan los nombres propios.

TABLA 1. *¿Tienen los nombres propios significado así como denotación?*

	Sí		*No*
Frege Russell (Quine)	Los nombres propios tienen el sentido de una descripción definida co-designativa conocida para el hablante.	Mill Ziff	Los nombres propios tienen denotación, pero no connotación y no son parte del lenguaje.
Wittgenstein Searle	Los nombres propios tienen el sentido de un subconjunto indeterminado de algún conjunto abierto de descripciones co-designativas.	Kripke	Los nombres propios son "designadores rígidos"; consideración causal del correcto uso de los nombres.
Burge (Davidson)	Los nombres propios como predicados.		

Los nombres propios como puramente denotativos

Hay una opinión para la que los nombres propios, en contraste con las descripciones definidas, son, por así decirlo, simples etiquetas: sirven para representar una persona, lugar o cosa. No pienso que las personas que adoptan esta opinión tengan la intención, o necesiten negar que en el caso de los nombres de personas hay convenciones acerca de qué nombres se dan a los machos y cuáles a las hembras, por ejemplo, acerca de que un niño tome el nombre de su padre, etc.; ni que los nombres tienen un "sentido" derivado de su etimología, como "'Pedro' significa 'piedra'". El nombre que tiene una persona puede transmitir, en virtud de convenciones como las mencionadas, alguna información sobre ella; lo que se niega, más bien, es que el nombre la *describa*.

Según Mill, 1843, los nombres propios tienen denotación, pero no connotación, es decir, no tienen significado. Ziff, en 1960, suscribe algo parecido a esta opinión; los nombres propios no tienen significado y, efectivamente, en cierto sentido ni siquiera son parte del lenguaje. Otro autor que niega que los nombres propios tengan significado es Kripke, quien, en 1972, esbozó una exposición de los aspectos semántico y pragmático de los nombres propios. Los nombres propios son "designadores rígidos", que es como decir que tienen la misma referencia en todos los mundos posibles. Por ejemplo, el nombre "Aristóteles" designa al mismo individuo en todos los mundos posibles, mientras que la descripción definida "El hombre más importante que estudió con Platón", aunque designe a Aristóteles en el mundo real, puede designar a otros individuos en otros mundos posibles; pues es posible que Aristóteles no estudiara con Platón. La idea es ésta: un nombre propio simplemente designa a un individuo específico, y puesto que no describe a ese individuo, lo designa, no en cuanto que es el individuo que..., sino simplemente *qua* ese individuo específico; y, por tanto, no importa lo diferente que fuera el individuo que el nombre designa del modo en que es en realidad, el nombre propio aún designaría a ese individuo —y es esto lo que Kripke significa al decir que designa al mismo individuo en todos los mundos posibles—. (Aparentemente Kripke identificaría a un individuo en cuanto a su origen; en el caso de personas por su fecha de nacimiento y linaje.)

Kripke no niega que la *referencia* de un nombre se pueda fijar mediante una descripción definida, ni que se podría introducir un nombre para denotar al referente, en el mundo real, de alguna descripción definida, fijar la referencia de "Fido", por ejemplo, como el primer perro que fue al mar. Lo que niega es que la descripción definida proporcione el *sentido* del nombre; probablemente Fido podría no haber sido el primer perro que fuera al mar, pero, en un mundo posible en el que no esté, mientras "el primer perro que fue al mar" designa a un perro diferente, "Fido" aún designa *Fido*, es decir, el perro que, en el mundo real era el primer perro que fue al mar.

La consideración semántica se complementa con una explicación causal de la pragmática del nombrar; el objeto es explicar cómo un hablante puede usar un nombre correctamente aun cuando sea bastante incapaz de dar una descripción que se aplique únicamente al individuo nombrado —alguien que sólo sabe de Feynman que es un físico—. Según Kripke, un hablante usa un nombre correctamente si hay una adecuada cadena de comunicación que conecta *su* uso del nombre con el individuo designado por un nombre en un "bautismo" inicial. Nace un niño, sus padres le ponen nombre, otras personas le conocen, se convierte en físico, escribe ensayos que otros leen y escriben sobre ellos..., y así sucesivamente; entonces un hablante usa el nombre "Feynman" correctamente, para referirse a Feynman, si su uso de este nombre está conectado causalmente en forma adecuada a la cadena de comunicación que vuelve al mismo Feynman. Por supuesto, no es literalmente necesario que haya un bautismo inicial y la cadena de comunicación puede ser, en efecto, muy larga, como el uso de "Julio César", por ejemplo. Kripke es consciente de que, "conectado causalmente *en forma adecuada*", tiene mucha necesidad de ampliación; ya que no propone ninguna explicación posterior, todavía no hay garantía de que la consideración causal no resulte trivial o falsa.

La conexión entre los cabos semántico y pragmático de la consideración de Kripke está, probablemente, en que sus criterios para el uso correcto de un nombre no recurren al conocimiento o a las creencias del hablante acerca del individuo designado, sino que sólo exigen que su uso del nombre esté causal y adecuadamente conectado con ese individuo; esto está en consonancia, con la insistencia, en la consideración semántica, en que un nombre sólo designa y no describe. Sin embargo, si la referencia de un nombre se puede fijar, como autoriza Kripke mediante una descripción definida, se podría abrir una brecha entre las consideraciones semántica y pragmática; pues si fijo la referencia de un nombre propio mediante una descripción definida que, de hecho, aunque no lo sepa, no designa nada (por ejemplo, si decido llamar "Smith" al hombre que robó mi maleta, cuando en realidad mi maleta no ha sido robada, sino transportada por un mozo), no puede haber una cadena causal adecuada con el portador del nombre, ya que no hay portador.

De la tesis de que los nombres propios son designadores rígidos se sigue que todos los enunciados de identidad de la forma "$a = b$" y donde "a" y "b" son nombres, son necesarios. Si "a" y "b" son nombres y "$a = b$" es verdadero, de manera que "a" y "b" designan al mismo individuo en el mundo real, entonces, ya que los dos nombres, al ser designadores rígidos, designan al mismo individuo en todos los mundos posibles, "$a = b$" es verdadero en todos los mundos posibles, que es como decir que es necesariamente verdadero.

Nombres equiparados a descripciones

Ahora bien, precisamente fue debido a un problema acerca de los enunciados de identidad por lo que Frege introdujo (1892a) su distinción entre sentido (*Sinn*) y referencia (*Bedeutung*) y argumentó que los nombres propios tienen sentido así como referencia. Pregunta Frege: ¿cómo puede

(i) $a = b$

diferir en "valor cognitivo", es decir, ser más informativo que

(ii) $a = a$

si a es b? Su respuesta es que, mientras si a es b, la referencia de "a" es la misma que la referencia de "b" (representan al mismo objeto), el sentido de "a" es diferente del sentido de "b", y esta diferencia explica el mayor grado de información de (i) sobre (ii)[1].

Frege explica el grado de información de los enunciados verdaderos de la forma "$a = b$" como originado a partir de la diferencia de sentido de los nombres "a" y "b". ¿Cómo explicaría esto Kripke, que no admite que los nombres *tienen* sentidos y para quien todas las identidades son necesarias? Su explicación es que, aunque los enunciados de la forma "$a = b$" son necesarios, no todos los enunciados necesarios pueden ser conocidos *a priori*; o sea, puede ser un *descubrimiento*, por muy necesariamente que a sea b. Por ejemplo, se dio el nombre "Hesperus" a cierto cuerpo celeste que se veía al atardecer, y el nombre de "Phosphorus" a un cuerpo celeste que se veía por la mañana; los dos son designadores rígidos y resultan ser designadores del mismo cuerpo celeste (el planeta Venus); pero los astrónomos tuvieron que *descubrir*, y no lo sabían *a priori*, que designan al mismo cuerpo celeste. (Kripke comenta que no hay nada especialmente digno de mención en saber de una proposición que es necesaria *si* es verdadera, y no saber todavía si *es* verdadera; la conjetura de Goldbach sería un ejemplo.)

---

[1] Aunque la distinción se introduce en principio, específicamente para nombres, se extiende para aplicarla a los predicados, y luego, según el principio de que el sentido (referencia) de una expresión compuesta depende del sentido (referencia) de sus partes, a las oraciones. Así:

expresión	*sentido*	*referencia*
nombre propio	significado del nombre	objeto
predicado	significado de la expresión predicativa	concepto
oración	proposición	valor de verdad

La referencia de una oración debe ser su valor de verdad, dice Frege, ya que si algún componente de una oración se reemplaza por otro con diferente sentido, pero igual referencia (como la "La Estrella de la Mañana es un Planeta" / "La Estrella de la Tarde es un Planeta"), es el valor de verdad lo que permanece inalterable. Siempre fuertemente antipsicologista, Frege insiste en que el sentido o significado, de una expresión ha de distinguirse de la idea que puede suceder que esté asociada con esa expresión; por tanto cuando dice que el sentido de una oración es el "pensamiento" (*Gedanke*) que expresa, quiere decir "proposición" más que "idea". En contextos "oblícuos", añade Frege (es decir contextos intensionales, por ejemplo habla indirecta) que las oraciones tienen no sólo su referencia acostumbrada sino también una referencia "indirecta", siendo la referencia directa el sentido acostumbrado, es decir, la proposición expresada. Por tanto en "Tom dijo que Mary vendría", la referencia de "Mary vendría" no es su valor de verdad sino la proposición de que Mary vendría.

Sin embargo, Frege considera que los nombres propios tienen sentido así como referencia. Por "nombre propio" entiende Frege *tanto* los nombres corrientes *como* las descripciones definidas (dice que un nombre es cualquier expresión que se refiera a un objeto definido, aunque de hecho también concibe la posibilidad de nombres, como "Odysseus", que no denotan un objeto real). Y equipara el sentido de un nombre corriente con el sentido de una descripción definida que se refiere al mismo objeto. ¿Con qué descripción definida codesignativa? Al parecer (1892a, pág. 58n, y cfr. 1918, página 517) una en la que piensa el hablante o que sucede que conoce. Frege se da cuenta de que esto tiene como consecuencia que personas diferentes pueden atribuir diferentes significados a un nombre, que dependerán de lo que sepan acerca de la persona nombrada; él comenta que tales variaciones de sentido, aunque deberían evitarse en un lenguaje perfecto, son tolerables con tal que permanezca la misma referencia. En vista del hecho de que una de las objeciones que con frecuencia hace en contra de la identificación del sentido de una expresión con la idea asociada, es que esto significaría que el sentido variaría de una persona a otra, esta tolerancia es sorprendente.

Russell, como Frege, identifica el significado de los nombres propios corrientes con el significado de algunas descripciones definidas relevantes (aunque, como se verá más tarde, difiere de Frege tanto en su opinión sobre el significado como en su opinión acerca de cómo podrían explicarse, a su vez, las descripciones definidas). Y, de nuevo como Frege, vio Russell que se seguía de esto que los nombres tienen diferente significado para hablantes diferentes.

Russell distinguió también, sin embargo, una categoría especial de *nombres lógicamente propios:* son expresiones cuyo papel es puramente denotar un objeto simple, y cuyo significado es el objeto denotado (así pues, en el caso de los nombres lógicamente propios Russell *equipara* significado y referencia). En la versión russelliana del atomismo lógico, los "objetos simples" son "objetos de conocimiento directo", así pues, los nombres lógicamente propios denotan objetos de conocimiento directo; según Russell no conocemos directamente los objetos corrientes, las personas, etc., sino únicamente los *sense-data*. Así pues, las únicas expresiones que admite que son nombres lógicamente propios son "esto", "eso" y (durante el periodo en que creía en un ego de introspección directa) "yo". Ningún nombre propio corriente es nombre lógicamente propio, pues ningún nombre propio corriente denota objetos de conocimiento directo. A veces Russell usa "conocimiento directo" de un modo más común, distinguiendo entre personas y lugares que realmente se han conocido o visitado y aquellos de los que sólo se ha oído hablar, y considera a los nombres o lugares de los que se tiene conocimiento directo en este sentido como nombres lógicamente propios. Pero está claro que éste es un uso relajado, y que la teoría estricta según la cual ningún nombre corriente es nombre lógicamente propio es la única que debe tomarse en serio.

Identificar el significado de un nombre propio con una descripción codesignativa conocida para el hablante tiene, como observaron Frege y Russell, la incómoda consecuencia de que el significado del nombre varía según los hablantes. Esta dificultad podría evitarse identificando el significado del nombre con el conjunto de todas las descripciones verdaderas para el portador. Pero esto tiene la infeliz consecuencia de que todo enunciado verdadero de la forma "*a* es (era) la persona que...", donde "*a*" es un nombre propio, es analítico, y todo enunciado falso de esa forma, contradictorio, pues "*a*", en este planteamiento, *significa* justamente "la persona que..." para todas las descripciones verdaderas para el portador. Este problema, a su vez, podría desviarse suavizando la conexión entre el significado del nombre y el conjunto de descripciones de su portador. Una idea de este tipo se encuentra en las *Philosophical Investigations* (Wittgenstein, 1953), donde se sugiere que un nombre no tiene un significado fijo e inequívoco, sino que está vagamente asociado con un conjunto de descripciones: por "Moisés" se puede significar el hombre que hizo la mayor parte o mucho de lo que la Biblia cuenta de Moisés, pero no está fijado cuánto, o qué parte de la historia tiene que ser falsa antes de que se diga que no hubo una persona tal como Moisés (§ 79). También Searle en 1969 propone algo parecido: mientras ninguno de los hechos que se consideran como establecidos sobre *a* requiere necesariamente ser verdadero, sin embargo, la disyunción de ellos debe serlo (página 138). No es analítico que Moisés fue encontrado entre los juncos, ni que condujo a los israelitas fuera de Egipto, ni..., etc., pero es analítico, de acuerdo con Searle, que o Moisés fue encontrado entre los juncos o..., etc. Como Wittgenstein, insiste Searle que es indeterminado cuántos de los disyuntos pueden ser falsos antes de que sea verdadero decir que *a* nunca existió.

Hasta aquí, pues, hay estas alternativas:

> sean $d_1 \dots d_n$ todas las descripciones supuestamente verdaderas para *a*

entonces, o:

> el significado de "*a*" es algún (algunos) miembros(s) del conjunto

o:

> el significado de "*a*" es la conjunción $d_1 \& d_2 \& \dots d_n$ de todos los miembros del conjunto

o:

> el significado de "*a*" es algún subconjunto del conjunto de descripciones, habiendo indeterminación acerca de cuáles y cuántas de las d_i incluye.

Estas propuestas identifican, o más bien asocian vagamente, el sentido de un nombre propio con el de las descripciones definidas relacionadas. Otra propuesta, de un espíritu parecido, es la presentada por Burge en 1973 (y apoyada por Davidson); según esta postura, en lugar de ser considerado el nombre como la abreviatura de una descripción definida se le toma como un predicado. Burge señala que los nombres propios, de hecho, raramente representan a un objeto único, que toman terminaciones de plural ("hay tres Jacks en la clase") y artículos definido e indefinido ("el Jack que escribió esto", "hay una Mary en la clase, pero ninguna Jane"). Burge está interesado en los usos literales de los nombres y no en los metafóricos, en "Callaghan es un James" más que en "Callaghan no es ningún Churchill". "Jack es alto", según la postura de Burge, es mejor considerarlo como una especie de oración abierta, con "Jack" como predicado gobernado por un demostrativo "este Jack es alto" (como "ese libro es verde") la referencia del cual está fijada por el contexto. Considerado, pues, como un predicado, "Jack" es, según Burge, verdadero de un objeto en el caso de que el objeto *sea un Jack*, en el caso de que al objeto se le haya dado ese nombre de una forma adecuada. La postura de Burge tiene algunas afinidades con una sugerencia que se encuentra en Kneale, 1962a, de que el significado del nombre "*a*" es la "persona llamada '*a*' ". Kripke objeta a la propuesta de Kneale que es viciosamente circular. Burge, sin embargo, señala que su interpretación de los nombres propios como predicados podría complementarse mediante una teoría del nombrar, una teoría que pondría las condiciones en las que un objeto es un Jack, es decir, las condiciones en las que es verdadero decir que "se le ha dado el nombre 'Jack' de una forma adecuada". No hay ninguna razón, por supuesto, por la que la clase de relación causal en la que insiste Kripke no pueda tener un papel a representar en este nivel.

Hay una tendencia a pensar en los nombres propios como, por así decirlo, medios por los cuales el lenguaje consigue su más directo asidero en el mundo; y quizás por esta razón hay una fuerte motivación para dar una imagen clara y ordenada de la forma en que funciona el nombrar. En las teorías que he esbozado, de la conexión entre nombres y los individuos nombrados emergen dos clases de imágenes: la puramente denotativa, o imagen "arpón", y la descriptiva, o imagen "red" (derivo de Fitzpatrick la útil metáfora, pero he cambiado su "flecha" por "arpón" para dar un lugar al papel de la cadena causal del nombrar de la interpretación kripkeana). Ya he sugerido que los nombres propios corrientes son muy diversos en los lenguajes naturales, y que funcionan contra un fondo de información compartida o parcialmente compartida, o información errónea. Confirmación para mi sospecha de que no puede haber un modo único en el que funcionen los nombres, puede encontrarse en la manera en que las dos imágenes, oficialmente representadas como rivales, parecen en realidad complementarse una con la otra: la imagen arpón explica cómo podemos arreglárnoslas para hablar sobre alguien aun cuando seamos ignorantes o estemos erróneamente informados acerca de él —se ha escapado, por así decirlo, por nuestra red; la imagen red explica cómo podemos hablar, sin confusión, sobre una de las muchas o varias personas del mismo nombre.

Los detalles de la imagen red dependerán, evidentemente, de qué interpretación se proporcione de las descripciones que, según esta opinión, están asociadas con un nombre. Paso a esta cuestión en lo que sigue.

3 Descripciones

Aunque Frege y Russell equipararon el significado de los nombres propios (corrientes) con el de las descripciones definidas que les corresponden, dan explicaciones bastante diferentes acerca del modo en que funcionan esas descripciones.

Según la *teoría de las descripciones* de Russell (Russell, 1905), las descripciones definidas, tales como "la montaña más alta del mundo", son "símbolos incompletos", es decir, que son eliminables contextualmente. Russell no da una definición explícita que permita reemplazar una descripción definida por un equivalente dondequiera que aparezca, sino una *definición contextual*, que permite reemplazar oraciones que contienen descripciones definidas por oraciones equivalentes que no contienen descripciones definidas:

$$E!(\iota x)Fx = \text{df.} \quad (\exists x)(y)(Fy \equiv x = y)$$

es decir, "el F existe" significa "hay exactamente un F", y

$$G((\iota x)Fx) = \text{df.} \quad (\exists x)((y)(Fy \equiv x = y) \,\&\, Gx)$$

es decir, "el F es G" significa "hay exactamente un F y cualquier cosa que sea F es G". La última tendrá, por tanto, dos "negaciones":

$$-(\exists x)((y)(Fy \equiv x = y) \,\&\, Gx)$$

es decir, "No es el caso que (hay exactamente un F y cualquier cosa que sea F es G)", y:

$$(\exists x)((y)(Fy \equiv x = y) \,\&\, -Gx)$$

es decir, "Hay exactamente un F y cualquier cosa que sea F no es G". De éstas, sólo la primera es la contradictoria de "F es G"; la última es su contraria. (De hecho, en general, ha de indicarse qué *alcance* tiene una descripción definida cuando está en una oración compuesta.)

Russell hace notar que la forma gramatical de oraciones como "La montaña más alta del mundo está en el Himalaya" es engañosa en cuanto a su forma lógica; lo que quiere decir es que, mientras que la oración en castellano contiene una expresión, "la montaña más alta del mundo", que parece como si tuviera el papel de designar un objeto, su representación formal no contiene ningún término singular, sino solamente variables ligadas, predicados e identidad. Y esto permite a Russell ocuparse del problema de las descripciones definidas tales como "el actual Rey de Francia", que no son verdaderas de ninguna cosa. El problema, como lo ve Russell, es éste: si "El actual Rey de Francia es calvo" es lógicamente, como lo es gramaticalmente, una oración de sujeto-predicado, entonces su término sujeto "el actual Rey de Francia" debe ser un nombre lógicamente propio, cuyo significado es el objeto que denota; pero ya que no hay ningún actual Rey de Francia, o "el actual Rey de Francia" denota un objeto irreal, o no denota nada, por tanto, y, en consecuencia, la oración completa es carente de sentido. Reacio a aceptar una de estas conclusiones, Russell resuelve el problema negando que "El actual Rey de Francia es calvo" es, lógicamente, de la forma sujeto-predicado; lógicamente es una oración existencial. Finalmente, pues, Russell niega que todos los nombres propios corrientes (o descripciones definidas) estén representados adecuadamente por los términos singulares de su lenguaje formal; este privilegio está restringido a los nombres lógicamente propios.

Russell consideró su teoría de las descripciones como ontológicamente liberadora; pues le liberó de la necesidad de admitir un dominio de entidades irreales como denotación de los nombres que aparentemente no denotan. (Véase su crítica de Meinong (1905), que admitió objetos no existentes, y cfr. § 4 más adelante.) Después de desarrollar la teoría, de hecho, Russell redujo con bastante severidad sus compromisos ontológicos. Antes, en rebelión contra el monismo de Bradley, admitió una ontología pluralista exuberante, creyendo, como dijo, en todo lo que Bradley no creía; pero, más tarde, influenciado por la defensa de la navaja de Ockham hecha por Whitehead, y equipado con la teoría de las descripciones que le liberó de la necesidad de admitir un objeto como denotación para asegurar la significatividad de todo nombre aparente, descartó como "ficciones" no sólo los objetos de Meinong, sino clases, propiedades e incluso objetos físicos. (Véase Quine, 1966b, para detalles sobre el desarrollo de las opiniones ontológicas de Russell.)

La propuesta de Quine (discutida en el cap. 4, § 2) de eliminar los términos singulares en favor de las descripciones definidas codesignativas está con toda claridad en el espíritu del planteamiento que hace Russell de los nombres propios; Quine no admite una categoría especial de nombres lógicamente propios y tampoco aceptaría los supuestos epistemológicos que se ocultan tras la doctrina del conocimiento de Russell, pero yo creo que simpatizaría con la visión russelliana de la teoría de las descripciones como un instrumento de restricción ontológica.

Para Frege, que no tiene ninguna categoría especial de nombres lógicamente propios, el significado de los cuales se identifica con su denotación, el problema de los nombres que no denotan parece algo diferente. Frege puede admitir que oraciones que contienen nombres o descripciones que no denotan, no obstante, tienen sentido perfectamente (expresan una proposición *bona fide*). Sin embargo, al dar su principio de que la referencia de una expresión compuesta depende de la referencia de sus componentes, está obligado a admitir que una oración como "El actual Rey de Francia es calvo", cuyo sujeto no tiene referencia, carece en sí misma de referencia, es decir, no tiene valor de verdad. Así pues, mientras según el análisis de Russell "El actual Rey de Francia es calvo" entraña que hay un actual Rey de Francia (pues que lo hay, es parte de lo que dice), según la interpretación de Frege "El actual Rey de Francia es calvo" *presupone* que hay un actual Rey de Francia, es decir, no es ni verdadera ni falsa a no ser que "El actual Rey de Francia existe" sea verdadera. Un adecuado tratamiento formal de la presuposición, es bastante evidente que requerirá una lógica no bivalente, una lógica en la que permitan los huecos de valor de verdad, es decir, en la que algunas fbfs no sean ni verdaderas ni falsas. Sin embargo, Frege no presenta una lógica de este tipo (pero véase Smiley, 1960, y van Fraassen, 1966, para las reconstrucciones formales de la idea de Frege); pues piensa en los términos singulares que no denotan como en una imperfección de los lenguajes naturales, que no debería permitirse en un lenguaje lógicamente perfecto, y, por tanto, recomienda que, en lógica formal, todos los términos singualres tengan asegurada una denotación, si es necesario aportando artificialmente un objeto —él sugiere el número 0— como su referente. (La elección del número 0 puede ser un poco desafortunada, ya que probablemente tendría la consecuencia de que "El mayor número primo es menor que 1", por ejemplo, sería verdadera.) En cualquier caso, mientras en la teoría de Russell las descripciones definidas y los nombres propios corrientes no son

genuinamente términos singulares, sino que son eliminados contextualmente, Frege considera los nombres corrientes y las descripciones como términos singulares *bona fide*, cada una con un único referente, y con términos "pícaros", como "el mayor número primo", que se refieren a 0. (Una teoría formal fregeana se encuentra en Carnap, 1942.)

En su influyente crítica de la teoría de Russell, Strawson, en 1950 (y véase Nelson, 1946, que anticipa alguno de los puntos de Strawson), emplea una noción de presuposición que tiene reminiscencias del análisis de Frege. Aunque hay diferencias que conviene destacar primero y derivan en gran parte de la insistencia de Strawson en la distinción entre oraciones y enunciados. Según Strawson, mientras son las expresiones lingüísticas las que tienen significado, son los usos de expresiones lingüísticas los que refieren, y en particular, los usos de oraciones —enunciados— los que son verdaderos o falsos. Así pues, su diagnóstico del problema de las descripciones que no denotan es así: aunque la expresión "el actual Rey de Francia es calvo" es bastante significativa, un uso de esa expresión no logra referir, y, en consecuencia, un uso de una oración que contiene esa expresión no logra formar un enunciado verdadero o falso. Strawson es ambiguo sobre si su diagnóstico es que un uso de la oración "El actual Rey de Francia es calvo" no logra formar un enunciado, o que un uso de este tipo forma un enunciado, pero un enunciado que no es ni verdadero ni falso. (La ambigüedad se expone claramente en Nerlich, 1965.) Hay también una ambigüedad en la tesis de Strawson de que una elocución de "El actual Rey de Francia es calvo" no entraña, como piensa Russell, sino que presupone que hay un actual Rey de Francia: algunos pasajes sugieren que es el hablante el que presupone que hay un actual Rey de Francia, otros, que la presuposición no es este tipo de relación epistemológica, sino una relación lógica que se mantiene entre el enunciado de que el actual Rey de Francia es calvo y el enunciado de que hay un actual Rey de Francia. En posteriores artículos (1954, 1964) Strawson se contenta con la segunda tesis: presuponer es una relación lógica entre enunciados, tal que S_1 presupone S_2 en el caso de que S_1 no es ni verdadero ni falso a no ser que S_2 sea verdadero. Ya que las relaciones lógicas sólo se mantienen, según Strawson, entre enunciados, esto resuelve también la primera ambigüedad observada antes —una elocución de "El actual Rey de Francia es calvo" debe admitirse que constituye un enunciado, pero un enunciado que no es ni verdadero ni falso. Obsérvese primero que, si no es por la insistencia en que existe una relación entre enunciados, la interpretación de Strawson de la presuposición es como la de Frege; y, segundo, que si una elocución de "El actual Rey de Francia es calvo", después de todo, constituye un enunciado, la crítica de Strawson de que el error de Russell estaba en no lograr distinguir entre oraciones y enunciados no se puede sostener. (Sobre este segundo punto véase la réplica de Russell (1959) a Strawson.)

Dudo de si la cuestión de si "El actual Rey de Francia es calvo" debería considerarse falsa o sin valor de verdad, podría o incluso debería resolverse recurriendo a "lo que diríamos corrientemente". La cuestión más bien se centra en si se está preparado para tolerar alguna artificialidad (bien en el caso de la teoría de las descripciones de Russell, en la traducción de los lenguajes naturales al formalismo, o, en el caso de la anterior teoría de Frege, en la elección del referente para las, de otro modo, expresiones que no denotan) para conservar la bivalencia, ya que la teoría de la "presuposición" fregeana defendida por Strawson exigiría una lógica fundamental no bivalente. Y si, por supuesto, se pensó que había otras razones para dudar de la bivalencia, esto tendría relevancia. (Los comentarios sobre estrategias rivales en formalización del cap. 9, § 1, son pertinentes para esta particular elección.)

Strawson se cuida bien de decir que son los usos de las expresiones los que refieren. Pero otra vez hay algo de ambigüedad sobre lo que entiende que son las condiciones para una referencia con éxito. Algunos pasajes insinúan una interpretación pragmática, según la cual es condición suficiente para que un uso de una expresión tenga éxito al referirse a un objeto que el hablante piense en un cierto objeto y su uso de la expresión concentre sobre el objeto la atención del que escucha —es decir, sin tener en cuenta si la expresión usada realmente denota ese objeto (cfr. Strawson, 1959, capítulo 1, y 1964). Pero, por lo general, Strawson prefiere una interpretación semántica, según la cual es necesario, para que un uso de una expresión tenga éxito al referirse a un objeto, que la expresión denote ese objeto.

Donellan, en 1966, concede lugar preeminente a la noción pragmática de referencia. Donellan distingue entre *usos atributivos* y *referenciales* de las descripciones definidas. (La misma descripción definida puede presentarse en cualquiera de los dos usos.) Una descripción definida se usa atributivamente si el hablante quiere aseverar algo acerca de cualquier persona o cosa que responda a la descripción y referencialmente si quiere, más bien, atraer la atención de su audiencia sobre alguna persona o cosa en particular y aseverar algo acerca de ella. Donellan pone como ejemplo el uso de la oración "El hombre que asesinó a Smith está loco"; atributivamente se usa para expresar que quienquiera que sea el que asesinó a Smith debe estar loco, referencialmente se usa para expresar que Jones (de quien el hablante y la audiencia saben que ha sido convicto de asesinato —quizás erróneamente—) está loco. Y se puede usar referencialmente, al modo de Donellan, una descripción definida, incluso si no es verdadera, de la persona o cosa a la que

se refiere y aun cuando el hablante y el que escucha sepan que no lo es; pues el criterio de uso referencial con éxito es simplemente que el hablante se las arregle para concentrar la atención de la audiencia sobre la persona o cosa en la que piensa. Sugiere Donellan que la interpretación de Strawson es aplicable sólo a los usos atributivos y no a los referenciales.

Es cierto —como ya sugerí antes— que la teoría de Strawson es, en última instancia, más semántica y menos pragmática de lo que su oficial insistencia sobre el uso de las expresiones podría habernos hecho esperar. Podría ser un provechoso artificio distinguir entre *referencia* y *denotación* o *designación* y usar la primera para la noción pragmática (lo que hacen los hablantes) y la última para la noción semántica (lo que hacen las expresiones); así pues, si se desea adoptar los estándares de referencia con éxito de Donellan, se puede decir que un hablante puede referirse a una persona o cosa mediante el uso de una expresión que no denota a esa persona o cosa. Una ventaja de esto es poner de manifiesto que no es necesario considerar la interpretación de Donellan del "uso referencial" de una descripción definida como rival de las teorías de Frege o Russell.

4 Nombres que no denotan: ficción

Aquí los temas son complejos y enmarañados, y no puedo esperar comentarlos todos. Algunos —las relaciones entre los términos singulares y las variables ligadas, y la posibilidad de eliminar los primeros en favor de las últimas— ya se han tocado (cap. 4, § 2). Otros recibirán más atención en capítulos siguientes —el papel de los términos singulares en los contextos modales y las consecuencias de las teorías rivales del nombrar para los problemas sobre la identidad de los individuos a través de los mundos posibles en el cap. 10, teorías de la presuposición en el cap. 11—. Abordaré ahora la cuestión de los nombres que no denotan.

En la discusión anterior surgió que hay dos tipos de discrepancias entre los nombres propios de los lenguajes naturales y los términos singulares de los lenguajes formales: mientras los términos singulares se asignan cada uno a *exactamente* un individuo del dominio, los nombres propios, a veces, tienen *más de un* portador y a veces *ninguno*. No es desconocido de los autores que simplemente se acabe con estas discrepancias, suponiendo "mediante argumentación" que los nombres propios corrientes denotan con toda seguridad un único individuo (véase McDowell, 1977); pero se eluden algunas cuestiones interesantes si se las descuida tan a la ligera. No discutiré aquí la primera discrepancia: que los nombres propios ("John Smith") tienen con frecuencia varios o muchos portadores, aunque es digno de observar que, de las teorías que he bosquejado, la de Burge es la única que toma esta posibilidad más en serio. Por el momento, me ceñiré a algunos comentarios sobre la otra discrepancia, el fenómeno de los nombres que no denotan, y, en relación con él, a algunos pensamientos acerca del discurso ficticio.

Los problemas suscitados por los nombres que no denotan pueden ponerse de relieve al considerarlos desde el punto de vista de la teoría de Russell. Tomemos el nombre de un personaje de ficción, por ejemplo, "Sherlock Holmes". Según Russell, el significado de un nombre *bona fide* ("lógicamente") propio tiene que equipararse con su denotación, por tanto, si "Sherlock Holmes" fuera un nombre *bona fide*, ya que no denota, sería carente de sentido y, por tanto, además, lo serían todas las oraciones sobre Sherlock Holmes, incluidas algunas tales como "Sherlock Holmes nunca existió", que, sin duda con alguna justificación, se toman como claramente verdaderas (el "problema de los existenciales negativos", cfr. Cartwright, 1960). Russell quiere evitar esta dificultad negando que "Sherlock Holmes" sea un nombre genuino: es una descripción definida disfrazada, y las oraciones sobre Sherlock Holmes son disfrazadamente existenciales, perfectamente significativas, y o claramente verdaderas o claramente falsas: "Sherlock Holmes nunca existió" es verdadera, mientras que otros enunciados sobre Sherlock Holmes, como "Sherlock Holmes era un detective" o "Sherlock Holmes era un policía", son falsos.

La interpretación de Russell proporciona una explicación de cómo nos es posible hablar significativamente sobre los no existentes, y decir con verdad que *son* no existentes, y al mismo tiempo una solución sencilla al problema de los valores de verdad de tales enunciados. Pero algunos se han dado cuenta de que la asignación de "falso" del modo en que se hace con "Sherlock Holmes era detective" y "Sherlock Holmes era un policía", es demasiado cruda y toma poco en cuenta la intuición de que el primero es "correcto" en un sentido en el que el último no lo es.

Sherlock Holmes es un personaje ficticio, y, según la obra de ficción en la que aparece, era un detective y no un policía. Se ha sugerido (cfr. Routley, 1963) que un lenguaje formal adecuado para representar el discurso sobre Sherlock Holmes podría exigir un dominio de entidades ficticias, de modo que el nombre "Sherlock Holmes" denotara, sólo que denotaría un objeto ficticio, no real. (Tales sistemas son conocidos como "lógicas libres", es decir, libres de compromiso existencial; véase Schock, 1968, y cfr. observaciones sobre elecciones alternativas de dominio en el cap. 4, § 1). Resulta interesante que este planteamiento se encuentra en el espíritu de la teoría de los objetos de Meinong, que tiene en cuenta el discurso significativo sobre los no existentes al admitir no sólo los

objetos reales, espacio-temporales, tales como los objetos físicos y las personas, y los objetos subsistentes, no espacio-temporales, tales como números y propiedades, sino también los objetos no existentes, no subsistentes e incluso imposibles, todos como genuinamente objetos (véase Meinong, 1904, y cfr. Parsons, 1974). Russell, en 1905, reconoce, como vimos, que esto presenta una alternativa a su propia teoría, pero de ella pensó que era ontológicamente objetable, quizás debido a sus afinidades con las extravagancias ontológicas a las que él mismo en un tiempo se había entregado (1903). Del modo curioso en el que las lógicas libres representan los términos no denotativos como terminos que denotan objetos no reales (algo así como la pretensión de que el tercer "valor" de algunas lógicas trivalentes represente una carencia de valor de verdad) podría pensarse, igualmente, que muestra una cierta ambivalencia ontológica.

Ahora bien, aunque la narración nos cuenta mucho sobre Holmes, hay también un buen número de enunciados sobre él, la verdad de los cuales *no* está fijada por la narración —si tenía una tía en Leamington Spa, por ejemplo—. Así pues, hay una motivación no sólo para ajustar el dominio para permitir las entidades ficticias, sino también para admitir que mientras algunos enunciados sobre Holmes son verdaderos y otros falsos, otros, además, no son ninguna de las dos cosas. Y esto significa que un lenguaje formal apropiado podría necesitar abandonar el principio de bivalencia, el principio de que todo enunciado es o verdadero o falso. En un tal lenguaje formal habría campo para la representación de la relación de presuposición de Frege, la cual, como señalé anteriormente, pide a gritos una lógica no bivalente.

Hay, por supuesto, una cuestión y es la de hasta qué punto toda habla sobre no existentes se debe ver sobre el modelo del habla acerca de las entidades ficticias; aunque Sherlock Holmes y el mayor número primo son iguales en lo que respecta al no existir, es discutible si son iguales en todos los aspectos lógicamente relevantes. Pero por el momento me ceñiré a la consideración de las entidades ficticias. De todas formas, está claro que hay una importante distinción entre, por una parte, el discurso *sobre* la ficción y, por otra, el discurso en la *ficción*. (No quiero sugerir, por supuesto, que *todo* discurso en o sobre la ficción es un discurso sobre entidades ficticias.) Lo que hago cuando hablo sobre Sherlock Holmes no es probablemente por completo análogo a lo que Conan Doyle hizo al escribir las historias de Holmes. En especial, mientras en el primer caso se pueden ver algunos motivos para la intuición de que hay un sentido en el cual lo que yo digo puede ser cierto o erróneo, en el último caso parece más apropiado decir que para Doyle la cuestión de obtener lo cierto o erróneo simplemente no surge. El tipo de respuesta considerada, pienso, parece más prometedora con respecto al discurso sobre la ficción que con respecto al discurso en la ficción.

Lo que es desacostumbrado acerca del discurso *en* la ficción, sospecho que no es en absoluto semántico, sino pragmático. Decir (o escribir) oraciones mientras se hace una narración difiere de decir oraciones mientras se informa de un suceso real; en el primer caso no se está como en el último aseverando, esto es, afirmando la verdad de las oraciones que se emiten (cfr. Plantinga, 1974, cap. 8, § 4, Woods, 1974; Searle, 1975; Haack, 1976b). Mientras que se podría sentir la necesidad de una lógica libre para el discurso sobre la ficción, se podría razonablemente esperar que las características distintivas del discurso en la ficción se trataran mediante una teoría de la pragmática. Pues, si mi presentimiento es acertado, la diferencia más significativa entre contar una historia y hacer un informe no radica, por así decirlo, en la diferencia entre la historia y el informe, sino en la diferencia entre el contar y el hacer.

A veces se da por supuesto que si las características distintivas de un tipo de discurso son pragmáticas, necesariamente esto le sitúa más allá del alcance de los métodos lógicos formales. La omnipresente importancia de los aspectos pragmáticos de todo discurso en los lenguajes naturales ha sido un tema recurrente con críticos como Schiller y Strawson, que consideran los métodos formales como seriamente inapropiados para las sutilezas del lenguaje natural. Así pues, quizás debería subrayar que, al alegar que las características distintivas del discurso en la ficción pueden ser más pragmáticas que semánticas, no doy por supuesto que esto necesariamente excluye la posibilidad de un tratamiento formal.

6

Oraciones, enunciados, proposiciones

1 Tres aproximaciones

Un tema recurrente en la filosofía de la lógica hace referencia a la pregunta acerca de la clase de ítem del que se ocupa la lógica, o quizás del que primariamente se ocupa. Las alternativas que normalmente se presentan son oraciones, enunciados y proposiciones, o, más raramente hoy en día, juicios o creencias. He planteado la pregunta de un modo deliberadamente impreciso, ya que parece estar involucrado más de un tema. Una vez más, como sucedió con la cuestión sobre los significados de las conectivas, cuantificadores, etc., el problema hace referencia a la relación entre argumentos formales e informales: ¿qué corresponde, en los argumentos informales, a las fbfs de los lenguajes formales? Puede ser útil distinguir tres aproximaciones a la pregunta:

(i) sintáctica: ¿qué es lo análogo, en los lenguajes naturales, de las "*p*", "*q*" de la lógica formal?

Hasta aquí, al hablar de "cálculo oracional", no quise plantear esta pregunta. Algunos prefieren hablar de "cálculo proposicional", "variables proposicionales"; y hasta aquí no he dicho nada para justificar mi preferencia por el anterior tratamiento.

(ii) semántica: ¿qué clase de ítem es capaz de verdad y falsedad?

Ya que los lenguajes formales pretenden representar aquellos argumentos informales que son válidos extrasistemáticamente, es decir, que son preservadores de verdad, esto se relacionará estrechamente con la primera pregunta.

(iii) pagmática[1]: ¿qué clases de ítem se podría asumir que son los "objetos" de creencia, conocimiento, suposición, etc?

[1] Llamo a ésta la aproximación pragmática porque la pragmática trata de las relaciones entre las expresiones y los usuarios de esas expresiones (la "sintaxis" y la "semantica" se explicaron en el cap. 2). Esta forma de separar los temas la tomo de Gochet, 1972.

(A veces, a "conocer", "suponer", etc., se les llama verbos de "actitud proposicional".) Ya que se puede conocer, creer o suponer algo verdadero o falso, esta tercera aproximación se relacionará bastante estrechamente con la segunda.

Sin embargo, por el momento, no discutiré (iii) (pero véase páginas 147-50 y cap. 12, § 3); primero comentaré muy brevemente (i) y, después, con muchos detalles (ii).

2 Oración, enunciado, proposición

Sin embargo, un preliminar necesario es especificar lo que entenderé por "oración", "enunciado" y "proposición"; pues una razón por la que la discusión de estos temas es con frecuencia confusa es la de que hay escasa uniformidad de tratamiento.

Por *oración* entenderé cualquier cadena de expresiones del lenguaje natural, gramaticalmente correcta y completa. Por ejemplo, "La nieve es blanca", "Cierra la puerta", "¿Está cerrada la puerta?" son oraciones; "sentado por en" y "rosa siendo" no lo son. Espero que esta consideración rápida y somera sea suficiente para expresar la idea en la que pienso; por supuesto, es imprecisa en cuanto que hay incertidumbre acerca de qué cadenas de expresiones se consideran como gramaticales. Necesitaré distinguir entre *tipos* de oración e *instancias* de oración. Una instancia de oración es un objeto físico, una serie de marcas en el papel o de ondas sonoras, que constituyen una oración escrita o hablada. Sin embargo, a veces se considera a dos o más instancias como en cierto sentido inscripciones o eluciones de la misma oración; "la misma oración" significa aquí "el mismo tipo de oración". Por ejemplo, las dos inscripciones:

Todos los filósofos están ligeramente locos
Todos los filósofos están ligeramente locos

son instancias del mismo tipo. Se podría pensar que un tipo de oración es o un modelo al que ejemplifican instancias similares, o una clase de instancias similares. La cuestión de qué criterio de identidad tomar para los tipos de oración, es discutida; unos exigirán similitud tipográfica o auditiva (probablemente también se necesitaría especificar las condiciones en las que una elocución lo sería del mismo tipo de oración que una inscripción), otros exigirían igualdad de significado. Me atendré al primer criterio, y admitiré la posibilidad de tipos ambiguos de oración. De nuevo necesito distinguir entre las oraciones las que son, por ejemplo, interrogativas o imperativas de las que son "declarativas". Las oraciones con el verbo principal en modo indicativo son declarativas, sólo que "declarativas" se considera que es bastante más amplio que "indicativas", para incluir, por ejemplo, las condicionales cuyo verbo principal está en subjuntivo. Intuitivamente, se podría decir que las oraciones declarativas son las seleccionables para la verdad y falsedad, mientras que las oraciones no declarativas no lo son; pero definir de este modo "declarativas", en el presente contexto, podría ser una petición de principio.

Por *enunciado* entenderé lo que se dice cuando se emite o se inscribe una oración. En su utilización no técnica, "enunciado" es ambiguo entre el suceso de la elocución o inscripción de una oración, y el contenido de lo que se inscribe o emite. Sólo el segundo sentido es relevante para los intereses del momento. Ahora surge la pregunta de si toda elocución o inscripción de una oración declarativa formará un enunciado. Strawson parece pensar que algunos usos de las oraciones declarativas —sus ejemplos incluyen elocuciones o inscripciones usadas durante la representación de una obra de teatro o al escribir una novela— *no* forman enunciados. También, como vimos en el capítulo anterior, parece insinuar que las elocuciones de oraciones cuyos términos para el sujeto no denotan nada no logran formar enunciados, aunque en otro momento sugiere, más bien, que tales elocuciones son enunciados, pero enunciados que no son ni verdaderos ni falsos. Estas preguntas obviamente serán importantes para el tema de los portadores de verdad. Ahora bien, ¿cuándo dos elocuciones o inscripciones forman el mismo enunciado? Normalmente se dice que esto es así justamente en el caso de que "digan la misma cosa sobre la misma cosa". Esta interpretación funciona bastante bien en casos sencillos. Por ejemplo, las elocuciones:

Tú tienes calor	(dicho por x a y)
Yo tengo calor	(dicho por y)
J'ai chaud	(dicho por y)

según estos estándars, formarían el mismo enunciado. Establecer el criterio preciso, sin embargo, parece ser difícil, por esto no siempre puede ser fácil especificar cuándo dos elocuciones lo son sobre la misma cosa, y podría ser todavía más arduo especificar cuándo dicen la misma cosa sobre su sujeto, ya que exigiría el recurso a la célebremente complicada noción de sinónimo.

Por *proposición* entenderé lo que es común a un conjunto de oraciones declarativas sinónimas. Según este sentido de "proposición", dos oraciones expresarán la misma proposición si tienen el mismo significado; así pues, de nuevo aquí, como con los enunciados, habrá que enfrentarse con el problema de la sinonimia. Otra interpretación, popular desde el advenimiento de la semántica, identifica una proposición con el conjunto de los mundos posibles en los que es verdadera, o con una función de los mundos posibles en valores de verdad. Sin embargo, no está claro que esto llegue a algo muy diferente de la interpretación que antes di, ya que distingue el mundo posible en el que p y q. (Si "Jack y Jill tienen en común uno de sus padres" expresa la misma proposición que "Jack y Jill son medio hermanos", entonces todos los mundos posibles en los que vale la primera son mundos posibles en los que vale la segunda, y si no es así, no.) Otra interpretación, que delimita una idea diferente, identifica la proposición con el contenido común de oraciones en diferentes modos verbales. Así pues:

> Tom, cierra la puerta.
> ¡Cierra la puerta, Tom!
> ¿Ha cerrado la puerta Tom?

tienen como contenido común la proposición "*el cerrar Tom la puerta*". Las proposiciones, en este sentido, son candidatos improbables a portadores de verdad, y, por esta razón, les prestaré escasa atención aquí. Sin embargo, tienen alguna relevancia para la interpretación de, por ejemplo, la lógica imperativa, sobre la que presentaré más adelante unos breves comentarios.

Es bastante fácil de comprobar que las oraciones, los enunciados y las proposiciones, tal como aquí se han caracterizado, son distintos, es decir, que se podría tener la misma oración/diferente enunciado/diferente proposición; el mismo enunciado/diferente oración/diferente proposición, la misma proposición/diferente oración/diferente enunciado (véase Cartwright, 1962).

La actitud frente a los enunciados o proposiciones puede quedar matizada por las propias opiniones metafísicas. Los nominalistas, a quienes no les gustan los objetos abstractos, o los extensionalistas, quienes sospechan que las nociones de significado padecen de una falta de claridad que las paraliza, es probable que estén mal dispuestos hacia los enunciados y las proposiciones y mejor dispuestos hacia las oraciones, mientras que los platónicos, que admiten los objetos abstractos, y los intensionalistas, cómodos con la teoría del significado, podrían admitir con ecuanimidad enunciados o proposiciones. (Compárese Quine, 1970, cap. 1, con Putnam, 1971, caps. 2, 3, 5, para contrastar actitudes.) Sin embargo, es necesario observar que aunque las instancias de oración son objetos físicos, los tipos de oración son abstractos; y que, mientras los criterios de identidad para enunciados y proposiciones exigen el recurso a la sinonimia, los criterios de identidad para los tipos de oración exigen el recurso a la noción de similitud, no totalmente libre de problemas. (Véase Goodman, 1970, para algunos de los problemas que rodean a los intentos de definir la similitud con precisión.)

3 "Letras de oración", "variables proposicionales" o ¿qué?

Cómo se entiendan las "p", "q"..., etc., de la lógica de oraciones dependerá obviamente de si se admite que las letras de oración se traten como genuinas variables para ser ligadas por cuantificadores, y si esto se hace, de cómo se interpreten esos cuantificadores.

Las habituales presentaciones de la lógica de oraciones no usan cuantificadores. A primera vista, sin embargo, parece razonable suponer que el cálculo de oraciones no cuantificado tiene una generalidad *implícita* que el cálculo de oraciones cuantificado simplemente hace *explícita*. Un teorema como "$p \rightarrow (p \vee q)$" normalmente se entiende que vale para todas las instancias de "p" y "q", tal y como, en las habituales presentaciones no cuantificadas del álgebra, "$a + b = b + a$" se entiende que vale cualquier cosa que puedan ser a y b. Así pues, las alternativas son o considerar a las formulaciones habituales no cuantificadas simplemente como una versión abreviada de la lógica de oraciones cuantificadas, o si no, encontrar algún otro medio de explicar la implícita generalidad del cálculo no extendido.

Quine, por razones a las que ya se hizo alusión en el cap. 4, § 3, prefiere la segunda alternativa. Propone que a "p", "q", etc., *no* se las trate como genuinas variables susceptibles de ser ligadas, sino que, en lugar de ello, se interpreten como "letras esquemáticas". Una fbf del cálculo de oraciones, tal como "$p \vee -p$", ha de entenderse "no como una oración, sino como un esquema o diagrama tal que todos los enunciados reales de la forma descrita son verdaderos" (1953a, pág. 109).

Sin embargo, si se *tratan* las letras de oración como genuinas variables, nos enfrentamos entonces con la cuestión de la interpretación de los cuantificadores. Si se adopta una interpretación objetual, en seguida nos enfrentamos con la cuestión de qué clase de objeto recorren los cuantificadores: los candidatos más habituales son las proposiciones, aunque Quine, en 1934, abogó por un dominio de oraciones. (Si sólo se trata del habitual cálculo de oraciones veritativo-funcional, incluso se podrían construir cuantificadores tales que recorrieran *valores de verdad*, es decir, que recorrieran los dos valores v y f; pues en lógica de oraciones veritativo-funcional solamente los valores de verdad de los componentes son relevantes para el valor de verdad del compuesto. Sin embargo, la adición de operadores de oración no veritativo-funcionales, como "necesariamente", quizás o "s cree que", excluiría esta alternativa.) En este caso es necesaria una adaptación a la lectura habitual: en "$(p)(p \vee -p)$", si se lee el cuantificador "para todas las proposiciones p", entonces "$p \vee -p$" debe interpretarse como término singular que denota una proposición compuesta ("la disyunción de una proposición con su propia negación"), y hay que aportar un predicado implícito ("es verdadera") para hacer la lectura gramatical. Por otra parte, si se adopta una interpretación sustitucional, "$(p)(p \vee -p)$" se leerá "Todas las instancias de sustitución de '... $\vee$ —...' son verdaderas", donde las instancias de sustitución adecuadas resultan de poner la misma oración en cada uno de los espacios en blanco. (No escapará a la consideración, estoy segura, que las "letras esquemáticas" de Quine se parecen mucho a las variables ligadas de la cuantificación sustitucional, que tienen oraciones como clase de sustitución.)

A este nivel, pues, parece haber varias opciones. Pero ¿qué hay de la cuestión de los portadores de la verdad?

4 Portadores de verdad

Si un argumento es válido, entonces si sus premisas son verdaderas, su conclusión debe ser verdadera también; así pues, probablemente las premisas y la conclusión necesitan ser la clase de ítem capaz de ser verdadero o falso. Por tanto, muchos autores han considerado esto como muy importante para decidir si aquello a lo que se llama "verdadero" o "falso" son oraciones, enunciados o proposiciones. El tema tiene muchas ramificaciones; por ejemplo, se ha sugerido que la confusión sobre los portadores de verdad es la base de las paradojas semánticas (Bar-Hillel, 1957; Kneale, 1971) que motiva las propuestas de lógica plurivalente (Lewy, 1946, Kneale y Kneale, 1962; Kripke, 1975, pág. 700n) que vicia la teoría de las descripciones de Russell (Strawson, 1950). Ya he comentado (cap. 5, § 3) la última de ellas; diré algo sobre la primera en el capítulo 8 y sobre la segunda en el cap. 11.

La discusión sobre los portadores de verdad se desarrolla normalmente en una línea como la que sigue: ya que probablemente la verdad es una propiedad, habría que ser capaz de identificar el tipo de ítem que la posee; normalmente se da por sentado, o que sólo uno de los candidatos puede ser el portador de verdad, o que uno es primario y los otros de algún modo derivados. El consiguiente debate acerca de qué ítems son los portadores de verdad o cuál es el primero de ellos, sin embargo, no ha sido, a mi entender, ni muy concluyente ni muy fructífero. En pocas palabras se verá lo que quiero decir.

Varios autores (Strawson, 1950; introducción a Pitcher, 1964; Putnam, 1971, por ejemplo) han afirmado que es impropio, o incluso carente de sentido, hablar de las oraciones como verdaderas o falsas. Los argumentos que se presentan para esta afirmación, sin embargo, parecen bastante poco concluyentes. Uno es que si las oraciones fueran verdaderas o falsas, algunas oraciones serían a veces verdaderas y a veces falsas; otro es que algunas oraciones,

las no declarativas, por ejemplo, no son capaces de verdad o falsedad, por tanto, no todas las oraciones pueden ser verdaderas o falsas. Pero, después de todo, de una puerta se puede decir con bastante corrección que es roja o verde, aunque puede tener un año un color y el siguiente otro; y a un cristal, por ejemplo, uno de color, se le puede adscribir adecuadamente predicados de color, a pesar del hecho de que algún cristal carece de color (véase Lemmon, 1966; R. J. y S. Haack, 1970).

Aunque estos argumentos desde luego no muestran que a las oraciones no se las puede propiamente llamar verdaderas o falsas, pueden sugerir una línea de pensamiento aparentemente más prometedora: que cualesquiera que sean los ítems que se elijan como portadores de verdad, deberían ser tales que (i) pueden transmitirse sin cambiar su valor de verdad, y (ii) todos los ítems del tipo relevante son o verdaderos o falsos. Por supuesto, la aceptabilidad de estos *desiderata* necesitará investigación. Pero incluso dejando aparte esa cuestión por el momento, resulta que los enunciados y las proposiciones apenas tienen más éxito que las oraciones con respecto a esto.

(i) El que un enunciado pueda cambiar de valor de verdad obviamente depende de cómo se entienda exactamente "decir la misma cosa sobre la misma cosa". Pero según una interpretación intuitiva, dos elocuciones que se refieren al mismo Jones, separadas por una distancia de medio minuto, de "Jones lleva un abrigo" dirían probablemente la misma cosa sobre la misma cosa. Sin embargo, una locución podría ser verdadera y la otra falsa, si Jones se pone o se quita el abrigo en el intervalo. Por supuesto, podríamos asegurarnos contra los cambios de valores de verdad de los enunciados, haciendo hasta tal punto estricto el criterio de identidad de enunciado que ninguna elocución no simultánea se considere que forma el mismo enunciado. Pero esto, en efecto, pondría en correlación de uno a uno a los enunciados con las instancias de oración, y entonces podría ser justificable el preguntarse para qué sirve presentar los enunciados como distintos de las oraciones.

Ya que el sentido de una oración puede permanecer estable durante un período considerable, la proposición expresada por una oración probablemente podría también cambiar su valor de verdad; por ejemplo, la proposición expresada por la oración "Luis XIV está muerto" en un tiempo fue falsa y ahora es verdadera. Algunos autores (Frege, 1918; Moore, 1953; Kneale, 1971, por ejemplo) han respondido a esta dificultad haciendo estricto el criterio de identidad proposicional para rechazar el cambio de valor de verdad; esto parece ser vulnerable a una objeción similar a la que antes se hizo a una maniobra parecida para prevenir los cambios de valores de verdad de los enunciados.

(ii) Ya que no se sabe si cada elocución de una oración declarativa se supone que hace un enunciado, no está claro, tampoco, si cada enunciado debe ser verdadero o falso. Strawson concede, sin embargo, que no es parte de la definición de "enunciado" que todo enunciado sea verdadero o falso (1952, pág. 69); y, como vimos, hay huellas en 1950, y una declaración explícita en 1964, de que elocuciones de oraciones de "referencia fallida" forman enunciados que no son ni verdaderos ni falsos. Por tanto, algunos enunciados carecerán de valor de verdad.

En algunos casos en los que una oración no es ni verdadera ni falsa, se podría argumentar convincentemente que no hay ninguna proposición con la que se corresponda, y hasta este punto a las proposiciones les va mejor que a las oraciones en cuanto a satisfacer (ii). Entre las oraciones que no son ni verdaderas ni falsas hay, así se dice con frecuencia, algunas que aunque gramaticalmente correctas son carentes de sentido ("La virtud es triangular", por ejemplo); al ser carentes de sentido, tales oraciones no expresan ninguna proposición. Las oraciones imperativas e interrogativas probablemente tampoco logran ser verdaderas o falsas, y una vez más podría afirmarse que tales oraciones no expresan proposiciones. Sin embargo, es dudoso si podría especificarse qué tipos de oración expresan proposiciones a no ser que nos limitemos a las oraciones declarativas (como antes en § 2); por tanto, este argumento no muestra que las proposiciones estén en mejores condiciones con respecto a (ii) que las oraciones. Y algunas oraciones declarativas (oraciones vagas y oraciones de futuros contingentes, por ejemplo) piensan algunos autores que no son ni verdaderas ni falsas, y no obstante ser bien significativas, expresan proposiciones que, por tanto, no son ni verdaderas ni falsas.

No trato de sugerir, por supuesto, que a las oraciones les va mejor que a los enunciados o proposiciones con respecto a (i) y (ii). Ya he mencionado antes varios tipos de oración que pueden no llegar a tener ningún valor de verdad; así pues, las oraciones no satisfacen (ii). En cuanto a (i): evidentemente, muchos tipos de oraciones cambian su valor de verdad ("Tengo hambre", por ejemplo, sería verdadero en boca de alguien en un momento, falso en la de otros en otro momento); incluso algunas instancias de oración pueden mostrarse capaces de cambiar su valor de verdad. (Una instancia de "Hay una persona en esta habitación", escrita en la pizarra de mi oficina, normalmente sería verdadera a las 12 del mediodía y falsa a las 12 de la noche.) Quine ha señalado que podemos especificar una clase de tipos de oración que no cambian su valor de verdad; ésta incluiría a las oraciones que manifiestan leyes físicas o matemáticas, para las que las consideraciones temporales, dice, son irrelevantes, y oraciones completamente especificadas en tiempo y en lugar, con verbos en forma temporal e índices como "ahora" sustituidos por verbos en forma no temporal, fechas y momentos. Quine llama a estos tipos estables "oraciones eternas" (véase cap. 9, § 3).

Portadores de verdad y teoría de la verdad.

Un argumento que se puede presentar a favor de la tesis de que las oraciones son portadores de verdad es éste: algunas teorías de la verdad, ciertas versiones (la de Wittgenstein, pero no la de Austin, por ejemplo) de la teoría de la correspondencia y, con mayor notabilidad, la teoría semántica de Tarski, explotan la estructura gramatical en la definición de verdad (para detalles, cfr. cap. 7). Por supuesto, las oraciones tienen estructura gramatical; los enunciados y las proposiciones, sin embargo, al ser extralingüísticos, no la tienen. Y ya que oraciones en diferentes lenguajes con estructuras gramaticales diferentes pueden formar el mismo enunciado al pronunciarse y estar expresadas por la misma proposición, será difícil para las proposiciones o los enunciados "tomar prestada" una estructura de las oraciones a las que expresan o que los forman. Sin embargo, mientras unos considerarían la plausibilidad de la teoría de Tarski como una razón para considerar las oraciones como portadores de verdad, otros, sobre la base de su convicción de que las oraciones no pueden ser portadores de verdad, están dispuestos a rechazar la teoría de Tarski (véase White, 1970, págs. 94-9). Además, unos argumentarían que el hecho de que la definición de verdad de Tarski tenga que ser relativa al lenguaje, que define "verdadero en L" más que "verdadero", es un punto en contra. Y otros proponen modificar la teoría de Tarski de modo que se hiciera aplicable a proposiciones (Popper, 1972) o enunciados (Davidson, 1967).

Después de pensarlo, se puede ver que exigencias tales como (i) y (ii), que a los que se oponen a las oraciones imponen implícitamente los portadores de verdad, se relacionan con supuestos —que resultan ser cuestionables— sobre la teoría de verdad: que una teoría correcta será bivalente y producirá verdad atemporal. Yo no discutiré aquí la cuestión del supesto carácter atemporal de la verdad, sino que simplemente remitiré al lector a Putnam, 1957, y Haack, 1974, págs. 69-70. Uno o dos breves comentarios acerca de la bivalencia pueden venir bien. Con frecuencia, un excesivo hincapié en la idea de que cada ítem de un tipo ha de tener un valor de verdad que sería verdadero o falso es síntoma de un indeseable tipo de conservadurismo acerca de las lógicas divergentes. Pues algunos autores reaccionan ante la sugerencia de que ciertas oraciones, al no ser ni verdaderas ni falsas, exigen quizás una lógica no bivalente, arguyendo que tales oraciones no pueden formar enunciados o no pueden expresar proposiciones, y, por tanto, ya que la lógica se ocupa de enunciados o proposiciones, están fuera de su ámbito (véase, por ejemplo, Lewy, 1946, y cfr. Kripke, 1975, pág. 700n; la tesis de "ningún ítem" se discute en Haack, 1974, págs. 47-53). Esta reacción es apta para trivializar temas serios.

Algunas teorías de la verdad —las descendientes de la teoría de la "redundancia" de Ramsey— sugieren una solución radical al problema de los portadores de verdad. La cuestión de que la verdad es una propiedad de algo que surge del supuesto —bastante natural— de que la verdad es una propiedad. Pero estas teorías (cap. 7, § 7) niegan que la verdad sea una propiedad, y, por tanto, desvían la cuestión acerca de qué algo es una propiedad. El pensar esto es excusable; en vista del poco satisfactorio estado de la cuestión, es una virtud de tales teorías evitarlo.

5 El problema reformulado

Los argumentos en contra de las oraciones parecen imponer requisitos sobre los portadores de verdad que los enunciados y las proposiciones tampoco logran, y que de todos modos son cuestionables; algunos argumentan a favor de las oraciones como portadores de verdad tal como exige de ellas la teoría de la verdad de Tarski, otros rechazan la teoría de Tarski porque requiere oraciones como portadores de verdad... Se empieza a pensar que la formulación del problema puede que necesite mejorarse. Yo creo que el problema que es la base del debate *puede* reformularse de un modo que lo hace mucho más manejable. Empecé, como puede recordarse, observando que las cuestiones acerca de oraciones, enunciados y proposiciones, etc., surgieron, como problemas filosóficos propios de la lógica, de preguntas acerca de las relaciones entre argumentos formales e informales. Ahora bien, supongamos que tenemos un argumento del cálculo de oraciones como:

$$\begin{array}{l} p \vee -q \\ \underline{-p \qquad} \\ -q \end{array}$$

y queremos saber qué argumentos informales pueden adecuadamente considerarse como instancias de él. Evidentemente, esto es algo que necesitamos saber si la lógica formal tiene que ayudarnos a valorar los argumentos informales. Una cuestión que se necesita contestar ahora es qué puede representar a las "p" y "q" en los argumentos informales. Bueno, queremos decir que cualquier oración que se quiera puede corresponderse con "p" y "q", con tal que la misma oración se corresponda en cada ocurrencia. Esto es un principio, pero son necesarias más restricciones: las oraciones declarativas se pueden corresponder con "p" y "q", pero no las oraciones interrogativas ni las imperativas; si las oraciones que se corresponden con "p" y "q" están en forma temporal, entonces la referencia al tiempo debería permanecer constante del principio al final del argumento; si contienen índices como "yo", "él", "ahora", su referencia debería permanecer constante en el argumento; y si son ambiguos, deberían usarse en el mismo sentido de principio a fin del argumento. De lo contrario, aunque el argumento formal sea válido, su supuesto análogo en el argumento informal es susceptible de ser *in*válido; si, por ejemplo, no se respeta la última condición, tenemos una "falacia de equivocidad".

Este modo de representar el problema tiene la ventaja de la neutralidad metafísica, de modo que no eriza el pelo ni a los nominalistas ni a los platónicos; y, sin embargo, parece contestar las cuestiones oportunas acerca de cómo aplicar la lógica formal al argumento informal. Y el hecho de que las restricciones sobre lo que, en el argumento informal, puede colocarse donde lo hacen "p" y "q" en la lógica formal, reflejen las condiciones de identidad propuestas para los diversos candidatos a portadores de verdad es una confirmación de mi afirmación de que se ha reformulado el problema original más que reemplazarlo por uno diferente. (Pero resulta, y es interesante, que en su versión reformulada el problema surge en teorías que no consideran la verdad como una propiedad.)

El problema reformulado no se refiere a los portadores de verdad directamente, sino, más bien, pregunta qué caprichos de las oraciones si se ponen en el lugar de "p" y "q" pueden interferir con la validez. En el caso de la lógica de oraciones clásica esto equivale a preguntar cómo se puede impedir que tenga la misma oración el mismo valor de verdad en las diferentes ocurrencias en un argumento; es porque una oración ambigua puede ser verdadera en las premisas y falsa en la conclusión, por lo que le equivocidad interfiere con la validez. Pero la mayor generalidad del problema reformulado podría inducirnos a lanzar otra mirada a las relaciones entre validez y verdad.

La validez de nuevo

Antes advertí que insistir en que la lógica sólo se ocupa de ítems que son verdaderos o falsos es dejar de lado sin caridad a las lógicas no bivalentes. Además, si iniciativas tales como la lógica imperativa o la erotética (lógica interrogativa) son viables, debe reconocerse que la lógica puede tratar oraciones incapaces de verdad o falsedad. Ahora bien, la interpretación extrasistemática de validez que di en el cap. 2 estaba en términos de preservación de verdad. Sin embargo, si se tiene en cuenta seriamente la posibilidad de lógicas que traten de ítems que no son portadores de verdad, es probable que sea necesaria una concepción extendida de validez. Si, por ejemplo, se desea manejar oraciones imperativas, puede resultar adecuado definir un análogo de verdad (un "valor designado" si se quiere) que *sea* aplicable a ellas. Ross, 1968, sugiere: "¡*p*!" se satisface sii "*p*" es verdadera. (Ejemplo: "¡Cierra la puerta!" se satisface[2] sii "La puerta estaba cerrada" es verdadera.) La validez, para la lógica imperativa, sería entonces preservación de satisfacción más que preservación de verdad.

No sería desalentador —ni siquiera muy sorprendente— que desarrollos tales como la lógica imperativa o la no bivalente puedan exigir cambios o extensiones de la concepción intuitiva de validez a la que el aparato lógico estándar dé expresión formal. Con frecuencia basta con la modificación o extensión de esas ideas que una ciencia cultiva.

[2] Este uso de "satisfacción" debería quedar como diferente del que se introduce en el siguiente capítulo, en la discusión de la teoría de la verdad de Tarski.

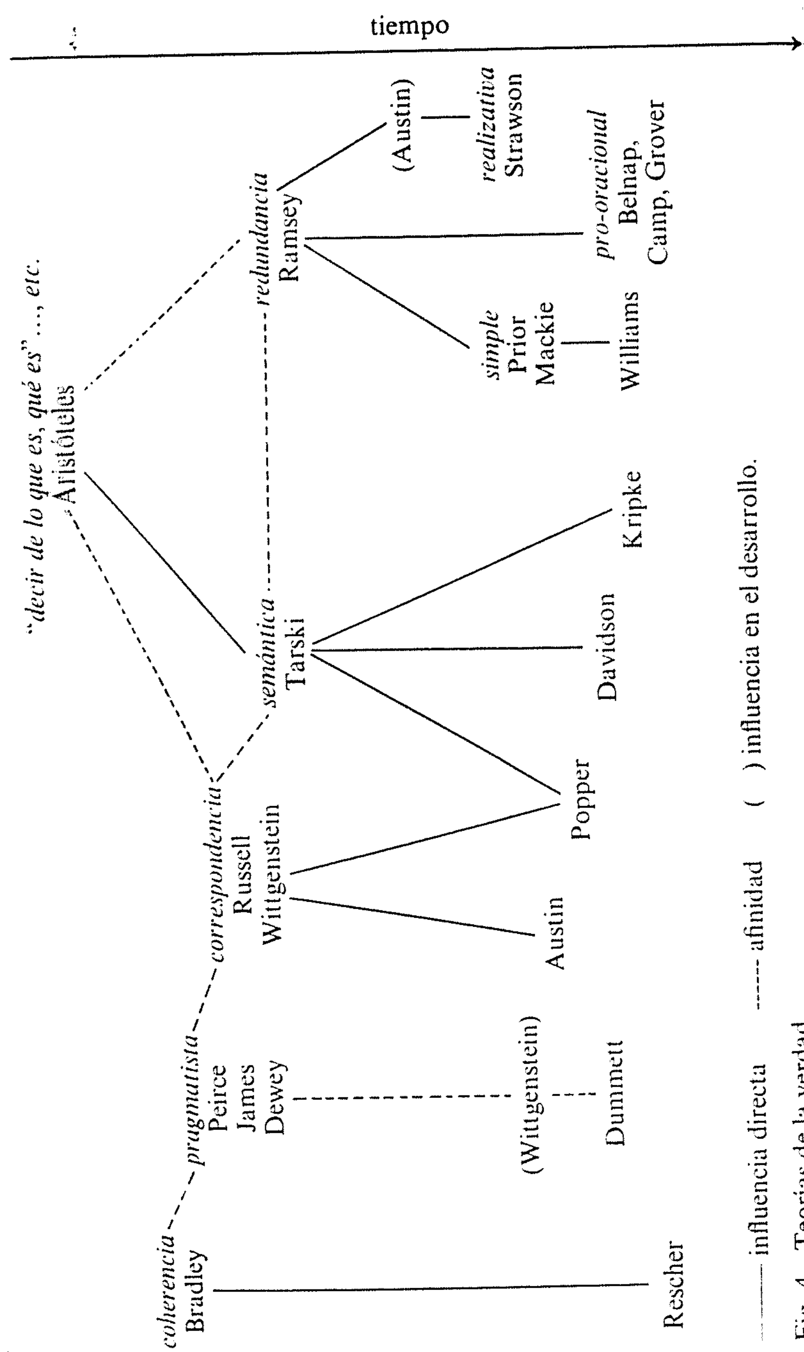
7
Teorías de la verdad

1 UN ESBOZO BREVE[1]

El objeto de este apartado es bosquejar los principales tipos de teorías de la verdad que han sido propuestos, e indicar cómo se relacionan unas con otras. (En los apartados siguientes se discutirán algunas de las teorías detalladamente.)

Para las teorías de la *coherencia*, la verdad consiste en las relaciones de coherencia entre un conjunto de creencias. Las teorías de la coherencia fueron propuestas, por ejemplo, por Bradley en 1914, y también por algunos positivistas oponentes del idealismo, como Neurath en 1932; más recientemente, Rescher en 1973 y Dauer en 1974, han defendido este tipo de enfoque. Para las teorías de la *correspondencia*, la verdad de una proposición consiste, no en sus relaciones con otras proposiciones, sino en su relación con el mundo, en su correspondencia con los hechos. Teorías de este tipo fueron sostenidas al mismo tiempo por Russell en 1918 y Wittgenstein en 1922 durante el período de su adhesión al atomismo lógico; Austin defendió una versión de la teoría de la correspondencia en 1950. La teoría *pragmatista*, desarrollada en los trabajos de Peirce (véase, por ejemplo, 1877), Dewey (véase, por ejemplo, 1901) y James (véase, por ejemplo, 1909), guarda afinidades tanto con la teoría de la coherencia como con la de la correspondencia, admitiendo que la verdad de una creencia deriva de su correspondencia con la realidad, pero insistiendo también en que la verdad de una creencia se manifiesta por la supervivencia ante la prueba de la experiencia, su coherencia con otras creencias; la explicación de la

[1] Los Ponentes de las teorías que discutiré toman diferentes puntos de vista acerca de qué clases de ítems son portadores de verdad. En lo que sigue hablaré indistintamente —dependerá de qué teoría esté discutiendo— de "creencias", "oraciones", "proposiciones", etc., como verdaderas o falsas; solamente cuando la elección de un término u otro sea significativa llamaré la atención sobre ella.



—— influencia directa ----- afinidad ( ) influencia en el desarrollo.

Fig. 4 Teorías de la verdad

verdad propuesta por Dummett en 1959 tiene, a su vez, grandes afinidades con el punto de vista pragmatista.

Aristóteles había observado que "decir de lo que es que no es, o de lo que no es que es, es falso, mientras que decir de lo que es que es, o de lo que no es que no es, es verdadero". Al proponer su teoría *semántica* de la verdad, Tarski, en 1931 y en 1944, intenta explicar el sentido de "verdadero" que este dictum captura. La verdad, en la explicación de Tarski, se define en términos de la relación semántica de satisfacción, una relación entre oraciones abiertas (como "$x > y$") y objetos no lingüísticos (como los números 6 y 5). La teoría de la verdad propuesta recientemente por Kripke en 1975 es una variante de la de Tarski, modificada esencialmente para dar cuenta de un modo más sofisticado de las paradojas semánticas. La explicación de la verdad de Popper y su teoría de la verosimilitud o proximidad a la verdad se basa en la teoría de Tarski, la cual considera Popper que suministra una versión más precisa de las teorías tradicionales de la correspondencia.

La teoría de la *redundancia* de la verdad, ofrecida por Ramsey en 1927, afirma que "verdadero" es redundante, pues decir que es verdadero que "p" es equivalente a decir que "p". Es evidente que esta explicación tiene algunas afinidades con el dictum de Aristóteles y, consecuentemente, con algunos aspectos de la teoría de Tarski. Han existido diversas variantes recientes de la teoría de Ramsey: la explicación "realizativa" de Strawson (1949); la teoría "simple" de la verdad sugerida por Prior en 1971 y ampliada por Mackie en 1973 y Williams en 1976, y la teoría "pro-oracional" presentada por Grover, Camp y Belnap en 1975.

Definiciones versus criterios de verdad

Corrientemente se hace una distinción (por ejemplo, por Russell en 1908b; Rescher en 1973, cap. 2; Mackie en 1973, cap. 1) entre *definiciones* de verdad y *criterios* de verdad; la idea es, aproximadamente, que mientras una definición da el significado de la palabra "verdadero", un criterio da un test mediante el cual decir si una oración (o lo que fuere) es verdadera o falsa —como, por ejemplo, se podría distinguir entre, por una parte, fijar el significado de "tener fiebre" como tener una temperatura superior a algún punto dado y, por otra parte, especificar procedimientos para decidir si alguien *tiene* fiebre.

Esta distinción necesita un tratamiento cuidadoso. Pueden venirle sospechas a uno por la existencia del desacuerdo acerca de qué teorías de la verdad se consideran como definicionales y cuáles como criteriales: por ejemplo, mientras el mismo Tarski rehúsa cualquier interés en proporcionar un criterio de verdad, y Popper considera como una ventaja de la teoría semántica el que ésta sea definicional en vez de criterial, Mackie explica la teoría de Tarski como una teoría que aspira a proporcionar un criterio y la critica por ello. Y las sospechas serían confirmadas por algunos usos claramente inapropiados de la distinción. Por ejemplo, Russell acusó a los pragmatistas de haber confundido la definición y el criterio de verdad, cuando los pragmatistas sostenían que el significado de un término es correctamente dado precisamente al suministrar criterios para su aplicación. (No es del todo insólito, me temo, para un filósofo que deliberadamente identifica las *A* y las *B* encontrarse frente a la crítica de que ha "confundido" las *A* y las *B*.)

Sin embargo, no puede uno simplemente decidir el abstenerse de usar la distinción, problemática como es, debido a su importancia respecto a cuestiones tales como la de si las teorías de la coherencia y de la correspondencia deberían ser consideradas como rivales entre las que uno está obligado a escoger, o como complementarias entre sí de tal manera que la correspondencia suministrase la definición y la coherencia el criterio. Esta cuestión está en litigio, incluso entre los proponentes de la teoría de la coherencia. Así Bradley, al reconocer que "La verdad para ser verdad debe ser verdadera de algo y este algo no es en sí mismo verdad" (1914, pág. 325), parece admitir que una explicación del significado de la verdad puede requerir apelación a algo semejante a la correspondencia, mientras la coherencia es más bien una marca, un test de la verdad. Blanshard, por contraste, insiste en que la verdad *consiste en* la coherencia, que es una definición tanto como un criterio. Esta insistencia parece basarse en la convicción de que debe haber alguna íntima conexión entre un criterio seguro y lo que es un criterio *de*. Él argumenta que la coherencia no podría ser el test de la verdad, sino que la correspondencia sería el significado de la verdad, puesto que entonces no hay ninguna explicación de por qué las creencias coherentes serían aquellas que se correspondiesen con los hechos; si la coherencia ha de ser un test fiable de la verdad, debe serlo porque es constitutiva del significado de la verdad (véase Blanshard, 1939, pág. 268).

Rescher propone (1973, caps. 1 y 2) eludir este argumento distinguiendo entre criterios *garantizadores* (infalibles) y *autorizadores* (falibles), y argumentando que solamente en el caso de los criterios garantizadores debe existir la conexión con la definición que Blanshard considera inevitable. Esta distinción ilumina algunas cuestiones anteriormente tratadas. Rescher considera *C* como un criterio garantizador de *x* si:

necesariamente (*C* sii *x* se da)

Pero —como observa Rescher— en este sentido cualquier definición de verdad proporcionaría también un criterio infalible de verdad. Por ejemplo, si la verdad consiste en la correspondencia con los hechos, entonces, necesariamente, si "*p*" se corresponde con los hechos, "*p*" es verdadera, por tanto, la correspondencia es un criterio infalible[2]. (La idea de que Tarski proporciona un criterio de verdad puede derivar de esta concepción de los criterios.)

Así, si uno posee una definición, posee por esa razón un criterio "garantizador". Sin embargo, la conversa es algo menos sencilla. Por ejemplo, es un criterio garantizador de que un número es divisible por 3 el que la suma de sus dígitos sea divisible por 3, pero esto, pienso yo, no es lo que significa que un número sea divisible por 3. Más bien: si uno tiene un criterio garantizador, entonces o *es* una definición o es *una consecuencia lógica de* una definición.

Sin embargo, un criterio autorizador es falible: no es necesariamente el caso de que (*C* sii *x* se da); así, o es verdadero, aunque no necesariamente, que *C* sii *x* se da, o quizás no es invariablemente verdadero que *C* sii *x* se da. (Rescher considera el segundo tipo del caso, pero no el primero.) Por consiguiente, un criterio autorizador de *x* es distinto de una definición de *x* —no necesita estar relacionado lógicamente con el significado de "*x*"[3].

Pero ahora bien, ¿por qué, si cualquier definición proporciona un criterio garantizador, deseamos siempre un criterio autorizador? Pienso que la respuesta es más bien clara, pero difícil de expresar con precisión: si deseamos averiguar si se da *x*, querríamos, en el mejor de los casos, un indicador fiable de la presencia de *x* que sea *más fácil de descubrir que se dé* que el mismo *x*. Una definición proporciona un indicador que es perfectamente fiable, pero exactamente tan difícil de descubrir que se dé como el mismo *x*; un criterio autorizador proporciona un indicador que puede resultar no del todo fiable, pero que, a modo de compensación, es más fácil descubrir que se dé. Por ejemplo, podríamos considerar las manchas características como un criterio autorizador del sarampión; no como un test infalible, ya que no es lógicamente necesario que uno tenga las manchas sii tiene el sarampión, sino mucho más fácilmente descubrible que, digamos, la presencia de una determinada bacteria

[2] Si uno identifica significado y criterio —como hacen los pragmatistas—, entonces se ve obligado a sostener que el criterio ha de ser garantizador. Esto resultará pertinente más adelante para la discusión, en el § 6, del argumento de Popper de que la teoría pragmatista de la verdad amenaza la falibilidad.

[3] Rescher no amplía explícitamente el "necesariamente" en la explicación que presenta de criterio garantizador, pero los indicios contextuales indican que él tiene en mente la necesidad lógica, que es la interpretación que yo he usado. Si se incluyesen tests físicamente necesarios, los apartados precedentes y algunos posteriores deberían ser reescritos para permitir criterios relacionados con aquello de lo que son un test por necesidad física, así como criterios relacionados por necesidad lógica para considerarlos como garantizadores. Por supuesto, la distinción entre necesidad lógica y física —y en realidad la distinción entre lo necesario y lo contingente— no deja de ser problemática.

que es (o así lo supondré en virtud del argumento) el criterio garantizador.

Hasta aquí, pues, triunfa la defensa que Rescher hace del punto de vista de Bradley de la coherencia como un criterio (esto es, un criterio autorizador), pero no una definición de la verdad, contra el argumento de Blanshard en pro de una conexión inevitable entre definición y criterio. Sin embargo, es pertinente que una versión más débil de la idea de Blanshard parece actuar incluso a favor de los criterios autorizadores. Parece plausible argüir que, si C es un criterio autorizador (incluso en el caso menos favorable de que su presencia no esté invariablemente correlacionada con la de x), entonces será necesario *algún* tipo de conexión entre x y C —no una conexión lógica, por supuesto, sino quizás una conexión causal, por ejemplo—. Consideremos, de nuevo, las manchas como un criterio autorizador del sarampión; hay una conexión causal entre las manchas y la enfermedad de la cual son el síntoma. Y, en verdad, esto es relevante para un aspecto de la explicación de Bradley que Rescher no tiene en cuenta. Es plausible pensar que Bradley creía que existe una conexión entre el hecho de que las creencias de uno sean coherentes y su correspondencia con la realidad (i.e., entre el criterio autorizador y la definición), porque él sostiene que la realidad es coherente.

El concepto de verdad es tan importante para la epistemología como para la filosofía de la lógica. Algunas teorías de la verdad tienen un importante componente epistemológico, se preocupan de la accesibilidad de la verdad; y la búsqueda de un criterio de verdad es a menudo una manifestación de esta preocupación. Es perceptible que, en conjunto, las teorías situadas en la parte izquierda del croquis de teorías de la verdad (fig. 4) toman la dimensión epistemológica más seriamente que las situadas en la parte derecha, siendo la teoría de la coherencia y la pragmatista epistemológicamente ricas, mientras que las teorías de la redundancia, por otra parte, carecen virtualmente de "carne" epistemológica (como dice Mackie).

2 Teorías de la correspondencia

Tanto Russell como Wittgenstein, durante sus períodos de "atomismo lógico"[4], ofrecieron definiciones de la verdad como correspondencia de una proposición con un hecho.

Las proposiciones, de acuerdo con Wittgenstein, son complejos

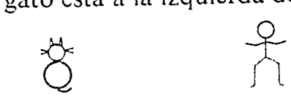
[4] Wittgenstein fue el creador del atomismo lógico, pero la versión de Russell apareció primero en sus conferencias de 1918, mientras que la de Wittgenstein fue presentada en 1922 en el *Tractatus*.

verbales; las proposiciones moleculares (como "$Fa \vee Gb$") están compuestas veritativo-funcionalmente de proposiciones atómicas (como "Fa"). El mundo consta de simples, o átomos lógicos, en diferentes complejos o disposiciones, que son hechos. Y en un lenguaje perfectamente perspicuo, la disposición de las palabras en una proposición atómica verdadera reflejaría la disposición de los simples en el mundo; la "correspondencia" consiste en este isomorfismo estructural. Las condiciones de verdad de las proposiciones moleculares pueden entonces darse: "$-p$" será verdadera sólo en el caso que "p" no sea verdadera, "$p \vee q$" será verdadera sólo en el caso de que o bien "p" sea verdadera o bien "q" sea verdadera, y así sucesivamente.

La versión de Wittgenstein del atomismo lógico es austera; Russell la enriqueció con una teoría epistemológica de acuerdo con la cual los simples lógicos, acerca de cuyo carácter Wittgenstein es agnóstico, son datos sensoriales, que Russell consideró que eran los objetos de conocimiento directo, y el significado de una proposición se supone que se deriva de que esté compuesta de nombres de objetos de conocimiento directo. Estas adiciones epistemológicas no afectan vitalmente al núcleo de la explicación de la verdad; pero algunas otras diferencias entre las versiones de Russell y Wittgenstein son más relevantes. La explicación de Russell tiene la virtud de reconocer las dificultades que hay en considerar todas las proposiciones moleculares, especialmente las proposiciones de creencias y proposiciones cuantificadas, como funciones de verdad de las proposiciones atómicas. Otras características de la versión de Russell, sin embargo, parecen crear dificultades innecesarias; por ejemplo, él admite (aunque no con total confianza a causa de la reacción adversa que esta tesis recibió en Harvard) hechos tanto negativos como positivos, de tal manera que la verdad de la negación de "p" puede consistir en su correspondencia con el hecho de que no "p", más bien que en que "p" deje de corresponder a los hechos; y la sugerencia de que hay dos relaciones de correspondencia, una de las cuales relaciona las proposiciones verdaderas con los hechos y la otra las proposiciones falsas, parece gratuita y, realmente, en vista de la admisión de hechos negativos parece doblemente gratuita.

Numerosos críticos han observado que la dificultad en la teoría de la correspondencia está en que su idea clave, la correspondencia, no se clarifica adecuadamente. Incluso en los casos más favorables, el isomorfismo que se requiere entre la estructura de una proposición y la estructura de un hecho implica dificultades; consideremos:

El gato está a la izquierda del hombre (la proposición)

 (el hecho correspondiente)

incluso aquí (como concede Russell, págs. 315-16) parece como si el hecho tuviera dos componentes y la proposición al menos tres; y, por supuesto, las dificultades serían mucho mayores en otros casos (consideremos "*a* es rojo", "*a* está casado con *b*", o en este caso "el gato está a la derecha del hombre"). La interpretación de la correspondencia como un isomorfismo estructural está íntimamente conectada con la teoría acerca de la estructura última del mundo y con el ideal de un lenguaje perfectamente perspicuo, tesis características del atomismo lógico. La cuestión que surge, por tanto, es si la teoría de la correspondencia se puede divorciar del atomismo lógico, y, si se puede, qué explicación se podría dar entonces de la relación de correspondencia.

Austin da en 1950 una nueva versión de la teoría de la correspondencia cuyo estudio ofrece algunas respuestas. La versión de Austin no cuenta ni con la metafísica atomista ni con el lenguaje ideal; la relación de correspondencia se explica no en términos de un isomorfismo estructural entre proposición y hecho, sino en términos de relaciones puramente convencionales entre las palabras y el mundo. La correspondencia se explica mediante dos tipos de "correlación":

(i) "convenciones descriptivas" que correlacionan palabras con *tipos* de situación

y

(ii) "convenciones demostrativas" que correlacionan palabras con situaciones *específicas*

La idea es que en el caso de un enunciado tal como "tengo prisa", proferido por *s* en *t*, las convenciones descriptivas correlacionan las palabras con situaciones en las cuales alguien tiene prisa, y las convenciones demostrativas correlacionan las palabras con el estado de *s* en *t*, y que el enunciado es verdadero si la situación específica correlacionada con las palabras por (ii) es del tipo correlacionado con las palabras por (i). Austin subraya el carácter convencional de las correlaciones; *cualquier* palabra se podría correlacionar con *cualquier* situación; la correlación no depende en modo alguno del isomorfismo entre palabras y mundo.

Una dificultad de esta explicación de la correspondencia, que apela esencialmente a *ambos* tipos de correlación, es que se aplica directamente sólo a los enunciados formados por oraciones indexadas (*indexical*), puesto que las convenciones demostrativas no tendrían un papel que jugar en el caso de oraciones como "Julio César era calvo" o "Todas las mulas son estériles", que no se pueden usar en enunciados que se refieren a situaciones diferentes. (Los comentarios de Austin sobre estos casos, pág. 23n, no son demasiado convincentes.)

Por otra parte, la versión de Austin creo que mejora la explicación de "los hechos" de Russell. Es difícil exponer la cuestión con claridad, pero, dada su gran importancia, vale la pena exponerla aunque sea con alguna vaguedad. Russell tiende a hablar como si la verdad de "*p*" consistiera *en su correspondencia con el hecho de que "p"*; pero el problema con esto es que la relación entre "*p*" y el hecho de que "*p*" es precisamente *demasiado* estrecha, que "*p*" no podría dejar de corresponder a *ese* hecho. Su evasiva acerca de los criterios de individuación de los hechos puede indicar que él era consciente de esta incomodidad. La versión de Austin, sin embargo, localiza la verdad del enunciado de que "*p*" no en su correspondencia con el hecho de que "*p*", sino más bien en que los *hechos* sean como dice "*p*", o, según lo expresa Austin, en las convenciones demostrativas que correlacionan "*p*" con una situación que es del tipo con el que las convenciones descriptivas lo correlacionan. (Austin es consciente de esta diferencia: véase 1950, pág. 23, y cfr. Davidson, 1973, y O'Connor, 1975.)

3 Teorías de la coherencia

Una teoría de la coherencia de la verdad fue sostenida por los idealistas (discutiré la explicación de Bradley, pero puntos de vista afines fueron sostenidos por sus antecesores filósofos alemanes Hegel y Lotze) y también por algunos de sus oponentes positivistas lógicos. Así la relación entre las teorías de la coherencia y el idealismo es más bien semejante a la que hay entre las teorías de la correspondencia y el atomismo lógico —en que en cada caso la teoría de la verdad llegó a divorciarse de la perspectiva metafísica con la que estaba original y característicamente asociada.

Será útil —porque de este modo se pueden aclarar algunas relaciones significativas entre las teorías de la coherencia y la correspondencia— comenzar por el medio con la defensa que hace Neurath del punto de vista de la coherencia. Una breve historia no estará de más: los positivistas lógicos, bajo la influencia del *Tractatus* de Wittgenstein, suscribieron al principio un punto de vista del carácter de la verdad como correspondencia. Estaban, sin embargo, fuertemente motivados por inquietudes epistemológicas y, en consecuencia, deseaban un test (un criterio autorizador) de la verdad —una manera de saber si una oración corresponde realmente o no a los hechos. Carnap y Schlick abordaron el problema en dos partes; argumentaban que los enunciados que registran una experiencia perceptual inmediata son incorregibles, es decir, podemos verificar directamente que corresponden a los hechos, y la verdad de otros enunciados se puede entonces comprobar por medio de sus relaciones lógicas con éstos. Un rasgo característico de la teoría de la correspondencia —que la verdad radica en una relación entre las creencias y el mundo— está ya modificado: el test de la verdad de todos los enunciados, excepto los perceptuales, deriva de sus relaciones con otros enunciados, los perceptuales, que se supone que se verifican por confrontación directa con los hechos. Neurath, sin embargo, planteó dudas acerca de la supuesta incorregibilidad de los "protocolos", y habiendo negado así la posibilidad de una comprobación directa incluso de la correspondencia de las creencias perceptuales con los hechos, mantuvo que el único test de la verdad consiste en las relaciones entre las mismas creencias. Nuestra búsqueda del conocimiento requiere un reajuste constante de creencias cuyo objetivo es un conjunto de creencias tan exhaustivo cuanto permita la consistencia. (Esto recuerda fuertemente el "método de los máximos y mínimos" de la epistemología de James (1907); la posición de Quine en "Dos dogmas del empirismo" (1951), donde suscribe la metáfora de Neurath del proceso de adquisición del conocimiento como el proceso del reparar una balsa mientras se flota sobre ella, es similar. Cfr. Hempel, 1935, para una explicación excelente del desarrollo del punto de vista positivista sobre la verdad, y Scheffer, 1967, cap. 5, para una viva descripción "golpe a golpe" de la controversia entre Schlick y Neurath.)

La posición final de Neurath tiene mucho en común con la explicación de Bradley del test de la verdad como "sistema", que él explica como requiriendo a la vez la *consistencia* y *exhaustividad* del conjunto de creencias. Y en Bradley como en Neurath la apelación a la coherencia está conectada con la negación de que nuestro conocimiento tenga cualquier base incorregible en los juicios de percepción. Sin embargo, la teoría de Bradley tiene íntimas conexiones con su idealismo absoluto. Brevemente y por encima, la realidad, según Bradley, es ella misma esencialmente un todo unificado coherente. (La metafísica pluralista del atomismo lógico de Russell estaba motivada por la reacción contra el monismo de los idealistas.) Y mientras Bradley concedía algo a la idea de la verdad como correspondencia con la realidad, sostuvo que, estrictamente hablando, nada menos que el conjunto de creencias plenamente exhaustivo y consistente al que aspiramos es realmente verdadero: en el mejor de los casos, conseguimos una verdad parcial —*parte* de la verdad no es completamente verdadera—. El objeto de estas observaciones es poner de manifiesto un punto anticipado arriba (§ 1) —que las conexiones entre la concepción que Bradley tiene de la verdad y su concepción de la realidad son lo suficientemente estrechas para que sea un tanto engañoso considerar simplemente que él ofrece la coherencia como test de la verdad mientras que deja la correspondencia como definición de la misma: más bien,

la explicación del éxito de la coherencia como test se deriva de una explicación de la realidad como esencialmente coherente en sí misma.

Una dificultad persistente de la teoría de la correspondencia, como observé más arriba (§ 2), ha sido la dificultad de suministrar una explicación precisa de "corresponde". Un problema similar persigue a la teoría de la coherencia; se necesita especificar exactamente cuáles deben ser las relaciones apropiadas entre creencias para que sean "coherentes" en el sentido requerido. Críticos no simpatizantes con las teorías de la coherencia —Russell, por ejemplo— han tendido a asumir que la simple consistencia es suficiente; Bradley, sin embargo, insistía ya (tan pronto como en 1909, contra la crítica de Stout; véase Bradley, 1914) en que se requiere tanto la exhaustividad como la consistencia.

Rescher, que defiende una epistemología coherentista (la coherencia como el test de la verdad), ofrece una explicación detallada de los requisitos gemelos del "sistema": consistencia y exhaustividad. El problema con el que se enfrenta el coherentista, como lo ve Rescher, es proporcionar un procedimiento para seleccionar a partir de los "datos" incoherentes y posiblemente inconsistentes ("candidatos a la verdad", no necesariamente verdades) un conjunto privilegiado, las creencias garantizadas, aquellas respecto de las cuales se tiene garantía al sostenerlas como verdaderas. Un "subconjunto consistente máximo" (S.C.M.) de un conjunto de creencias se define así: S' es un S.C.M. de S si es un subconjunto no vacío de S que es consistente y al cual no puede añadirse ningún miembro de S que no sea ya un miembro de S' sin generar una inconsistencia. Pero el conjunto de datos es probable que tenga más de un S.C.M.; esta es la base de la crítica de Russell de que la coherencia no puede distinguir la verdad de un cuento de hadas consistente. Para evitar esta dificultad, Rescher propone que los S.C.M. del conjunto de datos sean "filtrados" por medio de un índice de plausibilidad, dividiendo los datos entre aquellos que son y aquellos que no son inicialmente plausibles, y reduciendo así el número de los S.C.M. elegibles. Sin embargo, esto puede resultar insuficiente para distinguir un único S.C.M.; por tanto, Rescher recomienda la adopción de la disyunción de aquellos S.C.M. permitidos por el filtro de la plausibilidad.

Aunque el trabajo de Rescher ha contribuido significativamente a la elaboración detallada de una epistemología coherentista, subsisten dificultades. Un problema obvio es la especificación y justificación de los estándars de plausibilidad (la apelación de Schlick a la pretendida incorregibilidad de los protocolos podría verse como una respuesta alternativa a una dificultad similar). Una dificultad también importante, aunque menos obvia, es que el procedimiento recomendado es, por así decirlo, de carácter estático: nos dice cómo seleccionar un subconjunto privilegiado, "garantizado", a partir de un conjunto inicial de datos, pero correspondientemente subestima la importancia de buscar *nuevos* datos. (La insistencia de Bradley de que sólo el conjunto de creencias más plenamente exhaustivo —la verdad completa— es estrictamente hablando verdadero, podría verse como una respuesta a esta dificultad.) La coherencia constituirá, sin duda, parte, pero no la totalidad de una epistemología satisfactoria.

Hasta aquí he seguido a Rescher (con algunas matizaciones en el caso de Bradley) al considerar que la coherencia ha de entenderse como un test de la verdad que juega un papel epistemológico, mientras que a la correspondencia se le adjudica la parte metafísica. (Cfr. el gran papel desempeñado por la coherencia en la epistemología de Quine desde 1951 a 1970 con su adopción de la definición semántica de la verdad, 1970, cap. 3.) Los pragmatistas, sin embargo, cuestionan esta distinción con su característica teoría criterial del significado.

4 Teorías pragmáticas[5]

Peirce, James y Dewey ofrecen explicaciones característicamente "pragmáticas" de la verdad, que combinan elementos de la coherencia y de la correspondencia.

De acuerdo con "la máxima pragmática", el significado de un concepto viene dado por la referencia a las consecuencias "prácticas" o "experimentales" de su aplicación[6] —"no puede *haber* diferencia", como dijo James (1907, pág. 5), "que no *introduzca* diferencia"—. Así el enfoque que los pragmatistas hacen de la verdad consiste en preguntar por la diferencia que introduce el que una creencia sea verdadera.

De acuerdo con Peirce, la verdad es el final de la investigación, aquella opinión sobre la cual quienes usan el método científico concordarán, o quizás concordarían, si persistiesen el tiempo suficiente. La significación de esta tesis deriva de la teoría de la investigación de Peirce. Muy brevemente: Peirce considera la creencia como una disposición a la acción, y la duda como la interrupción de tal disposición debido a la terquedad por parte de la experiencia; la investigación es estimulada por la duda, que es un estado desagradable que uno intenta reemplazar por una creencia fija. Peirce argumenta que algunos métodos de adquisición de creencias —el método de tenacidad, el método de autoridad, el método *a priori*— son inherentemente inestables, pero el método científico permite que se adquieran (eventualmente) creencias estables, creencias que no serán sacudidas por la duda. Pues el método científico, argumenta Peirce, es el único entre los métodos de investigación exigido por una realidad que es independiente de lo que uno crea, y es por esto por lo que puede conducir al consenso. Así, ya que la verdad es la opinión sobre la que se asentará eventualmente el método científico, y ya que el método científico viene exigido por la realidad, la verdad es la correspondencia con la realidad. Se sigue también que la verdad es satisfactoria para la creencia en el sentido de que es estable, a salvo de la perturbación de la duda.

La mayor contribución de James fue una elaboración de esta idea. La ventaja de poseer creencias verdaderas, argumentaba, era que uno se encontraba de ese modo garantizado contra la experiencia recalcitrante, mientras que las creencias falsas serían eventualmente atrapadas ("La experiencia... tiene formas de *desbordarse*...", 1907, pág. 145). La explicación que hace James de la forma en que uno ajusta sus creencias cuando entra nueva experiencia, maximizando la conservación del antiguo conjunto de creencias al mismo tiempo que restaurando la consistencia —sorprendentemente semejante al punto de vista epistemológico de Quine en 1951—, introduce un elemento de coherencia. Las creencias verdaderas, comenta James, son aquellas que son verificables, i.e., aquellas que son a la larga confirmadas por la experiencia.

Hasta aquí he insistido en las afinidades entre los puntos de vista de Peirce y James, pero hay algunas diferencias que deben ser mencionadas. En primer lugar, mientras Peirce era un realista, James se inclinaba hacia el nominalismo (cfr. Haack, 1977), y por esta razón se vio turbado por las posibles-pero-todavía-no-realizadas verificaciones a las cuales le remitía la concepción de la verdad como verificabilidad; consecuentemente, aunque en principio admite que las creencias son verdaderas (falsas) aunque nadie las haya todavía verificado (falsificado), en la práctica está suficientemente persuadido de la inutilidad de insistir en esto que él introduce en el lenguaje inconsistentemente, como si nuevas verdades vinieran a la existencia cuando las creencias consiguen ser verificadas. (La idea de que la verdad es *hecha*, que crece, fue establecida por el pragmatista inglés F. C. S. Schiller.) En segundo lugar, James habla a menudo de que la creencia verdadera es la "buena" o la "conveniente" o la "útil" (por ejemplo, 1907, págs. 59, 145). Críticos poco compasivos (por ejemplo, Russell, 1908b; Moore, 1908) han considerado que James hace una crasa, por no decir moralmente objetable, identificación de la creencia verdadera con la que nos agrada. Los comentarios que provocaron esta crítica feroz, cuando se toman dentro del contexto, pueden a menudo leerse mucho más aceptablemente como apuntando a la superioridad de las creencias verda-

[5] Este apartado está inspirado en Haack, 1976c.

[6] Peirce subrayó la conexión de "pragmático" con el uso que hace Kant de "*pragmatische*" para lo empíricamente condicionado, y James acentuó la conexión con el griego "*praxis*", acción.

deras en cuanto están *a salvo de falsificación* (cfr. la propia defensa de James, 1909, pág. 192 —"Ante todo encontramos satisfactoria la *consistencia*"—). Pero James hace también otra afirmación: que puesto que en cualquier tiempo la evidencia de la que disponemos puede ser insuficiente para decidir entre creencias que compiten, nuestra elección puede depender de razones tales como la simplicidad o la elegancia (1907, pág. 142); afirmación ésta que tiene conexiones con su doctrina de "la voluntad de creer".

Dewey adopta la definición de Peirce como "la mejor definición de la verdad" (1938, pág. 345n). Él prefiere la expresión "aseverabilidad garantizada" a "verdad", y añade la tesis de que es precisamente la aseverabilidad garantizada la que caracteriza aquellas creencias a las que damos el título honorífico de conocimiento (cfr. Ayer, 1958). La concepción de la verdad de Dummett, cuya inspiración directa deriva de la obra del último Wittgenstein y del intuicionismo en la filosofía de la matemática, se parece a la concepción de Dewey en su insistencia sobre la aseverabilidad; véase Dummett, 1959.

Las tesis principales de la explicación pragmática pueden resumirse como sigue:

la verdad es:

el final de la investigación	Peirce	James	Dewey
la *correspondencia* con la realidad	Peirce	James	Dewey
la creencia satisfactoria (estable)	Peirce	James	Dewey
la *coherencia* con la experiencia — verificabilidad		James	Dewey
lo que autoriza a la creencia a denominarse "conocimiento"			Dewey

5 La teoría semántica

La teoría de la verdad de Tarski ha sido, entre las más recientes, probablemente la que más influencia ha ejercido y la más ampliamente aceptada. Su teoría se divide en dos partes: proporciona primero *condiciones de adecuación*, i.e., condiciones que cualquier definición aceptable de la verdad debe cumplir; y luego ofrece una definición de la verdad (para un lenguaje formal especificado) que muestra que es adecuada según sus propios estándars. Se examinarán ambas partes de este programa. La detallada formulación de la teoría se encuentra en Tarski, 1931; 1944 es una buena introducción.

No es difícil ver por qué la teoría de Tarski ha podido llegar a ejercer tanta influencia. Por una parte, sus condiciones de adecuación para las definiciones de verdad prometen una especie de filtro

para discriminar, entre las desconcertantes y numerosas teorías de la verdad, aquellas que reúnan condiciones mínimas de aceptabilidad y que por ello tienen alguna perspectiva de éxito. Además, los métodos empleados por Tarski en la definición de la verdad pueden aplicarse a una amplia clase de lenguajes formales. Pero los rasgos mismos de la teoría de Tarski que más contribuyen a su atractivo le crean también problemas, como veremos: ¿puede darse a las condiciones de adecuación de Tarski una motivación independiente? y ¿tienen sus métodos alguna aplicación interesante al problema de la verdad para los lenguajes naturales?

Condiciones de adecuación para las definiciones de verdad

El problema que el mismo Tarski se plantea es el de dar una definición de la verdad que sea a la vez *materialmente adecuada* y *formalmente correcta*; la primera de estas condiciones pone límites al contenido posible y la segunda a la forma posible de cualquier definición aceptable.

Adecuación material

Tarski espera que su definición "atrape el significado real de una vieja noción" (1944, pág. 53). Sin embargo, Tarski cree que la "vieja" noción de verdad es ambigua e incluso dudosamente coherente. Por ello restringe su interés a lo que él llama la "concepción aristotélica clásica de verdad", tal como se expresa en el dictum de Aristóteles:

> Decir de lo que es que no es, o de lo que no es que es, es falso, mientras que decir de lo que es que es, o de lo que no es que no es, es verdadero.

y propone, como condición de adecuación material, que *toda definición aceptable de la verdad tenga como consecuencia* todas las instancias del esquema (T):

(T) S es verdadera sii p

donde "p" puede ser reemplazada por cualquier oración del lenguaje para el cual se está definiendo la verdad y "S" ha de reemplazarse por un nombre de la oración que reemplaza a "p". Una instancia de (T) sería, por ejemplo:

"La nieve es blanca" es verdadera sii la nieve es blanca donde se hace referencia a la oración del lado derecho por medio de su "nombre entrecomillado" del lado izquierdo.

Tarski recalca que el esquema (T) *no es una definición* de la verdad —aunque a pesar de su insistencia ha sido mal entendido en este punto—. Es una *condición de adecuación material:* todas sus instancias deben ser entrañadas por cualquier definición de la verdad que tenga que considerarse como "materialmente adecuada". El *quid* del esquema (T) es que, si se acepta, fija no la intensión o significado, sino la *extensión* del término "verdadera". Supongamos, pues, que tuviésemos dos definiciones de la verdad, D_1 y D_2, y que cada una de ellas fuese materialmente adecuada. Entonces D_1 entrañaría todas las instancias de:

> *S* es verdadera_1 sii *p*

y D_2 todas las instancias de:

> *S* es verdadera_2 sii *p*

de manera que D_1 y D_2 son coextensivas. O, para expresar esencialmente el mismo punto de otro modo, la condición de adecuación material excluiría ciertas definiciones de la verdad, esto es, aquellas que *no* entrañasen instancias del esquema (T).

Pero exactamente ¿qué tipos de definición excluirá la condición de adecuación material? Al responder a esta pregunta usaré una versión débil del criterio: no que todas las instancias del esquema (T) *sean deducibles* de cualquier definición aceptable de la verdad (versión de Tarski), sino que la verdad de todas las instancias del esquema (T) *sea consistente* con cualquier definición aceptable de la verdad. La razón para esta modificación es simplemente que la condición débil de adecuación es mucho más fácilmente aplicable a las definiciones no formales de la verdad. Ahora hay que esperar —y quizás incluso contar con— que ello permitirá las clases de definición que han sido propuestas seriamente y rechazará lo que podríamos denominar teorías "bizarras". Pero el asunto resulta un tanto extraño. Consideremos la siguiente definición de la verdad, que me parece en definitiva extravagante: una oración es verdadera sii es aseverada en la Biblia. Ahora puede suponerse que esta definición (la llamaré "D_B" por abreviar) no entraña todas las instancias del esquema (T), ni, por ejemplo:

> "Varsovia fue bombardeada en la II Guerra Mundial" es verdadera_B sii Varsovia fue bombardeada en la II Guerra Mundial.

Ahora bien, es realmente el caso que alguien que no acepte D_B podría negar:

> "Varsovia fue bombardeada en la II Guerra Mundial" es aseverado en la Biblia sii Varsovia fue bombardeada en la II Guerra Mundial.

Pero una ulterior reflexión pone de manifiesto que un proponente de D_B podría perfectamente bien mantener que su definición entraña todas las instancias de (T): él puede admitir que "Varsovia fue bombardeada en la II Guerra Mundial" es verdadera, pero insistir en que *es* aseverada en la Biblia (quizás en un oscuro pasaje del Apocalipsis), o si él acepta que "Varsovia fue bombardeada en la II Guerra Mundial" no es aseverada en la Biblia, mantendrá también, si es inteligente, la falsedad del lado derecho de la instancia del esquema anterior. De esta manera, y más bien sorprendentemente, la condición de adecuación material de Tarski no puede considerarse como especialmente efectiva para excluir las definiciones de verdad "bizarras".

Sin embargo, la condición de adecuación material *sí* que excluye aparentemente una cierta clase importante de teorías de la verdad, esto es, aquellas según las cuales algunas oraciones (enunciados, proposiciones, fbfs o lo que fuere) no son ni verdaderas ni falsas. Supongamos que "*p*" no es ni verdadera ni falsa; entonces la parte izquierda de:

> "*p*" es verdadera sii *p*

será presumiblemente falsa, mientras que la parte derecha no será ni verdadera ni falsa. De este modo, todo el bicondicional será falso, o de todos modos no-verdadero. (Este argumento podría, sin embargo, evitarse si uno estuviera dispuesto a admitir que las aserciones metalingüísticas tales como " '*p*' es verdadera" no podrían ser ellas mismas ni verdaderas ni falsas.) Se puede argüir que la condición de adecuación material de Tarski excluiría al menos algunas versiones de la teoría de la coherencia; razonablemente *no* excluiría una teoría pragmatista, puesto que la concepción pragmatista del significado calificaría como carente de significado a cualquier oración que no sea ni verificable ni falseable, de modo que no habría oraciones significativas, sino carentes de valor de verdad. Ciertamente parece más bien extraordinario excluir las teorías no bivalentes de la verdad.

La idea que subyace a la condición de adecuación material de Tarski es, presumiblemente, que la verdad del esquema (T) es tan cierta y obvia que es conveniente que uno se sienta seguro al rechazar cualquier teoría de la verdad que sea inconsistente con ella.

Por mi parte, encuentro la certeza y evidencia iniciales del esquema (T) un tanto modificadas cuando resulta que no sólo alguna de las teorías de la verdad seriamente propuestas, sino también algunas teorías muy bizarras, son consistentes con él, mientras algunas otras teorías serias son inconsistentes con él (pero véase Davidson, 1973, para una defensa de la "convención T").

Corrección formal

El requisito formal que Tarski establece concierne a la estructura del lenguaje en el que ha de darse la definición de la verdad, los conceptos que pueden emplearse en la definición y las reglas formales a las que la definición debe conformarse.

Es notorio que los conceptos semánticos manejados imprudentemente tienden a originar paradojas (por ejemplo, la del Mentiroso: "Esta oración es falsa"; la paradoja de Grelling: " 'no verdadero de sí mismo' es verdadero de sí mismo sii no es verdadero de sí mismo", y así sucesivamente). Tarski investiga la paradoja del Mentiroso con cierto detalle y argumenta que la antinomia surge de los supuestos:

(i) Que el lenguaje usado contiene, además de sus expresiones, (*a*) los medios de referirse a esas expresiones y (*b*) predicados semánticos tales como "verdadero" y "falso". A este lenguaje Tarski lo llama "semánticamente cerrado".
(ii) Que las leyes lógicas usuales valen.

No deseando rechazar el supuesto (ii), Tarski concluye que una definición formalmente correcta de la verdad debe expresarse en un lenguaje que no sea semánticamente cerrado.

Específicamente esto significa que la definición de la verdad-en-O, donde O es el *lenguaje objeto* (el lenguaje *para* el cual se está definiendo la verdad), tendrá que darse en un *metalenguaje*, M (el lenguaje *en* el que se define la verdad-en-O). La definición de la verdad, argumenta Tarski, tendrá que ser relativa a un lenguaje, pues una y la misma oración puede ser verdadera en un lenguaje y falsa o carente de significado en otro. El peligro de las paradojas semánticas puede evitarse mediante el recurso a un metalenguaje; la oración del Mentiroso, por ejemplo, se convertirá entonces en la inocua "Esta oración es falsa-en-O", que es ciertamente una oración de M y consecuentemente no es paradójica. La distinción lenguaje objeto/metalenguaje es ciertamente relativa y se requeriría toda una jerarquía de lenguajes para definir la verdad a cada nivel. Puesto que por la condición de adecuación material todas las equivalencias de la forma (T) deben ser implicadas por la definición de

la verdad, M debe contener a O o traducciones de todas las oraciones de O como parte suya, más los medios de referirse a las expresiones de O; pues las instancias de (T) tienen en el lado izquierdo una expresión que denota una oración de O y en el lado derecho una oración de O o una traducción de una oración de O. Adviértase que, al especificar en el metametalenguaje que el metalenguaje, M, debería contener o bien al mismo lenguaje, O, o bien una traducción de cada oración de O, se emplean nociones semánticas (explícitamente en el último caso e implícitamente en el primero, puesto que M debe contener las mismas expresiones de O con las mismas interpretaciones que tienen en O).

Se requiere también que la estructura de O y M sea "formalmente especificable". Pues para definir "verdadera-en-O" será esencial lograr reconocer las fbfs de O, ya que éstas constituyen los elementos a los que se aplica "verdadera-en-O" (Ésta es una de las razones que da Tarski para sentirse escéptico sobre la posibilidad de definir "verdadera-en-español" — o "verdadera" para *cualquier* lenguaje natural; las oraciones de los lenguajes naturales no son, piensa, él, formalmente especificables. Posteriormente los seguidores de Tarski, sobre todo Davidson, se sienten más optimistas respecto de este punto. Se trata de un punto que necesitaré investigar más profundamente.)

Tarski exige también que "las usuales reglas formales de definición sean observadas en el metalenguaje" (1944, pág. 61). Estas reglas incluyen:

(i) ninguna variable libre puede figurar en el *definiens* sin figurar también en el *definiendum*

que excluye, por ejemplo, "Fx = df. ($x + y = 0$)", y

(ii) no pueden figurar en el *definiendum* dos ocurrencias de la misma variable

que excluye, por ejemplo, "Fxx = df. Gx". La condición (i) impide definiciones que podrían llevar a contradicción; la condición (ii) impide definiciones en las que el *definiendum* es ineliminable (cfr. Suppes, 1957, cap. 8).

Cualquier definición aceptable de la verdad debe entonces, según Tarski, satisfacer ambas condiciones: la adecuación material y la corrección formal. Él da una definición y muestra que, según estos estándares, es aceptable.

Definición de la verdad de Tarski

Podría pensarse que el esquema (T), aunque no es él mismo una definición de la verdad, proporciona un camino obvio para dar tal definición. El propio Tarski señala que podríamos concebir cada instancia de (T) como una definición *parcial* de la verdad en el sentido de que cada instancia especifica las condiciones de verdad de una cierta oración específica; de modo que una conjunción de *todas* las instancias del esquema (T), una para cada oración de O, constituiría una definición completa. Tarski, sin embargo, arguye que *no* es posible dar tal definición conjuntiva, porque el número de oraciones de un lenguaje puede ser infinito, y en este caso es imposible de hecho dar todas las instancias requeridas por el esquema (T).

Tampoco, arguye Tarski, puede convertirse el esquema (T) en una definición de la verdad mediante cuantificación universal. Se podría suponer que, usando en la parte izquierda un nombre entrecomillado de la oración usada en la parte derecha, podríamos sencillamente generalizar para obtener:

(D) (p)("p" es verdadera$_O$ sii p)

que aparentemente constituiría una definición completa, y además una definición con la garantía de ser materialmente adecuada, ya que todas las instancias de (T) son instancias suyas. Pero Tarski rechaza esta sugerencia porque cree que el resultado de cuantificar dentro de las comillas carece de significado. Pues, según Tarski (y también según Quine), la expresión obtenida entrecomillando una expresión es una unidad indivisible, análoga a un nombre propio, de modo que:

La nieve es blanca

no es más parte de:

"La nieve es blanca"

que (adaptando un ejemplo de Quine) "oro" lo es de "Teodoro". Tarski concede que si fuera factible considerar la cita como una función, entonces (D) no estaría menos bien-formada que, por ejemplo:

$(x)(x^2 = x \cdot x)$

Tarski piensa, sin embargo, que hay aplastantes objeciones para tratar la cita como una función y, en consecuencia, que (D) no está más bien-formada que, por ejemplo:

(x)(Texas es grande)

Así pues, Tarski piensa que el esquema (T) no sólo no es, sino que tampoco puede convertirse en una definición de la verdad. Así que construye su propia definición por un camino más sinuoso. Considera como un *desideratum* el que no se tome ningún término semántico como primitivo, de manera que cualquier noción semántica en términos de la cual se defina "verdadera" debería ella misma ser previamente definida. Puesto que él va a definir "verdadera" usando el concepto de satisfacción, que es un concepto semántico, esto significa que debe primero definir "satisface".

Explicación informal

El procedimiento es como sigue:

(a) especificar la estructura sintáctica del lenguaje, O, para el cual va a definirse la verdad

(b) especificar la estructura sintáctica del lenguaje, M, en el cual va a definirse verdad-en-O; M debe contener
  (i) o las expresiones de O, o traducciones de las expresiones de O
  (ii) el vocabulario sintáctico, incluyendo los nombres de los símbolos de O, un signo de concatenación (para formar "descripciones estructurales" de expresiones compuestas de O), y variables que tienen como rango las expresiones de O
  (iii) el aparato lógico usual

(c) definir "satisface-en-O", y

(d) definir "verdadera-en-O" en términos de "satisface-en-O"

¿Por qué define Tarski primero "satisface"? Bien, primero, porque considera deseable no emplear en su definición de la verdad ningún primitivo semántico; pues considera que las nociones semánticas no son ninguna de ellas lo suficiente claras preteóricamente para emplearlas con seguridad. Pero, ¿por qué "satisface"? Ésta es una noción adecuada para definir "verdadera" en términos de la misma porque las oraciones compuestas cerradas están formadas de oraciones atómicas *abiertas* más bien que de oraciones atómicas cerradas. Por ejemplo, "$(\exists x)Fx \vee Gx$" está formada de "Fx" y "Gx" por las operaciones de disyunción y cuantificación existencial; y las oraciones abiertas "Fx" y "Gx" no son verdaderas o falsas, sino satisfechas o no por los objetos. La definición de satisfacción es *recursiva* —esto es, se dan primero las definiciones para las oraciones abiertas más simples, y luego se establecen las condiciones en las que son satisfechas las oraciones compuestas abiertas—. (La definición podría, sin embargo, convertirse en una definición explícita.) Este procedimiento proporcionará una definición de la verdad aplicable a todas las oraciones de O.

"*Satisface*": las oraciones abiertas no son verdaderas o falsas; son satisfechas o no por ciertas cosas, pares de cosas, trios de cosas, etc. Por ejemplo: "x es una ciudad" es satisfecha por Londres; "x está al norte de y" es satisfecha por ⟨Londres, Brighton⟩; "x está entre y y z" es satisfecha por ⟨Londres, Brigton, Edimburgo⟩..., etc. ("⟨..., ...⟩" indica el *n-tuplo-ordenado* de los n elementos que aparecen entre los corchetes en ángulo.) El orden de los elementos es obviamente importante, puesto que ⟨Londres, Brighton⟩ satisface "x está al norte de y", pero ⟨Brighton, Londres⟩ no lo satisface. La satisfacción es una relación entre oraciones abiertas y n-tuplos ordenados de objetos. Para evitar las dificultades que surgen del hecho de que las oraciones abiertas pueden tener 1, 2, ... o *cualquier* número de variables libres, Tarski define la satisfacción como una relación entre oraciones abiertas y secuencias *infinitas*, bajo la convención de que "$F(x_1 \ldots x_n)$" va a ser satisfecha por la secuencia $\langle O_1 \ldots O_n, O_{n+1} \ldots\rangle$ solamente en el caso de que sea satisfecha por los primeros n miembros de la secuencia; los miembros subsiguientes se ignoran.

La negación de una oración abierta S_1 será satisfecha justamente por aquellas secuencias que no satisfacen a S_1; y la conjunción de S_1 y S_2 justamente por aquellas secuencias que satisfacen a S_1 *y* satisfacen a S_2. La cuantificación existencial de una oración abierta será satisfecha por una secuencia de objetos solamente en el caso de que haya alguna otra secuencia de objetos, que difiera de la primera en a lo sumo el i-ésimo lugar (donde el i-ésimo es la variable ligada por el cuantificador) que satisfaga la oración abierta resultante de la eliminación del cuantificador. Por ejemplo, la secuencia ⟨Inglaterra, Londres, Edimburgo...⟩ satisface "$(\exists x)(x$ es una ciudad entre y y $z)$" porque, por ejemplo, la secuencia ⟨York, Londres, Edimburgo⟩ satisface "x es una ciudad entre y y z".

"*Verdadera*": las oraciones cerradas son casos especiales de oraciones abiertas, a saber, las que *no* tienen variables libres. El primer miembro de una secuencia y todos los miembros subsiguientes son irrelevantes en cuanto a si la secuencia satisface o no una oración abierta de cero lugares, es decir, una oración *cerrada*. Así, Tarski define una oración como *verdadera solamente en el caso de que sea satisfecha por todas las secuencias*, y como *falsa solamente en el caso de que no sea satisfecha por ninguna*. Este procedimiento puede resultar menos misterioso considerando un ejemplo. La oración abierta de 2 lugares "x está al norte de y" es satisfecha, por ejemplo, por todas las secuencias ⟨Edimburgo, Londres, ...⟩, sean cuales sean su tercero y subsiguientes miembros. La oración abierta de 1 lugar "x es una ciudad" es satisfecha, por ejemplo, por todas las secuencias ⟨Edimburgo, ...⟩, sean cuales sean su segundo y subsiguientes miembros. Y la oración abierta (verdadera) de cero lugares "$(\exists x)(x$ es una ciudad)" es satisfecha por *todas* las secuencias ⟨..., ..., ...⟩, sean cuales sean su primer y subsiguientes miembros; pues hay una secuencia ⟨Edimburgo, ...⟩, por ejemplo, que difiere de cualquier secuencia arbitraria en a lo sumo el primer lugar y que satisface "x es una ciudad". Cualquier oración cerrada será satisfecha por *todas* las secuencias o por *ninguna*, y no puede ser satisfecha por algunas y no por otras. Consideremos un lenguaje bastante austero: el cálculo de predicados de primer orden sin términos singulares. En el caso más simple, una oración cerrada está formada por la cuantificación existencial de una oración abierta de 1 lugar. Tal oración existencialmente cuantificada es satisfecha por una secuencia arbitraria sólo si hay otra secuencia que difiera de ella en el primer lugar a lo sumo, y que satisfaga la oración abierta de 1 lugar que resulta de eliminar el cuantificador existencial inicial: y así, si la oración existencial es satisfecha por *cualquier* secuencia, será satisfecha por *toda* secuencia. Por tanto, una oración existencial cerrada será satisfecha o por todas las secuencias o por ninguna. La negación de una oración existencial cerrada, por la cláusula para la negación de la definición de satisfacción, será satisfecha por una secuencia sii la oración negada no es satisfecha por esa secuencia y así una vez más será satisfecha o por todas las secuencias o por ninguna: y similarmente para la conjunción de dos oraciones existenciales cerradas, la cual será satisfecha por una secuencia sii ambos miembros de la conjunción son satisfechos por esa secuencia y así será satisfecha también por todas las secuencias o por ninguna. Pero, ¿por qué se define "verdadera" como "satisfecha por todas las secuencias" y "falsa" como "no satisfecha por ninguna"? Pues bien, consideremos de nuevo la oración cerrada "$(\exists x)(x$ es una ciudad)": sea X una secuencia abierta de objetos. Por la cláusula de la definición de satisfacción que abarca las oraciones cuantificadas existencialmente, X satisface esta oración sii hay alguna secuencia Y que difiera de X en a lo sumo el primer lugar que satisface "x es una ciudad": ahora bien, un objeto, O, satisface "x es una ciudad" solamente en el caso de que O sea una ciudad, de modo que hay una tal secuencia solamente en el caso de que haya algún objeto que sea una ciudad. Así "$(\exists x)(x$ es una ciudad)" es satisfecha por todas las secuencias solamente en el caso de que algún objeto sea una ciudad. (Consúltese Rogers, 1963, para una ulterior discusión informal de la definición de Tarski).

Dos rasgos de la definición de Tarski merecen mención explícita en este punto. Primero, impone una *interpretación objetual* de los cuantificadores; como indica el ejemplo previo, "$(\exists x)Fx$" es verdadera sii algún objeto es *F*. Una interpretación sustitucional evitaría el rodeo vía satisfacción, pues permitiría definir la verdad de las oraciones cuantificadas directamente en términos de la verdad de sus instancias de sustitución (cfr. cap. 4, § 1). Segundo, en su escrito original, Tarski da una definición *absoluta* en vez de una definición en *términos de teoría de modelos*: "satisface", y, por tanto, "verdadera" se define con respecto a secuencias de objetos en el mundo real, no con respecto a secuencias de objetos en un modelo o "mundo posible" (por ejemplo, "hay una ciudad al norte de Birmingham" es verdadera absolutamente, pero falsa en un modelo en el que el dominio sea, pongamos por caso, {Londres, Exeter, Birmingham, Southampton}; cfr. págs. 137, 144s)[7].

Explicación formal

Tarski da su definición de la verdad para un cálculo de clases (el lenguaje objeto), y usa un metalenguaje formalizado. Yo daré, en cambio, una definición de la verdad para un lenguaje objeto más familiar, el cálculo de predicados de primer orden, y usaré el castellano más el lenguaje objeto (cfr. (b) (i), pág. 127) como metalenguaje. Esta definición de la verdad, sin embargo, seguirá en todo lo esencial a la de Tarski. (Y se atiene bastante fielmente a la explicación de Quine en 1970, cap. 3.)

Sintaxis de O

Las expresiones de O son:

variables: x_1, x_2, x_3 ... etc.
letras predicativas: *F*, *G*, ... etc. (cada una toma un número dado de argumentos)
conectivas oracionales: $-$, &
cuantificador: ($\exists$ ...)
paréntesis: (,)

[7] En 1957, Tarski y Vaught dan una definición en términos de teoría de modelos. La significatividad atribuida a la diferencia entre definiciones absolutas y definiciones en términos de teoría de modelos dependerá, en parte, de la actitud de uno ante los mundos posibles (véase págs. 202 y ss.). Quienes conciben el mundo real justamente como un mundo posible entre otros, considerarán la definición absoluta simplemente como un caso especial de una definición en términos de teoría de modelos. Sin embargo, no todos los autores consideran los dos enfoques de este tolerante modo. Queda abierta la cuestión de si una definición en términos de teoría de modelos satisface todas las restricciones usadas por Tarski en su artículo de 1931; y esto parece ser para algunos (Davidson, por ejemplo; véase más adelante) una razón importante para preferir la definición absoluta.

En términos de este vocabulario primitivo y austero pueden ser definidas, por supuesto, las otras funciones de verdad así como el cuantificador universal. Doy también por sentado que los términos singulares han sido eliminados. La ventaja de elegir un tal vocabulario mínimo es, como se verá, que reduce mucho el trabajo que conlleva la definición de la verdad.

Las oraciones atómicas de O son aquellas cadenas de expresiones que constan de un predicado de *n* lugares seguido por *n* variables.

(i) Todas las oraciones atómicas son fórmulas bien formadas (fbfs)
(ii) Si *A* es una fbf, $-A$ es una fbf
(iii) Si *A*, *B* son fbfs, $(A \& B)$ es una fbf
(iv) Si *A* es una fbf, $(\exists x)A$ es una fbf
(v) Ninguna otra cosa es una fbf

Definición de "satisface"

Tengan *X*, *Y* como rango las secuencias de objetos, tengan *A*, *B* como rango las oraciones de O, y denote X_i la *i*-ésima cosa en cualquier secuencia *X*.

Entonces la satisfacción puede definirse para las oraciones atómicas, dando una cláusula para cada predicado del lenguaje.

1. para predicados de 1 lugar:
   para todo *i*, *X*: *X* satisface "Fx_i" sii X_i es *F*
   para predicados de 2 lugares:
   para todo *i*, *X*: *X* satisface "Gx_ix_j" sii X_i y X_j se hallan en la relación *G*

y así sucesivamente para cada predicado.

2. para todo *X*, *A*: *X* satisface "$-A$" sii *X* no satisface "*A*"
3. para todo *X*, *A*, *B*: *X* satisface "*A* & *B*" sii *X* satisface *A* y *X* satisface *B*
4. para todo *X*, *A*, *i*: *X* satisface "$(\exists x)A$" sii hay una secuencia *Y* tal que $X_j = Y_j$ para todo $j \neq i$ e *Y* satisface "*A*"

(Obsérvese que cada cláusula de la definición de satisfacción corresponde a una cláusula en la definición de una fbf. Es por esto por lo que es tan conveniente trabajar con un vocabulario mínimo.) Una oración cerrada es una fbf sin variables libres; las oraciones cerradas serán satisfechas o por todas las secuencias o por ninguna.

Definición de "verdadera": Una oración cerrada de O es verdadera sii es satisfecha por todas las secuencias.

Tarski muestra que su definición es a la vez materialmente adecuada y formalmente correcta. Muestra también que a partir de su definición de la verdad se sigue que de cada par compuesto por una oración cerrada y su negación, una y solamente una es verdadera. Era de esperar en vista del hecho ya observado que la condición de adecuación material excluye las teorías de la verdad no bivalentes.

6 Comentario sobre la teoría semántica

La teoría de Tarski tiene la característica de haber sido criticada tanto por decir demasiado poco:

> la neutralidad de la definición[8] de Tarski con respecto a las teorías filosóficas de la verdad en liza es suficiente para demostrar su falta de relevancia filosófica. (Black, 1948, página 260.)

como por decir demasiado:

> la teoría de Tarski... pertenece al análisis fáctico más bien que al conceptual... La teoría de Tarski tiene mucha enjundia, mientras que el análisis conceptual correcto de la verdad tiene muy poca. (Mackie, 1973, pág. 40.)

La cuestión de la significación filosófica de la teoría de Tarski es evidentemente una cuestión difícil; yo la abordaré en tres etapas: me ocuparé en primer lugar de la propia estimación que Tarski hace de la significación de su teoría y luego trataré el uso que de la teoría han hecho dos autores —Popper y Davidson— que mantienen mayores esperanzas respecto de la misma que el propio Tarski.

a) *Estimación del propio Tarski*

Tarski expresa la esperanza (1944, págs. 53-4) de que su definición hará justicia a la concepción aristotélica de la verdad, pero ve poco aliciente en la cuestión de qué sea el concepto "correcto", y está dispuesto en realidad a usar la palabra "ferdadera" en lugar

[8] Aquí Black confunde aparentemente la condición de adecuación material con la definición, aunque en otro lugar del mismo artículo establece la distinción con bastante claridad.

de "verdadera" en el caso de que la decisión fuera contra él en este asunto (pág. 68).

Tarski es también modesto respecto a las pretensiones epistemológicas de su teoría; dice que no comprende realmente lo que podrá ser "el problema filosófico de la verdad" (pág. 70), pero de todos modos:

> podemos aceptar la concepción[9] semántica de verdad sin renunciar a cualquier actitud epistemológica que podamos haber tenido, podemos seguir siendo realistas ingenuos o idealistas, empiristas o metafísicos... La concepción semántica es completamente neutral respecto a todas estas cuestiones. (Pág. 71.)

Field sugiere (1972) que Tarski puede haber concedido importancia metafísica a la condición en la que insistía (pero cfr. pág. 130n) de que se defina la verdad sin el uso de primitivos semánticos: condición ésta que él justificaba (1931, págs. 152-3) apelando a la mayor claridad de las nociones sintácticas. Un comentario en otro artículo, "The establishment of scientific semantics", sugiere que puede haber pensado también en una importancia más profunda; después de repetir que el uso de primitivos semánticos amenazaría la claridad, prosigue:

> este método despertaría ciertas dudas desde un punto de vista filosófico general. Me parece, pues, que sería difícil armonizar este método con los postulados de la unidad de la ciencia y del fisicalismo (pues los conceptos de la semántica no serían ni conceptos lógicos ni conceptos físicos) (1936, pág. 406).

La conjetura de Field es que la intención de Tarski era poner la semántica en consonancia con las demandas del fisicalismo, la tesis de que no hay nada más que cuerpos físicos con sus propiedades y relaciones; y que esto ha de llevarse a cabo *definiendo* conceptos no físicos tales como verdad y satisfacción. Ello es confirmado en el pasaje, 1944, págs. 72-4, donde Tarski defiende la concepción semántica de la verdad frente a la crítica de que la semántica involucra elementos metafísicos, recalcando que su definición usa como primitivos sólo nombres lógicos, expresiones del lenguaje objeto, y nombres de esas expresiones. La otra cuestión de si la teoría de Tarski posee realmente esta significación es espinosa. Field cree que Tarski no tuvo éxito realmente al reducir la semántica a primi-

[9] El contexto sugiere que Tarski está aquí preocupado principalmente por la condición de adecuación material.

tivos "fisicalistamente" aceptables. Para las oraciones complejas abiertas, Tarski define la satisfacción recursivamente y para las oraciones atómicas abiertas en términos de la satisfacción, pero para las oraciones atómicas abiertas define la satisfacción *enumerativamente*, una cláusula para cada predicado primitivo del lenguaje objeto (como sería, "X satisface 'x_i es una ciudad' sii X_i es una ciudad, X satisface 'x_i está al norte de x_j' sii X_i está al norte de X_j..." y así sucesivamente). Puesto que Field sostiene que una reducción lograda requiere algo más que la equivalencia extensional del *definiens* y *definiendum*, que es todo lo que la definición de Tarski garantiza, encuentra que Tarski no ha vindicado, como esperaba, el fisicalismo. Parece digno de observar que existe una fuerte tendencia en los fisicalistas a ser extensionalistas y que hay cierta razón, por tanto, para suponer que Tarski habría considerado la equivalencia extensional como restricción suficiente. Queda la cuestión, por supuesto, de si la equivalencia extensional es realmente un requisito suficiente para las reducciones, o si, como sugiere Field, es conveniente un requisito más fuerte.

b) *Afirmaciones de Popper en favor de la teoría de Tarski*

> Popper saluda a la teoría de Tarski por haber rehabilitado la teoría de la correspondencia de la verdad absoluta u objetiva... Él vindicó el libre uso de la idea intuitiva de verdad como correspondencia con los hechos... (1960, página 224).

y usa las ideas de Tarski al elaborar su propia exposición del papel de la verdad como un ideal regulador de la investigación científica [10].

¿Es la teoría de Tarski una teoría de la correspondencia?

Según Popper, Tarski ha proporcionado justamente lo que faltaba en las teorías tradicionales de la correspondencia —un sentido preciso de "corresponde" (1960, pág. 223; 1972, pág. 320)—. Al menos, inicialmente esto resulta enigmático, pues Tarski comenta explícitamente que la teoría de la correspondencia es insatisfactoria (1944, pág. 54) y observa que "no se sorprendió en modo alguno" al conocer que en un sondeo llevado a cabo por Ness, sólo el 15 por 100 estaba de acuerdo en que la verdad es correspondencia con la realidad, mientras que el 90 por 100 estaba de acuerdo en que "está nevando" es verdadera si y sólo si está nevando (1944, pág. 70; y véase Ness, 1938).

[10] Esta sección amplía y modifica algunos puntos de Haack, 1976d; y cfr. Sellars, 1967, cap. 6, para alguna discusión al respecto.

Entonces, ¿qué es lo que lleva a Popper a pensar que Tarski ha vindicado la teoría de la correspondencia? Algunos comentarios (por ejemplo, 1960, pág. 224) sugieren que lo que él específicamente tiene en la mente es la insistencia de Tarski sobre la necesidad de un metalenguaje en el cual pueda uno referirse a expresiones del lenguaje objeto y decir lo que el lenguaje objeto dice. Es como si en cada instancia del esquema (T) tal como:

"La nieve es blanca" es verdadera sii la nieve es blanca

concibiera que la parte izquierda se refiere al lenguaje y la parte derecha a los hechos. Pero esto parece una razón bastante inadecuada para considerar la teoría de Tarski como una teoría de la correspondencia, pues la condición de adecuación material, aunque su papel consiste en excluir algunas definiciones, ciertamente no caracteriza a la teoría de la correspondencia como la única correcta: permite presumiblemente, por ejemplo, una definición de la redundancia tal como la de Mackie:

(p) (el enunciado de que p es verdadera sii p)

Es precisamente por esta razón por la que el mismo Tarski recalca la neutralidad epistemológica del esquema (T).

Sin embargo, aunque Popper no se refiere explícitamente a ellos, hay rasgos de la *definición* de la verdad de Tarski que evocan las teorías de la correspondencia. Existe aquí la dificultad de que no está muy claro qué se requiere para considerar realmente a la definición de Tarski como una versión de la teoría de la correspondencia; y la dificultad se agrava por la insistencia de Popper en que hasta Tarski no había habido ninguna teoría de la correspondencia genuina ni satisfactoria. Con todo, podemos conseguir algún progreso comparando la definición de Tarski primero con la versión del atomismo lógico ofrecida por Russell y Wittgenstein, y luego con la versión de Austin.

Tarski define la verdad en términos de satisfacción, y la satisfacción es una relación entre oraciones abiertas y secuencias de objetos; la explicación de satisfacción guarda cierta analogía con la concepción de la verdad de Wittgenstein que consiste en la correspondencia entre la disposición de los nombres en una proposición y la disposición de los objetos en el mundo. Por otra parte, la definición de la *verdad* de Tarski no apela a secuencias específicas de objetos, ya que las oraciones verdaderas son satisfechas por todas las secuencias, y las oraciones falsas por ninguna. Es sintomático que la definición de Tarski englobe tanto la verdad lógica como la fáctica; suponer que la verdad lógica consiste en la correspondencia con los hechos es seguramente menos plausible que suponer que la verdad "fáctica" consiste en la correspondencia con los hechos. Dos observaciones históricas parecen necesarias aquí: primera, que Wittgenstein pensaba que las fbfs cuantificadas podían considerarse como conjunciones/disyunciones de proposiciones atómicas, y si ello fuera realmente así, el rodeo de Tarski a través de la satisfacción sería innecesario; y, segunda, que Russell admitió "hechos lógicos".

El desarrollo que hace Tarski de la estructura de las oraciones en la definición recursiva de satisfacción es, por tanto, una semejanza de los comentarios que Russell y Wittgenstein hacen sobre "corresponde". Es igualmente una *de*semejanza de la explicación de Austin. Austin insiste en que los enunciados, no las oraciones, son los portadores primarios de verdad. Esto tiene al menos dos consecuencias relevantes: Tarski ignora los problemas planteados por oraciones que tienen palabras índice, como "yo" y "ahora", sobre los que se centra Austin; y mientras que la definición de satisfacción de Tarski tiene en cuenta la estructura sintáctica de las oraciones abiertas, la explicación de correspondencia de Austin acentúa su carácter arbitrario puramente convencional —en otro lenguaje, el enunciado de una tontería, dice él, sería verdadero precisamente en la circunstancia de que el enunciado en inglés de que los Liberales Nacionales son la elección del pueblo fuese verdadero[11]. Hay, sin embargo, un punto de analogía que es digno de mención. La explicación de Austin, sugerí antes, evita localizar la correspondencia en la conexión demasiado estrecha entre el enunciado de que p y el hecho de que p, explicándola más bien como consistente en la situación a la que se refiere el enunciado de que p que es del tipo del que el enunciado dice que es. Se puede ver aquí, sin demasiado esfuerzo, una semejanza con la explicación enumerativa de satisfacción dada por Tarski para las oraciones atómicas abiertas: por ejemplo, X satisface "x_i es blanco" sii la i-ésima cosa de la secuencia X es blanca.

Así: Tarski no se estima a sí mismo como suministrador de una versión de la teoría de la correspondencia y su condición de adecuación material es neutral entre la correspondencia y las otras definiciones. Sin embargo, la definición, si no la de la verdad, sí la de satisfacción de Tarski, conlleva cierta analogía con las teorías de la correspondencia: las cláusulas para las oraciones atómicas abiertas con la versión de Austin, las cláusulas para las oraciones moleculares abiertas con las versiones de Russell y Wittgenstein.

[11] Aquí, por tanto, se da un caso en el que el problema de los portadores de verdad adquiere una importancia real. (No me resisto a la tentación de llamar la atención sobre la queja de Austin (1950, pág. 30) de que su compañero de simposio, Strawson, no había acertado al hacer la distinción crucial entre oración y enunciado.) Me ocuparé de la cuestión de cómo la teoría de Tarski puede adaptarse para tratar las oraciones indexadas en el apartado sobre Davidson.

¿Es la teoría de Tarski "absoluta" y "objetiva"?

Se estimen o no las afinidades lo suficientemente fuertes como para considerar la teoría de Tarski como una versión de la teoría de la correspondencia, vale la pena preguntar si la definición semántica de la verdad posee, en cualquier caso, lo que Popper considera que son las mayores virtudes de la teoría de la correspondencia, su carácter "absoluto" y "objetivo".

Tarski subraya que la verdad puede definirse sólo *relativamente a un lenguaje* —lo que él define no es "verdadera" (punto), sino "verdadera-en-O"—. Esto es por dos razones: que la definición debe aplicarse a oraciones (que, a diferencia de "ítems" extralingüísticos tales como proposiciones, tienen la estructura sintáctica que él desarrolla) y una y la misma oración puede ser verdadera en un lenguaje y falsa o carente de significado en otro; y que sólo una jerarquía del lenguaje objeto, del metalenguaje y del metametalenguaje puede evitar las paradojas semánticas. En este sentido, por tanto, la definición de la verdad de Tarski no es una definición absoluta, sino relativa. Popper, sin embargo, que tiende a tomar una actitud un tanto indiferente respecto a la cuestión de los portadores de verdad (1972, págs. 11, 45, 319n), no se preocupa de este sentido de "absoluto". Tampoco muestra ningún interés respecto al hecho de que la definición original de Tarski sea absoluta más bien que en términos de teoría de modelos.

Sin embargo, parece que Popper considera equivalentes "absoluto" y "objetivo" contrastándolos con "subjetivo", esto es, "relativo a nuestro conocimiento o creencia". A este respecto, Popper piensa que la teoria de la correspondencia es superior a

> la teoría de la coherencia... [la cual] confunde consistencia con verdad, la teoría de la evidencia... [la cual] confunde "conocido como verdadero" con "verdadero" y la teoría pragmatista o instrumentalista [la cual] confunde utilidad con verdad. (1960, pág. 225)

No necesito comentar, creo yo, la exactitud de la caracterización que Popper hace de las teorías rivales; de todos modos, el meollo de su argumento afortunadamente no depende de estos detalles. Las teorías rivales, arguye Popper, están fundadas en el "extendido, pero equivocado dogma de que una teoría satisfactoria debería producir un criterio de creencia verdadera" (1960, pág. 225). Y una teoría criterial de la verdad es subjetiva debido a que no puede admitir la posibilidad de que una proposición sea verdadera incluso aunque ninguno la crea, o falsa incluso aunque todos la crean.

¿Qué encuentra exactamente Popper objetable acerca de las teorías criteriales de la verdad? Popper no lo dice muy claro; pero yo pienso que el problema puede ser encauzado. La dificultad crucial estriba no en el intento de suministrar un criterio de verdad en sí mismo, sino en la adopción de una teoría criterial del significado de "verdadera". (Su actitud aparece quizás más clara en el apéndice a la edición de 1961 del vol. 2 de *The Open Society and its Enemies*.) Si damos el significado de "verdadera" en términos de nuestros criterios de verdad, no dejamos lugar para la posibilidad de que una proposición sea falsa aunque pase nuestros tests de verdad, o verdadera aunque los suspenda. Este es un problema particular de los pragmatistas, puesto que plantea una amenaza a su falibilismo oficial; aunque todavía queda lugar para cometer equivocaciones en la *aplicación* incluso de test infalibles de verdad. El infalibilismo en *sí mismo* no es subjetivista; pero la otra afirmación de que decir que una proposición es verdadera (falsa) *significa justate* que pasa (suspende) nuestros tests plantea una amenaza al objetivismo.

Tarski renuncia expresamente a la aspiración de suministrar un criterio de verdad (1944, págs. 71-2); y ciertamente su definición no hace ninguna referencia a nuestros tests de la verdad. (Irónicamente, el pasaje en el que Tarski llama la atención sobre estos aspectos es considerado como una refutación de la "objeción" de que su teoría es una clase de teoría de la correspondencia que involucra la lógica en "el más acrítico realismo".)

Por tanto, la teoría de Tarski es una teoría objetiva en el sentido de Popper. Pero, ¿por qué concede Popper tanta importancia a este punto? La explicación depende del uso epistemológico al cual él se propone someter el concepto de verdad.

La verdad como ideal regulador: verosimilitud

Popper se describe a sí mismo como un "falibilista absolutista": falibilista porque niega que dispongamos de método alguno garantizado de adquirir conocimiento; absolutista porque insiste en que hay una cosa tal como la verdad objetiva a que aspira la investigación científica. La teoría de Tarski ha de suministrar una explicación adecuadamente objetiva de este "ideal regulador" de la ciencia.

Esto requiere, por supuesto, que la teoría de Tarski sea aplicable a los lenguajes —presumiblemente, fragmentos más o menos completa y rígidamente reglamentados del lenguaje natural y matemático en los que se expresan las teorías científicas. No discutiré aquí las cuestiones planteadas por este requisito; en parte, debido a que el mismo Tarski manifiesta (1944, pág. 74) un optimismo cauteloso respecto a las posibilidades de aplicación de su trabajo a las ciencias empíricas y, en parte, debido a que en el próximo apartado, cuando trate el uso que hace Davidson del trabajo de Tarski, tendré que considerar las razones que Tarski da para dudar de si sus métodos se aplican al lenguaje "coloquial".

Según Popper, el quehacer de la ciencia consiste en idear y contrastar conjeturas; los científicos no pueden estar seguros de que sus conjeturas corrientes sean verdaderas, ni tampoco de que alcanzarán alguna vez la verdad y ni siquiera pueden estar seguros de saber, si han alcanzado la verdad, que la han alcanzado. Pero si la verdad no ha de ser solamente un ideal, sino un ideal *que guíe* o "regulador", debería ser posible saber, cuando una teoría reemplaza a otra, si la ciencia se está acercando a la verdad. Así pues, el problema de Popper es explicar entre dos teorías que pueden ser las dos falsas en qué sentido puede la una estar más cerca de la verdad que la otra. Su solución consiste en la extensión que Popper hace de las ideas de Tarski en la teoría de la "verosimilitud" o semejanza de la verdad.

La explicación que da Popper de la verosimilitud es la siguiente:

> *Suponiendo que el contenido de verdad y el contenido de falsedad de dos teorías t_1 y t_2 sean comparables, podemos decir que t_2 es más exactamente similar a la verdad... que t_1 si y sólo si o bien:*
> (a) *el contenido de verdad, pero no el contenido de falsedad de t_2 excede al de t_1*
> [o bien]
> (b) *el contenido de falsedad de t_1, pero no su contenido de de verdad excede al de t_2*. (1963, pág. 233).

El contenido de verdad (falsedad) de una teoría es la clase de todas sus consecuencias verdaderas (falsas) y sólo ellas. El contenido de verdad o de falsedad de una teoría puede exceder al contenido de verdad o de falsedad de otra sólo si su contenido de verdad o de falsedad incluye el de la otra en términos de teoría de conjuntos, de modo que esta explicación se aplica solamente a las teorías que se solapan de esta forma. Popper sugiere también (1963, págs. 393-6; 1972, págs. 51, 354) medidas de contenidos de verdad y de falsedad en términos de probabilidad lógica, de modo que se podrían comparar dos contenidos cualesquiera. Pero me centraré en la versión primera, "la cualitativa", más bien que en la segunda, "la cuantitativa".

La definición de verosimilitud no puede mostrar que la ciencia progresa hacia la verdad: pero Popper espera (1972, pág. 53) que apoye a su metodología falsacionista, que recomienda elegir la conjetura más falsable, la de más contenido, porque una teoría con más contenido tendrá mayor verosimilitud, a menos que, añade Popper, tenga más contenido de falsedad así como más contenido de verdad.

Sin embargo, se ha mostrado[12] que una teoría t_2 tiene mayor verosimilitud que otra, t_1, de acuerdo con las cláusulas (a) y (b) de Popper solamente si t_2 es una teoría *verdadera* de la que se sigue el contenido de verdad de t_1. Esto significa que *la definición de verosimilitud de Popper no se aplica a comparaciones entre teorías que entrambas son falsas;* pero ese, por supuesto, era el principal objetivo de la teoría, que, por tanto, falla en su propósito epistemológico. Este fracaso es, creo, importante para la vuestión de la viabilidad del absolutismo falibilista (véase Haack, 1977b); y, a mi entender, se debería también apoyar la valoración más bien modesta que hace Tarski de la significación epistemológica de la teoría semántica de la verdad en contra de la de Popper que es más ambiciosa.

c) *El uso que hace Davidson de la teoría de Tarski*

Verdad y significado. Davidson piensa que cualquier teoría adecuada del significado debe explicar cómo los significados de las oraciones dependen de los significados de las palabras (de lo contrario, argumenta él, no se podría aprender el lenguaje). Una teoría del significado debe ser consistente con —o explicar como él dice a veces— "la productividad semántica": la capacidad que los hablantes tienen de producir y comprender oraciones que nunca han oído antes. Lo que esto significa, afirma él, es que la teoría debería producir todas las oraciones de la forma:

S significa *m*

donde "*S*" es una descripción reveladora de la estructura de una oración del lenguaje para el que se da la teoría, y "*m*" es un término que denota el significado de esa oración. Pero la apelación aquí

[12] Miller, 1974; y cfr. Tichy, 1974, y Harris, 1974. Muy brevemente, la estrategia de Miller consiste primero en mostrar, si t_1 y t_2 son comparables por el contenido de verdad, de qué modo son también comparables por el contenido de falsedad; y después en mostrar que para que t_2 esté más cerca de la verdad que t_1, t_2 debe ser una teoría verdadera de la que se sigue el contenido de verdad de t_1, ya que de otro modo t_2 excederá a t_1 en contenido de falsedad así como en contenido de verdad, de modo que sus verosimilitudes no serán comparables.

a significados implícitos, sugiere Davidson, no contribuye a nada útil; y reformulando el requisito así:

S significa que *p*

donde "*p*" es una oración que tiene el significado de la oración descrita por "*S*", deja un problema con el "significa que", el cual, por tanto, reformula Davidson como "es *T* sii" donde "*T*" es cualquier predicado arbitrario que, dadas las condiciones anteriores para "*S*" y "*p*", satisface:

S es *T* sii *p*

Pero, por supuesto, cualquier predicado que satisfaga esta condición será, en virtud de los estándars de Tarski, un predicado *de verdad* materialmente adecuado. Davidson concluye que lo que es requerido por una teoría del significado es precisamente la definición de un tal predicado *de verdad* (Davidson, 1967).

Significado como condiciones de verdad

Aunque el camino por el cual Davidson llega a esta conclusión es un tanto indirecto, el resultado —que el significado de una oración puede ser dado especificando sus condiciones de verdad— no es desconocido. Lo que es nuevo en la versión de Davidson es la imposibilidad de las restricciones "tarskianas" en la explicación de las condiciones de verdad[13].

El atractivo de una teoría del significado en términos de condiciones de verdad puede quizás ser valorado recordando la clasificación hecha por Quine de las nociones semánticas en dos grupos, el extensional, que él considera que constituye el objeto de la "teoría de la referencia", y el intensional, que él considera que constituye el objeto de la "teoría del significado", así:

ideas semánticas

teoría de la referencia (extensional)	*teoría del significado* (intensional)
por ejemplo, "designa"	por ejemplo, "es significativo"
"satisface"	"es sinónimo de"
"es verdadero"	"analítico"
	"significa"

[13] Dummett insiste en una teoría del significado en términos más bien de condiciones de asertabilidad que de condiciones de verdad (también una comparación con los pragmatistas, ahora con su teoría criterial del significado, sugiere lo mismo). Para discusiones críticas véase Haack, 1974, págs. 103 y ss., y cfr. Brandom, 1976.

Quine argumentaba en 1953a que la teoría de la referencia estaba considerablemente mejor configurada que la teoría del significado. Un rasgo atractivo de la teoría de condiciones de verdad es que augura una explicación del significado (desde el más problemático lado derecho) en términos de verdad (desde el menos problemático lado izquierdo).

Teoría de la interpretación

Ulteriormente (1974) Davidson añade una nueva teoría, la de la interpretación del discurso de otro en otro lenguaje o incluso en el mismo que el propio de uno; esencialmente, ello consiste en una explicación de cómo saber cuando "*p*" es una oración que tiene el significado que "*S*" describe. Brevemente, la idea es que para comprobar empíricamente si una oración de la forma

"*Es regnet*" es verdadera sii está lloviendo

es una oración-T, es decir, cumple la especificación de Tarski de que la oración de la derecha traduce la oración nombrada en la parte izquierda, hay que comprobar si los hablantes del lenguaje en cuestión (en este caso el alemán) consideran verdadera "*Es regnet*" sii está lloviendo. El punto de interés respecto a lo que los hablantes nativos *consideran como verdadero* está en alcanzar el significado de sus emisiones mediante, por decirlo así, la comprensión de sus creencias constantes. En consecuencia, se requiere una suposición, el principio de caridad, a efectos de que los hablantes de otros lenguajes estén de acuerdo generalmente con nosotros acerca de lo que es el caso. El carácter holista de la explicación de Davidson, su insistencia en que la "unidad de interpretación" es el lenguaje completo, puede derivar al menos en parte del holismo epistémico, la idea "duhemiana", subrayada también por Quine, de que las creencias son verificadas/falsificadas no aisladamente, sino conjuntamente.

Aunque hay muchas preguntas importantes que hacer sobre esta teoría de la interpretación, me centraré ahora en la explicación del significado de Davidson ya que es allí donde la teoría de la verdad de Tarski juega un papel crucial.

Si la tarea de una teoría del significado es realmente, como piensa Davidson, definir un predicado de verdad "tarskiano", ¿qué trabajo sería necesario hacer además del ya realizado por Tarski? Davidson busca una teoría del significado *para lenguajes naturales*, tales como el inglés; Tarski, por supuesto, es totalmente escéptico acerca de la aplicabilidad de su teoría a los lenguajes naturales. Así pues, si el programa de Davidson es factible, la primera tarea

es mostrar que los métodos de Tarski *pueden* ser ampliados. Ésta es una cuestión importante incluso independientemente de las especiales ambiciones de Davidson referente a los métodos de Tarski, pues el concepto de verdad tiene significación filosófica en muchos contextos en los que debe admitirse que "verdadera" se aplique a oraciones de los lenguajes naturales —por ejemplo, en epistemología—. A pesar de la modestia oficial de Tarski a este respecto, me parece que la utilidad de su trabajo quedaría lamentablemente restringida si el concepto que él define resultara ser completamente diferente del concepto de verdad de los lenguajes naturales.

¿Es aplicable la teoría de Tarski a los lenguajes naturales?

Según Tarski:

> *La posibilidad misma de un uso consistente de la expresión "oración verdadera" que esté en armonía con las leyes de la lógica y el espíritu del lenguaje ordinario parece ser muy cuestionable y consecuentemente la misma duda afecta a la posibilidad de construir una definición correcta de esta expresión.* (1931, pág. 165)

El pesimismo de Tarski tiene dos fuentes principales: su condición de corrección formal excluye la posibilidad de una definición adecuada de la verdad para lenguajes que no son ni (i) semánticamente abiertos ni (ii) formalmente especificables. Los lenguajes naturales, argumenta Tarski, fallan en ambos puntos de modo que no hay ninguna perspectiva de una definición adecuada de la verdad para ellos.

(i) Tarski sugiere que los lenguajes naturales contienen sus propios metalenguajes de modo que la verdad no puede definirse sin topar con la paradoja; aunque a veces él da más bien a entender que, debido a que los lenguajes naturales no son formalmente especificables, no se puede responder a la cuestión de su clausura semántica. Davidson no tiene una respuesta muy satisfactoria para este problema, pero dice que "no es lícito proseguir sin haber desinfectado esta fuente de ansiedad conceptual" (1967, pág. 10). Parece proponer que el trabajo prosiga en aquellos fragmentos semánticamente abiertos de los lenguajes naturales donde no aparece el peligro de la paradoja. Existe cierta dificultad en compaginar la actitud que Davidson mantiene ante las paradojas (no se preocupa demasiado de ellas, se centra en el resto de la tarea) con su holismo, la insistencia en que una teoría adecuada del significado debe ser una teoría para todo el lenguaje; aunque también deja ver que él tiene dudas sobre si los lenguajes naturales son realmente universales.

(ii) Parece haber aquí toda una familia de dificultades; el problema de dar una explicación cabal exactamente de qué cadenas se consideran como oraciones de un lenguaje natural es agravado por el hecho de que los lenguajes naturales no están estáticos, sino en crecimiento; y la frecuencia en los lenguajes naturales de fenómenos tales como vaguedad, ambigüedad, indexicalidad. Tarski es pesimista:

> Quienquiera que desee, a pesar de todas las dificultades, dedicarse a la semántica del lenguaje coloquial con la ayuda de métodos exactos se verá obligado primero a emprender la ingrata tarea de una reforma de este lenguaje... Sin embargo, puede ponerse en duda si el lenguaje de la vida cotidiana, después de ser "racionalizado" de este modo, conservaría todavía su naturalidad y si no asumiría más bien los rasgos característicos de los lenguajes formalizados. (1931, pág. 267)

El núcleo de la réplica de Davidson a esto es que, aunque será necesario cierto "ordenamiento" antes de que puedan aplicarse los métodos de Tarski a un lenguaje natural, esta necesidad no será tal como para transformarlo más allá de todo reconocimiento. Sostendría, pienso, que los logros de la gramática transformacional (véase, por ejemplo, Chomsky, 1957) augurarían vencer el primer problema; y es optimista en el sentido de que más fragmentos de lenguajes naturales podrán ser tratados dentro del ámbito de los métodos "tarskianos", semejantemente a como el trabajo de Frege sobre "(x)" y "$(\exists x)$" ya ha reglamentado rígida y adecuadamente a "todo", "ninguno" y "alguno"[14].

Lo que Tarski considera como una "tarea ingrata", Davidson la acomete de buen grado observando que "es bueno saber que no nos cansaremos de la tarea". Su principal tarea, de hecho, es proporcionar un análisis adecuado de aquellas locuciones de los lenguajes naturales que son inicialmente recalcitrantes al tratamiento "tarskiano". Y es en el éxito o fracaso de esta tarea donde debe basarse la valoración de la respuesta de Davidson al escepticismo de Tarski. Es digno de observarse que Davidson insiste en usar el concepto "absoluto" de verdad en lugar del concepto de verdad en términos de teoría de modelos; y que algunos de estos problemas (por ejemplo, los problemas creados por la introducción de nuevos predicados a medida que se desarrolla el lenguaje natural) resultan más difíciles desde un enfoque absoluto de lo que lo habrían sido desde un enfoque en términos de teoría de modelos; cfr. Field, 1972.

[14] Es discutible que la explicación de Frege reglamente rígida y propiamente los cuantificadores del lenguaje natural; recuérdese (cap. 4, § 1) que Montague y Hintikka, como el primer Russell, subrayan sus afinidades con los términos singulares, mientras que, según Frege, pertenecen a una categoría sintáctica enteramente diferente.

Forma lógica

Davidson se describe a sí mismo como buscador de "la forma lógica" de las locuciones del lenguaje natural. Por ejemplo, recuérdese (cap. 2, § 4) que, según Davidson, las construcciones adverbiales del lenguaje natural están mejor representadas involucrando la cuantificación sobre eventos con adverbios interpretados como adjetivos de términos de eventos. La forma lógica de "John untó de mantequilla la tostada con un cuchillo", afirma Davidson, es algo así como "Hay un evento que es el hecho de untar mantequilla la tostada por John y que es realizado con un cuchillo". La seguridad de Davidson de que cada construcción del lenguaje natural posee una única forma lógica brota de la creencia de que una representación formal a la que se aplica el método de definición de la verdad de Tarski representa la estructura esencial de una forma idealmente perspicua. (Es notable la analogía con el proyecto de Russell y Wittgenstein en su período de atomistas lógicos de idear un lenguaje ideal que representase la forma *real* de los lenguajes naturales.) De manera interesante, Cargile se ha preguntado (1970; y cfr. la réplica de Davidson en el mismo volumen) por qué la conexión entre un predicado y su forma modificada adverbialmente debe necesariamente ser considerada como una cuestión de forma más bien que de contenido. No está, sugiere él, tan claro como Davidson debe suponer, qué es lo que debe considerarse como esqueleto y qué como carne; él solicita de hecho una concepción más flexible de la forma lógica, más próxima a la presentada en el capítulo 2.

El programa de Davidson

Así pues, Davidson considera que la tarea de la teoría del significado no es suministrar una explicación del significado de las palabras individuales, sino analizar la estructura de las oraciones. (Esto no es *totalmente* correcto, porque algunas partículas —"in", por ejemplo— tienen un carácter estructural.) Por ejemplo, Davidson no requiere una teoría del significado para dar el significado de "bueno", pero la requiere para analizar la estructura, verbigracia, de "Bardot es una buena actriz" al explicar de tal modo por qué no es equivalente a "Bardot es buena y Bardot es una actriz" mientras que "Bardot es una actriz francesa" sí que es equivalente a "Bardot es francesa y Bardot es una actriz" (cfr. "pequeño elefante", y el ambiguo "pobre violinista"). El atractivo del método de Tarski, que consiste en definir la satisfacción para las oraciones complejas abiertas en términos de la satisfacción de oraciones simples abiertas, radica en su promesa de explicar cómo los significados de oraciones compuestas dependen del significado de sus partes; el reto consiste en analizar oraciones tales como "Bardot es una buena actriz" de manera que el método de Tarski se aplique a ellas lo mismo que a las menos recalcitrantes tales como "Bardot es una actriz francesa". Davidson admite que el trabajo es considerable ya que:

> queda una asombrosa lista de dificultades y enigmas. (1967, pág. 321)

Él incluye ("por nombrar algunos") contrafácticos, subjuntivos, enunciados de probabilidad, enunciados causales, adverbios, adjetivos atributivos, términos de masa, verbos de creencia, percepción, intención, acción. Obviamente mi estudio de los detalles del programa de Davidson tendrá que ser selectivo.

Indices

La teoría de Tarski necesita ser relativizada a hablantes y tiempos, sugiere Davidson, porque los lenguajes naturales contienen índices. El esquema (T) revisado exigirá que la teoría entrañe oraciones como:

"yo estoy cansado" (s, t) es verdadera sii s está cansado en t

La verdad, dice Davidson, es un predicado más bien de elocuciones que de oraciones. (Esta sugerencia es relevante para la tesis discutida por Strawson y, antes que él, por Schiller de que los métodos formales son inherentemente inadecuados para tratar la dependencia del contexto de los enunciados en los lenguajes naturales.)

Pero el estudio que Davidson hace de los índices (indexicals) va dirigido también hacia los problemas planteados por el análisis de citas y verbos de "actitud proposicional" ("dice que", "sabe que", etc.); pues él piensa que todas estas construcciones involucran demostrativos encubiertos. Un análisis de estos índices ("esto", "eso") realizado por Weinstein en 1974 ha sido ratificado por Davidson. Según esta explicación, "eso es un gato", por ejemplo, es verdadera sólo en el caso de que el objeto indicado por el hablante en el momento de la elocución satisfaga "... es un gato".

Oratio obliqua

Mientras los compuestos veritativo-funcionales no plantean ningún problema, habrá obviamente una dificultad referente a la aplicación de los métodos de Tarski a oraciones compuestas del castellano cuyos valores de verdad no dependen de ningún modo obvio de los valores de verdad de sus partes. Las oraciones de la clase *oratio obliqua* pertenecen a este tipo intensional problemático; pues el valor de verdad de "Galileo dijo que la tierra se mueve", por ejemplo, no depende en modo alguno directo del valor de verdad de "la tierra se mueve"; y hay fallo de sustitutividad, pues de "Tom dijo que la luna es redonda" y "La luna = el único planeta de la tierra" no se puede con seguridad inferir "Tom dijo que el único planeta de la tierra es redondo".

El primer paso en la dirección correcta, insiste Davidson, es analizar:

Galileo dijo que la tierra se mueve

a tenor de:

Galileo dijo que.
La tierra se mueve.

El "que" debe interpretarse no como un pronombre relativo, sino como un pronombre demostrativo que se refiere a una elocución —como yo podría decir "he escrito que", señalando a un mensaje en el tablón de anuncios—[15]. Por supuesto, Galileo no profirió la misma elocución que genera el hablante; realmente Galileo no hablaba castellano; por tanto, se requieren más explicaciones. Davidson amplía su análisis de esta manera:

La tierra se mueve.
$(\exists x)$ (la elocución de Galileo x y mi última elocución hacen que estemos diciendo lo mismo)

[15] Merece la pena hacer dos puntos de comparación. He mencionado ya las afinidades entre la definición de satisfacción de Tarski y la explicación de verdad de Wittgenstein en el *Tractatus*; los verbos de actitud proposicional, que presentan un problema para la aproximación de Wittgenstein y la de Davidson, se tratan en 5.542. El análisis de Wittgenstein resulta, sin embargo, notoriamente oscuro. Alun Jones me ha indicado que la lista de "dificultades y enigmas" que da Davidson para *su* empresa y la lista de problemas de Anscombe (1959) para la empresa de Wittgenstein son muy similares. Kotarbinski (1955) ha sugerido un análisis del discurso indirecto notablemente semejante al de Davidson. El objetivo de Kotarbinski era respaldar la tesis de que sólo existen cuerpos materiales ("pansomatismo") mediante el análisis de las referencias aparentes de objetos abstractos tales como proposiciones; esto puede ser significativo en vista de la conjetura de que Tarski estaba motivado por la simpatía con el materialismo.

Galileo y yo estamos diciendo lo mismo, hemos dicho precisamente en el caso en que él profirió una oración que significa en su boca lo que alguna elocución mía significa en mi boca.

La aplicación de los métodos de Tarski, tal como han sido ampliados por Weinstein para enfrentarse con los índices, dan un resultado a tenor de:

> "Galileo dijo que la tierra se mueve" (*s*, *t*) significa que es verdadera sii
>
> Galileo profirió en *t'* (*t'* anterior a *t*) una oración que significa en su boca lo que la elocución proferida por *s* en *t''* (*t''* justo detrás de *t*) significa en la boca de *s*, donde la elocución proferida es "La tierra se mueve".

Puede ser útil detenerse brevemente en contrastar la explicación de Davidson con algunas explicaciones alternativas de las (denominadas) actitudes "proposicionales". Frege, por ejemplo, consideraría "que *p*" en "*s* dijo (cree) que *p*" como refiriéndose a una proposición (véase cap. 5, § 2). Carnap analizaría "*s* dijo (cree) que *p*" como "*s* profirió (está dispuesto a asentir a) alguna oración intensionalmente isomórfica de '*p*' en castellano" (véase 1947). Scheffler trata a "que *p*" más bien como un adjetivo que como un nombre: "*s* dijo (cree) que *p*" equivale a "*s* profirió (cree) una emisión de-que-*p*" donde hay un predicado separado correspondiente a cada oración "*p*" (véase 1954); Quine va aún más lejos en la misma dirección, tratando el todo de "dijo (cree)-que-*p*" como un predicado de *s* (véase 1960a, § 44).

Davidson cree que su explicación tiene las ventajas siguientes: a diferencia de los análisis que consideran a "que *p*" como refiriéndose a una proposición, ella no requiere apelar a entidades intensionales; a diferencia del análisis de Carnap, ella no requiere referencia explícita a un lenguaje; y a diferencia de los análisis que tratan a "dice (cree)-que-*p*" como un único predicado, ella concede que lo que sigue al "que" es una oración con "estructura significativa", estructura que puede ser desarrollada por una teoría del significado.

Su explicación, argumenta Davidson, admite, como parece más propio, que "*s* dijo que *p*" entraña "*s* dijo algo", ya que el análisis nos da "*s* profirió una oración que...". Al mismo tiempo, ello explica, como también se requiere, por qué "*s* dijo que *p*" no entraña "*p*", pues lo que parecía ser una oración ("*p*") dentro del alcance de un operador oracional ("*s* dijo que") se convierte en una única oración ("*s* dijo que") que contiene un demostrativo ("que") el cual se refiere a una elocución de otra oración ("*p*"). Y así como, aunque los gatos arañan, la oración "eso es un gato", que se refiere a un

gato, no araña, así, aunque "*p*" entraña "*p*", la oración "*s* dijo que *p*", que se refiere a una elocución de "*p*", no entraña "*p*".

Desde luego, como el último ejemplo evidencia, en los casos regulares considerados por Weinstein, a lo que se refiere "que" es un ítem no lingüístico, un gato, por ejemplo. Cuando la explicación se extiende a "que" del habla indirecta, los referentes serán elocuciones de oraciones, y estas oraciones tienen estructura significativa (entre las instancias de "*s* dijo que *p*" estarían, por ejemplo, "*s* dijo que *q* y *r*" y "*s* dijo que *s'* dijo que *q'*") en virtud de cuya estructura viene dado su significado.

Haremos aquí algunos comentarios acerca de cómo el análisis de Davidson difiere del de Carnap: en la explicación de Davidson "*s* dijo que *p*" involucra referencia a una elocución del hablante relacionada con alguna elocución de *s* por *la dicción de lo mismo*; en la de Carnap, involucra referencia a una oración relacionada con una oración del castellano por *isomorfismo intensional*. Una elocución (aquí Davidson pone de manifiesto que se refiere a un acto de habla, al evento de proferir una oración) es una elocución de alguna oración en un lenguaje específico con un contexto específico; y, por tanto, se evita la necesidad de especificar el lenguaje relevante[16].

Esto confiere a la explicación de Davidson un carácter inesperado porque el concepto de elocución (acto de habla) pertenece más bien a la pragmática que a la semántica. Es igualmente sorprendente, y metodológicamente también inquietante, que la explicación de Davidson, como la de Carnap, requiera un primitivo semántico (respectivamente, "la dicción de lo mismo" y el isomorfismo intensional) en el metalenguaje. *s* y *s'* dicen lo mismo, explica Davidson, sólo en el caso de que alguna elocución de *s signifique lo mismo que* alguna elocución de *s'*. Ahora bien, Davidson insiste en que las condiciones de verdad se den en términos de una definición de verdad absoluta, es decir, una definición que no use primitivos semánticos. Y él evita "*S* significa *m*" y la fórmula "*S* significa que *p*" debido a su carácter intensional. Aparentemente Davidson considera como admisible el recurso a decir lo mismo porque es *local*: la explicación general del significado apela solamente a las condiciones de verdad tarskianas, aunque la explicación específica de "dice que" requiere el decir lo mismo como primitivo semántico. Es cuestionable, empero, si el recurso *es* local en el sentido relevante: pues seguramente "dice que" se considera como estructura más

[16] Davidson habla a veces como si fuera la referencia a una elocución (más bien que a una oración) lo que impide que "*s* dijo que *p*" entrañe "*p*". Pero esto queda con seguridad suficientemente explicado apelando al hecho de que (una elocución de) "*p*" *se refiere a por*, y no *está contenida en* "*r* dijo que *p*". El sentido de "elocución" en donde, según Davidson, la verdad es una propiedad de las elocuciones, presumiblemente tiene que ser el sentido del "contenido" y no, como en este contexto, el sentido del "acto".

que como vocabulario en el sentido en el que la dependencia del significado de "bueno" en cuanto al significado de "actriz" en "Bardot es una buena actriz" es estructural (Davidson pone objeciones a la explicación *del discurso indirecto* de Frege porque requiere objetos intensionales). El problema es cuáles serían exactamente las restricciones en la empresa de Davidson: ¿qué aparato se debería permitir usar y dónde? Es pertinente que el atractivo de su empresa derive en buena medida de la austeridad del método que parece prometer al principio.

Desde que se lanzaron a la empresa, Davidson y sus seguidores han abordado con diferentes grados de fortuna muchas de las "dificultades y enigmas" señalados en 1967. En 1973 Davidson habla de "progreso bastante impresionante" apuntando a la investigación sobre actitudes proposicionales, adverbios, citas (Davidson, 1967, 1968a, b), nombres propios (Burge, 1973), "debe" (Harman, 1975), términos de masa y comparativos (Wallace, 1970, 1972).

El éxito del programa de Davidson justificaría, en gran medida, la aplicabilidad de la teoría de Tarski a los lenguajes naturales; pero la valoración de este programa depende obviamente del estudio detallado de los análisis específicos ofrecidos. Y como he sugerido con referencia al análisis de la *oratio obliqua*, este estudio a su vez plantea ciertas cuestiones metodológicas que resultan de todos modos lo suficientemente difíciles como para no poder decir con toda confianza que Davidson *ha* mostrado que la teoría de Tarski se aplica al inglés.

7 La teoría de la redundancia

Ramsey

La teoría de la redundancia (aunque sugerida anteriormente por algunos comentarios de Frege en 1918) se deriva primariamente de la obra de F. P. Ramsey en 1927. Ramsey ofrece un esbozo de una teoría en un pasaje muy breve (págs. 142-3) en el curso de una discusión sobre el análisis apropiado de la creencia y el juicio: el contexto es significativo de la estimación que Ramsey da a la importancia del tema: "no hay", piensa él, "realmente ningún problema separado de la verdad, sino meramente un enredo lingüístico".

En resumen, su idea es que los predicados "verdadero" y "falso" son redundantes en el sentido de que pueden ser eliminados de todos los contextos *sin pérdida semántica*[17]; admite que desempeñan un papel pragmático por "énfasis o razones estilísticas". Ramsey considera dos tipos de casos donde "verdadero" y "falso" ocurren típicamente. Los casos que usa para introducir la teoría son del tipo más sencillo en el que la proposición a la que se adscribe verdad o falsedad es dada explícitamente: "es verdadero que *p*", argumenta Ramsey, *significa lo mismo que* "*p*", y "es falso que *p*" *significa lo mismo que* "no *p*". Los casos en los que la proposición relevante no es realmente ofrecida sino solamente descrita presentan, en cambio, más dificultad inicial, pues, como Ramsey reconoce, no podemos simplemente eliminar "es verdadero" de, por ejemplo, "lo que él dice es siempre verdadero"; propone superar esta dificultad usando el aparato de la cuantificación proposicional para ofrecer en el caso mencionado algo así como "Para todo *p*, si él asevera *p*, entonces *p*"[18]. La cuestión de si los cuantificadores de segundo orden que Ramsey necesita pueden ser adecuadamente explicados resulta ser una cuestión clave para la factibilidad de la teoría de la redundancia; pero comenzaré por señalar algunas de las ventajas de la teoría antes de pasar a sus problemas.

Portadores de verdad

En vista de las dificultades causadas por los atributos —hechos y proposiciones— de la teoría de la correspondencia, la austeridad de la teoría de la redundancia es atractiva. Ramsey considera comprensiblemente como una virtud de su teoría que evite las cuestiones planteadas por una explicación de la correspondencia acerca de la naturaleza y la individualización de los hechos. "Es un hecho que...", insiste, tiene la misma redundancia semántica y el mismo uso enfático que "Es verdadero que...".

Además, puesto que el efecto de las teorías estilo Ramsey es negar que, en "es verdadero que *p*", "... es verdadero ..." haya de concebirse como un predicado que adscribe una propiedad *bona fide* a cualquier cosa que "*p*" represente, la cuestión de los portadores de verdad es similarmente eludida; si la verdad no es una propiedad, no se necesita preguntar *de* qué es una propiedad. Observo, sin embargo, que todavía surge lo que he argüido (cap. 6, § 5) que es el tema real que subyace a las disputas acerca de los portadores de verdad —la cuestión de las apropiadas restricciones impuestas

[17] Hay una alusión aquí a la doctrina de los "símbolos incompletos" de Russell, es decir, símbolos que son contextualmente eliminables. Cfr. cap. 5, § 3, para una discusión de esta doctrina con referencia a la teoría de las descripciones de Russell.

[18] Tarski escribe (1944, págs. 68-9) como si la teoría de Ramsey sencillamente no dispusiera de ningún medio para tratar este tipo de casos; Ramsey presumiblemente analizaría los dos casos problemáticos que presenta Tarski —"La primera oración escrita por Platón es verdadera" y "Todas las consecuencias de oraciones verdaderas son verdaderas"— como "(*p*) (si la primera cosa que Platón escribió era que *p*, entonces *p*)" y "(*p*) (*q*) (si *p*, y si *p* entonces *q*, entonces *q*)".

a las instancias de las letras oracionales, i.e., lo que podemos poner en lugar de "p"—. (La preferencia de Ramsey por la locución "es verdadero que p" en vez de por "'p' es verdadera" tiene cierta significación a este respecto.) Yo consideraría como una ventaja de mi diagnóstico del tema acerca de los portadores de verdad el que sea aplicable incluso a las teorías de la redundancia, y como una ventaja de la teoría de la redundancia el que allí aparezca el tema en su forma fundamental.

Desde luego, esto será una genuina economía sólo si es cierto que no se necesitan proposiciones (o lo que fuere) para otros propósitos además del de "portar la verdad". Quienes creen que necesitamos proposiciones como objetos de creencias, por ejemplo, tienden a sentirse menos impresionados por la capacidad que la teoría de la redundancia tiene para prescindir de ellas como portadores de verdad. Es significativo, por tanto, que Prior, que acepta la teoría de Ramsey, insista (1971, cap. 9) en una explicación de la creencia, según la cual, "s cree que ..." en "s cree que p" es un operador para formar oraciones sobre oraciones como "no es el caso que ..." en vez de ser "cree" un símbolo relacional con los argumentos "s" y "que p", denotando el último una proposición. Nuevamente podría suponerse que las proposiciones (o lo que fuere) pudieran ser requeridas como portadoras de *otras* propiedades y que la teoría de la redundancia está, por tanto, en peligro de sacrificar la analogía entre "... es verdadero", y, pongamos por caso, "... es sorprendente" o "... es exagerado" sin, al final, compensación alguna por vía de economía ontológica genuina. Y es significativo a este respecto que Grover y otros, en un ensayo (1975) que insiste sobre las pretensiones de una teoría estilo redundancia, argumenten que es sólo una apariencia engañosa el que "... es verdadero" "... es sorprendente" sean adscripciones a la mismísima cosa.

La distinción lenguaje objeto/metalenguaje

El teórico de la redundancia niega que "es verdadero que p" verse sobre la oración "p": "Es verdad que los leones son tímidos", como "No es el caso que los leones sean tímidos", versa, en su opinión, sobre los leones, no sobre la oración "Los leones son tímidos". Esto significa que no ve necesidad ninguna de insistir sobre la distinción entre lenguaje objeto y metalenguaje que es tan vital para la semántica tarskiana (Prior muestra sumo conocimiento de este punto; por ejemplo, 1971, cap. 7). Esto plantea algunas cuestiones acerca de la capacidad de la teoría de la redundancia para manejar los problemas en los que la distinción lenguaje objeto/metalenguaje desempeña aparentemente un papel importante.

La idea de que la verdad es un predicado metalingüístico parece, por ejemplo, contribuir a las explicaciones usuales de la semántica de las conectivas oracionales, como: "'$-p$' es verdadera sii 'p' es falsa", "'$p \vee q$' es verdadera sii 'p' es verdadera o 'q' es verdadera". ¿Cómo puede ofrecer la teoría de la redundancia una teoría alternativa adecuada? Puesto que esa teoría iguala tanto "Es falso que p" como "Es verdadero que $-p$" con "$-p$", todo lo que queda de la "explicación" de la negación parece ser "$-p$ sii $-p$". El teórico de la redundancia podría insistir en que hay realmente menos de lo que aparece a primera vista en las explicaciones usuales de la negación, pues hay, según él, menos de lo que aparece a primera vista en las explicaciones usuales de la verdad. (Cfr. Dummett, 1958, y el reconocimiento de Grover y otros de que "no es el caso que ..." puede que no sea eliminable.)

Otra dificultad relacionada es que el teórico de la redundancia parece ser incapaz de admitir una distinción aparentemente genuina entre la ley de tercero excluido ("$p \vee -p$") y el principio metalingüístico de bivalencia ("para toda p, 'p' es o verdadera o falsa"). Pues si "'p' es verdadera" significa lo mismo que "p", y "'p' es falsa" significa lo mismo que "$-p$", entonces "'p' es o verdadera o falsa" significa "$p \vee -p$". Una vez más el teórico de la redundancia podría aceptar la consecuencia de que ésta es una distinción sin diferencia; pero, puesto que es una distinción aparentemente con algún poder explicativo, esto le deja con alguna explicación que dar. (Por ejemplo, ¿insistiría él en que los lenguajes "superevaluacionales" de van Fraassen, en donde "$p \vee -p$" es un teorema, pero la semántica permite lagunas de valor de verdad, deben estar confundidos? Cfr. Haack, 1974, págs. 66 y ss., y cap. 11, § 4, más adelante.)

Señalé antes (págs. 122-3) que el esquema (T) parece requerir la bivalencia y esto plantea la cuestión de si una teoría de la redundancia no está también comprometida con la tesis de que "p" debe ser o verdadera o falsa. Pero esta consecuencia es evitable, pues el teórico de la redundancia puede negar, si no es verdadero ni falso que p, que sea falso que es verdadero que p; después de todo, puesto que su teoría es que "es verdadero que p" significa lo mismo que "p", él podría razonablemente insistir en que, si no es ni verdadero ni falso que p, tampoco es ni verdadero ni falso que es verdadero que p. Así no está obligado a negar la posibilidad de lagunas de valor de verdad y, por tanto, el argumento previo no entraña que esté obligado a insistir en la ley de tercero excluido.

En la obra de Tarski, por cierto, el papel más importante de la distinción lenguaje objeto/metalenguaje era asegurar la adecuación formal, específicamente para evitar las paradojas semánticas. Por tanto, su capacidad de habérselas con las paradojas será una cuestión muy crucial para la valoración que uno ofrezca respecto a la factibilidad de la teoría de la redundancia. Esta cuestión debe esperar hasta el cap. 8; pero algunas de las consideraciones sobre los cuantificadores proposicionales de las que me voy a ocupar ahora resultarán relevantes para la misma.

Los cuantificadores: "(p) (si él asevera que p, p)"

Ramsey propone eliminar "verdadero" donde lo que se dice que es verdadero no es ofrecido explícitamente, sino sólo referido oblicuamente por medio de la cuantificación de segundo orden: "Lo que él dice es siempre verdadero", por ejemplo, ha de explicarse como significando "Para todo p, si él asevera p, entonces p". Admite que hay algo raro en este análisis, pues, piensa él, el giro vernáculo parece pedir un "es verdadero" final (como: "(p) (si él asevera p, entonces p es verdadero")) para convertir la "p" final en una oración *bona fide;* pero este aparente obstáculo a la eliminación se supera, arguye Ramsey, si recordamos que "p" es ella misma una oración y que ya contiene un verbo. Suponiendo que todas las proposiciones tuvieran la forma lógica "$a\ R\ b$", sugiere, se podrían observar las propiedades gramaticales escribiendo "para todo a, R, b, si se asevera $a\ R\ b$, entonces $a\ R\ b$". Pero, por supuesto, como bien sabe Ramsey, todas las proposiciones *no* son de la forma "$a\ R\ b$" ni tampoco hay muchas perspectivas de dar una disyunción finita de todas las formas posibles de proposición, de modo que esto difícilmente resuelve el problema.

El descontento de Ramsey es comprensible, pues el problema es real. Si en su fórmula:

(p) (si él asevera p, entonces p)

el cuantificador se interpreta en el estilo objetual estándar, se tiene:

Para todos los objetos (¿proposiciones?) p, si él asevera p, entonces p.

Aquí las "p" ligadas son sintácticamente como términos singulares y la "p" final tiene, por tanto, que ser entendida elípticamente como conteniendo implícitamente un predicado para convertirla en algo de la categoría de una oración capaz de estar a la derecha de "entonces", de la manera siguiente:

Para todas las proposiciones p, si él asevera p, entonces p *es verdadera.*

Pero si resulta que el análisis contiene el predicado "es verdadera", la verdad, después de todo, no ha sido eliminada y no es, después

de todo, redundante. (Ésta es la dificultad que Ramsey ve; es formulada bastante claramente por referencia a la versión de Carnap de la teoría de la redundancia en Heidelberger, 1968.) Si, por otra parte, el cuantificador es interpretado sustitucionalmente, se tiene:

Todas las instancias de sustitución de "Si él asevera..., entonces..." *son verdaderas*

y una vez más "verdadero" aparece en el análisis y, por tanto, no ha sido realmente eliminado.

Así pues, esto está claro: para que la teoría de Ramsey funcione, se necesitará alguna *otra* explicación de los cuantificadores de segundo orden, puesto que bajo cualquiera de las interpretaciones usuales "verdadero" no parece que quede eliminado. Prior ve la dificultad como el resultado de una deficiencia del lenguaje vernáculo, que carece de locuciones coloquiales adecuadas para leer los cuantificadores de segundo orden y nos obliga a recurrir a locuciones de engañosa apariencia nominal tales como "Toda *cosa* que dice". Sugiere, por tanto (1971, pág. 37), "en cualquier tiempo" y "en algún tiempo" como lecturas de "(p)" y "($\exists p$)", y lee "(p) ($p \rightarrow p$)", por ejemplo, como "Si en cualquier tiempo, entonces en ese tiempo".

Grover piensa también que a los cuantificadores se les pueden proporcionar lecturas adecuadas y ofrece cierto aparato gramatical adicional para este propósito. La dificultad de dar una lectura apropiada surge, como sugiere Prior, del hecho de la falta de palabras y expresiones que estén en lugar de oraciones de la manera que los pronombres están en lugar de nombres y descripciones; lo que se necesita, como lo expresa Grover, son *pro-oraciones.*

Pronombres y oraciones son dos tipos de *proformas;* cfr. proverbos, como "do" en inglés, y proadjetivos, como "tal". Una proforma debe ser capaz de ser usada anafóricamente en referencia cruzada o bien como los pronombres de pereza (Geach, 1967), como en "*María* tenía intención de venir a la fiesta, pero *ella* se puso enferma", o bien como los pronombres "cuantificacionales", como en "Si *cualquier coche* se calienta demasiado, no *lo* compres". Las pro-oraciones se parecen a los pronombres en que ocupan posiciones que podrían ocupar las oraciones, como los pronombres ocupan posiciones que podrían ocupar los nombres, y cumplen un papel anafórico similar. La propuesta de Grover es que se lea "(p) (si él asevera que p, entonces p)" como:

Para todas las proposiciones, si él asevera que esso, entonces esso

donde "esso" es una pro-oración. Adviértase que lo que se propone es una nueva *lectura;* ello es, arguye Grover, compatible tanto con una explicación objetual como con una sustitucional al nivel de la interpretación formal.

Esta ingeniosa propuesta plantea una serie de cuestiones para las que solamente puedo ofrecer respuestas provisionales. En primer lugar, recordemos que el problema con el que comencé era si es posible dar una lectura de los cuantificadores proposicionales de Ramsey que sea gramatical y que no reintroduzca el predicado "verdadero". ¿Satisface estos requisitos la lectura de Grover? Pues bien, sería un tanto extraño preguntar si su lectura es gramatical puesto que no es ciertamente española; pide expresamente una *adición* al español. Sería más apropiado preguntarse si hay analogías gramaticales suficientemente fuertes para justificar su innovación; pero, en vista del "suficientemente fuertes", ésta no es una cuestión nada precisa. El inglés, como admite Grover, no tiene ninguna pro-oración atómica —aunque sí tiene, creo yo, expresiones compuestas que desempeñan ese papel: "It is", por ejemplo, que podríamos describir como una oración compuesta de un pronombre y un proverbo—. Y la segunda parte de la cuestión, si la lectura de Grover elimina auténticamente "verdadero", es igualmente engañosa. De hecho, hay dos puntos a plantear aquí. El primero es que aun cuando se ofrezca una *lectura* apropiada, esto deja abierta la cuestión de si no hay todavía una apelación implícita a la verdad al nivel de la interpretación formal. (Y exactamente, *¿de* qué se debe eliminar "verdadero" para mostrar que es redundante?) La segunda cuestión es si nuestra comprensión de "esso" requiere implícitamente la noción de verdad.

La "teoría pro-oracional de la verdad"

La aplicación que la propia Grover hace de su explicación de la cuantificación proposicional a la teoría de la verdad puede arrojar alguna luz sobre este problema. Grover y otros, 1975, proponen una versión modificada de la teoría de la redundancia según la cual "eso es verdadero" se aplica como siendo ella misma una pro-oración. Según su explicación, las adscripciones veritativas son eliminables en favor de "Es verdadero" *en cuanto pro-oración atómica*, i.e., una oración en la que "verdadero" no es un predicado separable[19].

¿Qué muestra esto sobre si la teoría "pro-oracional" elimina realmente "verdadero"? "Verdadero", se nos dice, es eliminable; no del español, con seguridad, sino del español + "esso". Pero ¿cómo hemos de entender "esso"? Pues bien, no hay nada que sea *exactamente* como ella en español, pero funciona como "Eso es verdadero" excepto en que es atómica más bien que compuesta...

Queda abierto a la duda, pienso yo, si se ha justificado la esperanza de Ramsey de eliminar enteramente la referencia a la verdad. No obstante, hay algo importante que aprender de la discusión de la teoría pro-oracional: que el predicado de verdad desempeña un papel crucial permitiéndonos hablar *generalmente*, esto es, hablar sobre proposiciones que no exhibimos realmente, sino a las que sólo nos referimos indirectamente, papel que comparte con el aparato de los cuantificadores de segundo orden ("proposicionales" u "oracionales"). Esta semejanza de función resultará relevante para el diagnóstico de las paradojas semánticas.

[19] Ramsey pensaba que toda referencia a la verdad es eliminable; Grover y otros admiten que hay un residuo. En algunos casos la eliminación de "verdadero" pide la modificación de la oración contenida, como "Solía ser verdadero que Roma era el centro del mundo conocido"/"Roma solía ser el centro del mundo conocido" o "Podría ser verdadero que haya vida en Marte"/"Podría haber vida en Marte". Y donde este fenómeno se combina con la cuantificación, como en "Algunas oraciones solían ser verdaderas pero ya no son verdaderas", se ven obligados a introducir nuevas conectivas, como "$(\exists p)$ (solía-ser-el-caso-que p pero ya-no-es-el-caso-que p)" que admiten que son, en efecto, locuciones veritativas. Sus comentarios acerca de "podría ser verdadero que", por otra parte, sugieren una alternativa interesante a la idea de que la verdad necesaria, igual que la verdad, es una propiedad de las oraciones o las proposiciones.

8
Paradojas

1 La paradoja del mentiroso y afines

La importancia que la paradoja del mentiroso tiene para la teoría de la verdad se ha hecho ya patente; en Tarski las condiciones de adecuación formal de las definiciones de verdad están motivadas, en gran parte, por la necesidad de evitarla. Ha llegado ya el momento de prestar atención directa por sí mismas a la paradoja del mentiroso y afines.

¿Por qué la "paradoja del mentiroso"? La oración del mentiroso, junto con principios acerca de la verdad aparentemente obvios, conduce a contradicción mediante un razonamiento aparentemente válido; por eso es por lo que se la llama paradoja (del griego "*para*" y "*doxa*", "más allá de la creencia")[1].

La paradoja del mentiroso se presenta en diversas variantes; la versión clásica versa sobre la oración:

(*O*) Esta oración es falsa

Supongamos que *O* es verdadera; entonces lo que dice es el caso; por tanto, es falsa. Supongamos, por otra parte, que *O* es falsa; entonces, lo que dice no es el caso; por tanto, es verdadera. Luego *O* es verdadera sii *O* es falsa. Las variantes incluyen indirectamente las oraciones autorreferenciales tales como:

La oración siguiente es falsa. La oración anterior es verdadera.

y la "paradoja de la tarjeta postal", cuando uno supone que en una cara de la tarjeta se ha escrito:

La oración que hay en la otra cara de esta tarjeta es falsa

[1] Las "paradojas" de la implicación material e implicación estricta —discutidas prolijamente en el cap. 11— son, en el peor de los casos, contraintuitivas, pero no contradictorias como la del mentiroso; por eso se las cita poco.

y en la otra cara:

La oración que hay en la otra cara de esta tarjeta es verdadera.

Otra variante es la paradoja de "Epiménides" que se refiere a un cretense llamado Epiménides al que se le atribuye haber dicho que todos los cretenses son siempre mentirosos. Si un mentiroso es alguien que siempre dice lo que es falso, entonces si lo que dice Epiménides es verdadero, es falso. La de Epiménides es, sin embargo, algo menos paradójica que la del mentiroso, ya que se puede suponer consistentemente que es falsa, pero no que es verdadera (cfr. Anderson, 1970). Están también las variantes de "los decidores de la verdad" ("Esta oración es verdadera") y las variantes imperativas ("Desobedece esta orden").

Otras paradojas involucran "verdadero (falso) de ..." más bien que "verdadero (falso)". "Heterológico" significa "no verdadero de sí mismo"; así, por ejemplo, "alemán", "largo", "cursiva" son adjetivos heterológicos, mientras que "castellano", "corto", "impreso" son autológicos, verdaderos de sí mismos. Ahora bien, ¿"heterológico" es "heterológico"? Pues, si heterológico *es* heterológico, no es verdadero de sí mismo; por tanto, no es heterológico. Pero, si *no* es heterológico, es verdadero de sí mismo; por tanto, es heterológico. Así pues, "heterológico" es heterológico sii "heterológico" no es heterológico (paradoja de Grelling).

Otras involucran además "definible" o "especificable". El número tres es susceptible de especificar mediante un nombre de una sílaba, el número cinco mediante un nombre de dos sílabas, el número catorce por un nombre de tres sílabas. Considérese, entonces, el menor de los números enteros no susceptibles de ser nombrados con menos de treinta y tres sílabas. Ese número es susceptible de ser nombrado con treinta y dos sílabas mediante "el menor de los números enteros no susceptibles de ser nombrados con menos de treinta y tres sílabas" (paradoja de Berry). Sea E la clase de decimales que pueden ser definidos por medio de un número finito de palabras y sean ordenados sus miembros como el primero, segundo, tercero, ..., etc. Sea ahora N un número tal que si la n-sima cifra del n-simo decimal de E es m, entonces la n-sima cifra de N es $m + 1$, o 0 si $m = 9$. En ese caso N es diferente de todos los miembros de E y, sin embargo, ha sido definido por medio de un número finito de palabras (paradoja de Richard).

Otras paradojas involucran el concepto de conjunto. Algunos conjuntos son miembros de sí mismos, mientras que otros, no (por ejemplo, el conjunto de los objetos abstractos, al ser él mismo un objeto abstracto, es miembro de sí mismo; el conjunto de las vacas, al no ser él mismo una vaca, no es miembro de sí mismo). Considérese ahora el conjunto de los conjuntos que no son miembros de sí mismos. ¿Es él miembro de sí mismo o no? Si *es* miembro de sí mismo, entonces posee la propiedad que poseen todos sus miembros, es decir, *no* es miembro de sí mismo; si, por otra parte, *no* es miembro de sí mismo, entonces posee la propiedad que califica a un conjunto de pertenecer a sí mismo, por tanto, *es* miembro de sí mismo. En consecuencia, el conjunto de todos los conjuntos que no son miembros de sí mismos es miembro de sí mismo sii no es miembro de sí mismo (paradoja de Russell). Otras paradojas de la teoría de conjuntos son la paradoja de Cantor: ningún conjunto puede ser mayor que el conjunto de todos los conjuntos, pero, para cualquier conjunto, hay otro, el conjunto de todos sus subconjuntos, que es mayor que él; y la de Burali-Forti: la serie de todos los números ordinales tiene un número ordinal, a saber, Ω, pero la serie de todos los ordinales hasta llegar a un ordinal cualquiera dado, e incluido éste, excede en una unidad al ordinal dado, de modo que la serie de todos los ordinales hasta Ω, e incluido éste, tiene un número ordinal $\Omega + 1$.

De ningún modo agotan estos ejemplos el campo de las paradojas que se encuentran en la literatura (para más ejemplos, cfr. Russell, 1908a; Mackie, 1973, apéndice). Espero, no obstante, que mi lista sea suficientemente representativa para ilustrar el tipo de problemas con los que debe enfrentarse una solución a las paradojas; el motivo de considerar cierto número de variantes es para que nos permita comprobar si las soluciones propuestas alcanzan suficientemente sus objetivos.

¿Paradojas de la "teoría de conjuntos" versus paradojas "semánticas"?

Aunque algunas de estas paradojas habían sido conocidas mucho tiempo antes, comenzaron a adquirir interés filosófico serio después del descubrimiento que Russell hizo de su paradoja. Frege había reducido la aritmética al cálculo de oraciones, al cálculo de predicados y a la teoría de conjuntos. Russell, sin embargo, mostró que su paradoja era en realidad un teorema del sistema de Frege, sistema que resultó ser, por ende, inconsistente. (Dado que Frege esperaba proporcionar los fundamentos de la aritmética reduciéndola a principios autoevidentes, el hecho de que sus axiomas lógicos "autoevidentes" resultasen contradictorios, supuso naturalmente una conmoción epistemológica muy fuerte; cfr. cap. 1, § 2.) Las paradojas no pueden desecharse como meros trucos o rompecabezas, puesto que se siguen de los principios de la teoría de conjuntos intuitivamente obvios y, en consecuencia, amenazan a los mismos fundamentos de la teoría de conjuntos. En virtud de que de una contradicción se deriva cualquier cosa, las consecuencias de las paradojas para una teoría en la que éstas sean derivables son completamente intolerables (pero cfr. cap. 11, § 6, para otras consideraciones sobre "$p \,\&\, -p \vdash q$"). La paradoja de Russell actúa como constricción clave sobre las tentativas para idear teorías de conjuntos consistentes; similarmente, la paradoja del mentiroso actúa como constricción clave sobre las tentativas para idear teorías semánticas consistentes.

Pero esto da lugar a una cuestión importante, aunque difícil. Como el mismo comentario sobre la analogía entre el papel de la paradoja de Russell en la teoría de conjuntos y el papel de la paradoja del mentiroso en la teoría semántica sugiere, es posible clasificar las paradojas en dos grupos distintos: las que involucran esencialmente conceptos de la teoría de conjuntos tales como "∈" y "número ordinal", y las que involucran esencialmente conceptos semánticos tales como "falso", "falso de ..." y "definible". De hecho, es común distinguir entre paradojas de la *teoría de conjuntos* y *paradojas semánticas* (la distinción se remonta a Peano; su uso corriente se deriva de la propuesta de Ramsey en 1925):

paradojas de la teoría
de conjuntos
(Ramsey: "*lógicas*")
Paradoja de Russell
Paradoja de Cantor
Paradoja de Burali-Forti
(Involucran esencialmente
"conjunto", "∈", "número ordinal")

paradojas
semánticas
(Ramsey: "*epistemológicas*")
Paradoja del mentiroso y sus variantes
Paradoja de Grelling
Paradojas de Berry y Richard
(Involucran esencialmente
"falso", "falso de", "definible")

El segundo grupo tiene interés inmediato para la teoría semántica.

El mismo Russell, sin embargo, no consideró que las paradojas se distribuyesen en dos grupos distintos, *ya que pensaba que todas ellas surgían como resultado de una falacia*, de las violaciones del "principio del círculo vicioso". Si se supone que unas paradojas nacen a causa de alguna peculiaridad de los conceptos de la teoría de conjuntos y otras a causa de alguna peculiaridad de los conceptos semánticos, entonces será aceptable la clasificación en dos grupos; pero si se piensa, como Russell, que la dificultad radica en algo más profundo común a todas las paradojas, se encontrará uno con que resulta engañosa. Es difícil negar, pienso, que todas las paradojas esbozadas guarden *prima facie* afinidad entre sí y que una solución a todas ellas sería con seguridad más satisfactoria que una solución solamente para algunas; y en vista de esto, lo más seguro parece ser *no* plantear preguntas que habrían de quedar sin respuesta y centrarnos exclusivamente en las paradojas "semánticas".

2 "Soluciones" a las paradojas

Requisitos para una solución

Antes de intentar valorar las soluciones que se han dado, pienso que es conveniente tratar precisamente de aclarar en qué consistiría una "solución". ¿Cuál es exactamente el problema? —que las conclusiones contradictorias se siguen mediante un razonamiento aparentemente intachable de premisas aparentemente intachables—. Esto sugiere dos requisitos para una solución: que se debería dar una teoría formal consistente (de semántica o teoría de conjuntos, según el caso) —en otras palabras, que se indique qué premisas o principio de inferencia aparentemente intachables deben rechazarse (la solución *formal); y*, además, que se debería proporcionar una explicación de por qué dicha premisa o principio, a pesar de las apariencias, es recusable (la solución *filosófica)*—. Es difícil precisar justamente lo que una tal explicación requiere, pero, más o menos, lo que se pretende es que se muestre que la premisa o principio rechazados pertenecen a una clase para la que hay objeciones independientes —esto es, objeciones independientes del hecho de conducir a la paradoja—. Es importante, aunque difícil, evitar las supuestas "soluciones" que simplemente *etiquetan* las oraciones infractoras de una forma que parece explicativa, pero que realmente no lo es. Hay otros requisitos concernientes al alcance de la solución; no debería ser tan amplia que mutilase el razonamiento que queremos defender (el principio de "no te arranques la nariz por despecho contra tu cara"); pero debería ser suficientemente amplia para bloquear todos los argumentos paradójicos relevantes (el principio de "no saltes de la sartén al fuego"); lo "relevante", desde luego, encubre algunos problemas. A nivel formal, el segundo principio pide sencillamente que la solución sea tal que se restablezca la consistencia. La respuesta de Frege ante la inconsistencia encontrada por Russell en su teoría de conjuntos fue una restricción formal que evita la paradoja de Russell, pero que que permite todavía paradojas muy afines y, por ende, viola este requisito (véase Frege, 1903; Quine, 1955; Geach, 1956). A nivel filosófico, el principio de la "sartén y el fuego" apremia a que la

explicación ofrecida sea lo más profunda posible; esto es, por supuesto, lo que sirve de base a mi presentimiento de que una solución para *ambos* grupos de paradojas "semánticas" y de "teoría de conjuntos", si fuera posible, sería preferible a una solución local para un grupo.

La fuerza de estos requisitos quizás pueda apreciarse considerando brevemente algunas de las soluciones propuestas que fallan al encontrarse con ellos.

Se ha sugerido a veces que las paradojas se resolverían prohibiendo la autorreferencia; pero esta sugerencia es a la vez demasiado amplia y demasiado restringida. Infringe el principio de "no te arranques la nariz por despecho contra tu cara": pues no solamente son autorreferenciales muchas oraciones completamente inofensivas ("Esta oración está en castellano", "Esta oración está en tinta roja") (cfr. Popper, 1954; Smullyan, 1957), sino también algunos argumentos matemáticos, incluyendo la prueba de incompletud de la aritmética de Gödel, usan esencialmente oraciones autorreferenciales (cfr. Nagel y Newman, 1959, y Anderson, 1970); de modo que las consecuencias de prohibir la autorreferencia serían muy serias. Y, puesto que no todas las variantes de la oración del mentiroso son francamente autorreferenciales (ninguna oración en "La oración siguiente es falsa, La oración anterior es verdadera" se refiere a sí misma), esta propuesta es al mismo tiempo demasiado estrecha todavía.

El argumento de que la oración del mentiroso lleva a contradicción usa la suposición de que "Esta oración es falsa" es o verdadera o falsa; y por eso, naturalmente, se ha sugerido a menudo que la forma de bloquear el argumento es negar esta suposición. Bochvar propuso (1939) tratar al Mentiroso adoptando una lógica trivalente en la que el tercer valor, "paradójico", se asigne a las oraciones recalcitrantes (véase también Skyrms, 1970a, 1970b, y el cap. 11, § 3). Esta propuesta tiene el peligro también de ser demasiado amplia y demasiado restringida: demasiado amplia porque exige un cambio en los principios de la lógica elemental (cálculo de oraciones); y demasiado restringida todavía porque deja problemas con la paradoja del "mentiroso reforzado" —la oración:

> Esta oración es o falsa o paradójica

que es falsa o paradójica si es verdadera, verdadera si es falsa, y verdadera si es paradójica.

Hay otro enfoque que niega también que la oración del mentiroso sea verdadera o falsa, sin sugerir, sin embargo, que tenga un tercer valor de verdad, argumentando que no es un ítem del tipo adecuado para tener un valor de verdad. Se argumenta que sólo los enunciados son verdaderos o falsos, y una elocución de la oración del mentiroso no constituiría un enunciado. (Véase Bar-Hillel, 1957; Prior, 1958; Garver, 1970; y cfr. —*mutatis mutandis* con "proposición" para "enunciado"— Kneale, 1971.) Este tipo de enfoque adolece, pienso, de una explicación inadecuada —no proporciona una base racional adecuada para negar un valor de verdad a las oraciones infractoras. Incluso concediendo por mor del argumento que solamente los enunciados o proposiciones pueden ser o verdaderos o falsos (pero concedido *solamente* por mor del argumento —cfr. el cap. 6—), se necesitaría un argumento de por qué no se dispone de un ítem del tipo apropiado en el caso del mentiroso. Después de todo, la oración del mentiroso no padece ninguna deficiencia obvia de gramática o vocabulario. Los requisitos mínimos serían; primero, una explicación clara de bajo qué condiciones la elocución de una oración constituye un enunciado; segundo, un argumento de por qué ninguna elocución del mentiroso puede cumplir estas condiciones; tercero, un argumento de por qué solamente los enunciados pueden ser o verdaderos o falsos. En otro caso, tendrá uno derecho a quejarse de que la solución no es suficientemente explicativa.

Solución de Russell: la teoría de los tipos, el principio del círculo vicioso

Russell ofrece (1908a) tanto una solución formal, la teoría de los tipos, como una solución filosófica, el principio del círculo vicioso.

Hoy en día es costumbre distinguir en la solución formal de Russell la teoría simple y la teoría ramificada de los tipos. *La teoría simple de los tipos* divide el universo del discurso en una jerarquía; individuos (tipo 0), conjuntos de individuos (tipo 1), conjuntos de conjuntos de individuos (tipo 2), ..., etc., y, correspondientemente, variables suscritas con un índice de tipos, de tal forma que x_0 corresponde al tipo 0, x_1 al tipo 1, ..., etc. Así pues, las reglas de formación se restringen de tal modo que una fórmula de la forma "$x \in y$" está bien formada sólo si el índice de tipos de y es superior en uno al de x. Así, por tanto, "$x_n \in x_n$" está mal formada, y la propiedad de no ser miembro de sí mismo, esencial en la paradoja de Russell, no puede expresarse. *La teoría ramificada de tipos* impone una jerarquía de órdenes de "proposiciones" (oraciones cerradas) y "funciones proposicionales" (oraciones abiertas) así como la restricción de que ninguna proposición (función proposicional) puede ser "acerca de", i.e., que contenga un cuantificador que abarque a proposiciones (funciones proposicionales) del mismo orden o superior que el suyo propio. "Verdadero" y "falso" deben ser también suscritos, dependiendo del orden de la proposición a la cual

se aplican; una proposición del orden n será verdadera (falsa) $n + 1$. La oración del mentiroso que afirma de sí mismo que es verdadera, se torna entonces inexpresable, precisamente lo mismo que ocurría en la teoría simple con la propiedad de no ser miembro de sí mismo (he simplificado considerablemente; véase Copi, 1971, para una explicación más detallada).

Russell, sin embargo, no clasificó las paradojas en dos grupos distintos, pues pensaba que *todas* las paradojas provenían de una y la misma falacia, i.e., de violaciones de lo que él, siguiendo a Poincaré, llamó "el principio del círculo vicioso" (P.C.V.):

> "Lo que presupone *el todo* de una colección no debe formar parte de la colección": o, conversamente, "Si, supuesto que una cierta colección, tuviera un total, hubiera de tener miembros definibles solamente en términos de ese total, entonces la referida colección no tiene total". [Nota de pie de página: Quiero decir que los enunciados acerca de *todos* sus miembros carecen de sentido.] (1908a, pág. 63.)

Formula el P.C.V. de diversas formas obviamente equivalentes: por ejemplo, una colección no debe "involucrar" o "ser definible solamente en términos de" sí misma. El P.C.V. motiva las restricciones del tipo/orden impuestas a la teoría formal, mostrando que lo que las fórmulas calificadas de mal formadas dicen es demostrablamente carente de significado. Es importante que la misma base racional se da tanto en la teoría simple como en la teoría ramificada. Ciertamente, puesto que Russell sostuvo que los conjuntos son en realidad construcciones lógicas de las funciones proposicionales, consideró las restricciones de la teoría simple como un caso especial de las de la teoría ramificada (cfr. Chihara, 1972, 1973).

La explicación de Russell tanto a nivel formal como filosófico tiene sus dificultades. Formalmente, hay algún peligro de que Russell se haya cortado la nariz por despecho contra su cara; las restricciones evitan las paradojas, pero también bloquean ciertas inferencias deseadas. Recuérdese que Russell intentó completar el programa, comenzado por Frege, de reducir la aritmética a "lógica", i.e., al cálculo de oraciones, al cálculo de predicados de primer orden y a teoría de conjuntos. Sin embargo, las restricciones de los tipos bloquean la prueba de la infinitud de los números naturales y las restricciones de los órdenes bloquean la prueba de ciertos teoremas afines. En los *Principia Mathematica* se preservaron éstas mediante la introducción de nuevos axiomas, el axioma de infinitud y el axioma de reducibilidad, respectivamente; esto asegura la derivabilidad de los postulados aritméticos de Peano; pero el carácter *ad hoc* de estos axiomas merma la plausibilidad de la tesis de que la aritmética haya sido reducida a una base *puramente* *lógica.* Con todo, podría pensarse que estas dificultades, aunque pongan en duda la viabilidad del logicismo de Russell, no muestran necesariamente que su solución a las paradojas sea equivocada.

Pero mis sospechas quedan confirmadas por las dificultades del nivel filosófico. En primer lugar, el P.C.V. ciertamente no está establecido con toda la precisión que sería de desear; y, por tanto, es difícil ver exactamente lo que está equivocado junto con las violaciones del mismo. Ramsey comentó que no veía nada objetable en especificar a un hombre como, digamos, el máximo bateador de su equipo —especificación que aparentemente viola el P.C.V.—. No todos los círculos excluidos por el P.C.V., señalaba, son verdaderamente viciosos (obsérvese la analogía con las dificultades de la propuesta de prohibir todas las oraciones autorreferenciales).

Sin embargo, a pesar de estas dificultades, el diagnóstico y la solución de Russell han continuado ejerciendo su influencia; más adelante, en el § 3, argumentaré que el enfoque de Russell es ciertamente correcto en algunos aspectos. Pero lo que me preocupa de momento son otras soluciones que se parecen a la de Russell en aspectos interesantes. Su diagnóstico repercutió en el enfoque de Ryle. Éste argumenta (1952) que "El enunciado corriente es falso" debe descomponerse en "El enunciado corriente (a saber, que el enunciado corriente... [a saber, que el enunciado corriente... {a saber, ..., etc.}]) es falso", y nunca se alcanza ningún enunciado completamente especificado. Ryle, lo mismo que Russell, piensa que la "autodependencia" de la oración del mentiroso de algún modo la priva de sentido. Mackie, 1973, está de acuerdo con Russell y Ryle en que el problema radica en la "autodependencia viciosa" del mentiroso, pero prefiere decir, en virtud de que la oración del mentiroso es en apariencia correctamente construida a partir de componentes *bona fide*, que el resultado no es carente de significado, sino "vacío de contenido". Sin embargo, puesto que él pone especial cuidado en distinguir "vacío de contenido" de vacío de significado *y* de vacío de valor de verdad, le deja a uno un tanto perplejo para comprender precisamente qué clase de vacío es el vacío de contenido. Y el enfoque de Tarski a las paradojas semánticas, al que volveré a continuación, guarda ciertas semejanzas significativas (observadas por Russell, 1956; y cfr. Church, 1976) con la jerarquía russelliana de los órdenes de proposiciones.

Solución de Tarski: la jerarquía de lenguajes

Tarski diagnostica las paradojas semánticas (a las que restringe su atención) como resultado de dos suposiciones:

(i) que el lenguaje es semánticamente cerrado, i.e., contiene:

(a) los medios para referirse a su propia expresión, y (b) los predicados "verdadero" y "falso"

(ii) que las leyes usuales de la lógica son válidas

y, estando poco dispuesto a negar (ii) (pero cfr. los comentarios anteriores sobre la propuesta de Bochvar), niega (i), proponiendo como una condición de adecuación formal que la verdad se defina para lenguajes *semánticamente abiertos.* Así Tarski propone una *jerarquía de lenguajes:*

el lenguaje objeto, O,
el metalenguaje, M,
que contiene (a) medios para referirse a expresiones de O
y (b) los predicados "verdadero en O" y "falso en O",
el metametalenguaje, M',
que contiene (a) medios para referirse a expresiones de M
y (b) los predicados "verdadero en M'" y "falso en M'", etc.

Puesto que en esta jerarquía de lenguajes, la verdad para un nivel dado es expresada siempre por un predicado del siguiente nivel, la oración del mentiroso solamente puede aparecer en la forma inofensiva "Esta oración es falsa en O", que debe ser ella misma una oración de M, y por eso no puede ser verdadera en O y es simplemente falsa en vez de paradójica.

Aunque el atractivo de la teoría de la verdad de Tarski ha obtenido bastante apoyo para esta propuesta, también ha recibido críticas por su "artificialidad". La jerarquía del lenguaje y la relativización de "verdadero" y "falso" evitan las paradojas semánticas, pero parece faltarles justificación intuitiva independientemente de su utilidad a este respecto. En otras palabras, el enfoque de Tarski parece dar una solución formal, pero no filosófica. La razón que aduce Tarski para exigir la abertura semántica es, simplemente, que la clausura semántica conduce a la paradoja. *Hay* una base racional independiente para la relativización de "verdadero" y "falso" respecto a un lenguaje —que Tarski está definiendo "la verdad" para oraciones (fbfs), y una y la misma oración (fbf) puede tener diferente significado y, por ende, diferente valor de verdad en lenguajes diferentes; pero esta base racional no proporciona ninguna justificación independiente para sostener que "verdadero en L" es siempre un predicado, no de L, sino del metalenguaje de L.

Intuitivamente, uno no considera "verdadero" como sistemáticamente ambiguo de la forma en que Tarski sugiere que debería ser. Tal vez esta contraintuición no debería ser por sí misma una consideración contundente. Pero Kripke (1975) señala que estas atribuciones ordinarias de verdad y falsedad ni siquiera pueden ser asignadas a niveles *implícitos*. Supongamos, por ejemplo, que Jones dice:

> Todas las elocuciones de Nixon sobre Watergate son falsas.

Esto tendría que asignarse al siguiente nivel superior al más alto nivel de cualquiera de las elocuciones de Nixon sobre Watergate; pero no sólo no dispondremos ordinariamente de ninguna forma de determinar los niveles de las elocuciones de Nixon sobre Watergate, sino que además en circunstancias desfavorables hasta puede ser realmente imposible asignar niveles de forma consistente —supóngase que una de las elocuciones de Nixon sobre Watergate es:

> Todas las elocuciones de Jones sobre Watergate son falsas

entonces la elocución de Jones tiene que estar en un nivel por encima de todas las de Nixon, y las de Nixon en un nivel por encima de todas las de Jones.

Kripke argumenta que el enfoque de Tarski no tiene en cuenta suficientemente el carácter "arriesgado" de las atribuciones de verdad. Aserciones completamente corrientes sobre verdad y falsedad, indica Kripke, pueden resultar paradójicas si los hechos empíricos son desfavorables. Supongamos, por ejemplo, que Nixon hubiera dicho que todas las elocuciones de Jones sobre Watergate son verdaderas; entonces, la aserción de Jones de que todas las aserciones de Nixon sobre Watergate son falsas sería falsa, si verdadera, y verdadera, si falsa (cfr. la "paradoja de la tarjeta postal", § 1). La moraleja, sugiere Kripke, es que uno apenas puede esperar que las oraciones recalcitrantes sean distinguidas por cualquier característica sintáctica o semántica, sino que se debe buscar una base racional que permita que la paradoja pueda surgir respecto de cualquier atribución de verdad si los hechos resultan mal[2].

Solución de Kripke: fundamentación

Kripke busca proporcionar una explicación del origen de la paradoja que sea más satisfactoria en este aspecto y, entonces, intenta construir una teoría formal sobre esta base. (Mi presentimiento es que este es el camino correcto para avanzar.) Su propuesta depende del rechazo de la idea —dada por supuesta por Tarski— de que el predicado de verdad debe ser totalmente definido, es decir, que toda oración adecuadamente bien formada debe ser o verdadera o falsa. Por tanto, esto guarda afinidades con la propuesta de Bochvar de una lógica trivalente y con las propuestas del *no ítem* anteriormente consideradas. Pero Kripke acentúa que su idea *no* es que las oraciones paradójicas tengan algún valor de verdad no clásico, sino que no tienen *ningún* valor de verdad.

El concepto de *fundamentación*, introducido primero por Herzberger en 1970, es la idea clave en la explicación de cómo se asignan los valores de verdad a las oraciones ordinarias —y de cómo las oraciones extraordinarias no obtienen ningún valor—. Kripke explica la idea así:

Supongamos que uno está intentando explicar la palabra "verdadero" a alguien que no la entiende. Se podría introducir mediante el principio de que uno puede aseverar que una oración es verdadera solamente cuando tiene derecho a aseverar esa oración y que puede aseverar que una oración no es verdadera solamente cuando tiene derecho a negarla. Ahora bien, dado que el aprendiz tiene derecho a aseverar que

> La nieve es blanca

esta explicación le dice que tiene derecho a aseverar que

> "La nieve es blanca" es verdadera.

Ahora él puede extender ese uso de "verdadero" a otras oraciones como, por ejemplo, "La nieve es blanca" ocurre en Tarski, 1944; la explicación le permite aseverar que

> Alguna oración en "La concepción semántica de la verdad" es verdadera.

Y puede también extender su uso de "verdadero" a oraciones que ya contienen "verdadero", por ejemplo, aseverar que

> "'La nieve es blanca' es verdadera" es verdadera

o

> "Alguna oración en 'La concepción semántica de la verdad' es verdadera" es verdadera.

La idea intuitiva de *fundamentación* es que una oración está fundamentada precisamente en el caso de que eventualmente alcanzase un valor de verdad en este proceso. No todas las oraciones *alcanza-*

[2] Kripke pone también la objeción técnica de que la jerarquía de Tarski no ha sido extendida a niveles transfinitos y que, además, existen dificultades para tal extensión.

rán un valor de verdad de esta forma; entre las oraciones "no fundamentadas" que no lo alcanzarán están:

Esta oración es verdadera

y

Esta oración es falsa.

Esta idea tiene afinidades con la noción —expresada en el P.C.V. de Russell y por Ryle y Mackie— de que lo que sucede en las oraciones paradójicas es una especie de autodependencia viciosa. Sin embargo, a las oraciones no fundamentadas se les permite ser significativas, mientras que la idea de Russell es que la violación del P.C.V. produce carencia de significado.

Formalmente, esta idea se representa (simplifico bastante) en una jerarquía de lenguajes interpretados donde, en cualquier nivel, el predicado de verdad es el predicado de verdad para el nivel inferior inmediato. En el nivel inferior, el predicado "*V*" es completamente indefinido. (Esto corresponde al estadio inicial de la explicación intuitiva.) En el siguiente nivel, el predicado "*V*" se asigna a fbfs, que en sí mismas no contienen "*V*". Se ha supuesto que esta asignación estará de acuerdo con las reglas de Kleene que proporcionan la asignación de valores a fbfs compuestas proporcionando la asignación —o falta de asignación— a sus componentes: "$-p$" es verdadera (falsa) si "p" es falsa (verdadera), indefinida si "p" es indefinida; "$p \vee q$" es verdadera si al menos uno de sus miembros es verdadero (aunque el otro sea verdadero, falso o indefinido), falsa si los dos miembros son falsos, si no indefinida; "$(\exists x)Fx$" es verdadera (falsa) si Fx es verdadero para algunas (falso para todas) las asignaciones de x, si no indefinido. (Esto corresponde al primer estadio en el que el aprendiz asigna "verdadero" a una oración si tiene derecho a aseverar la oración.) En cada nivel, las fbfs a las que se les asignó "*V*" y "*F*" en un nivel previo retienen esos valores, pero se asignan valores a nuevas fbfs para las que "*V*" fue previamente indefinido —"*V*" se hace *más definido* a medida que avanza el proceso—. Pero el proceso no continúa indefinidamente con nuevas oraciones que alcanzan valores en cada nivel; eventualmente —"en un punto fijo"— el proceso se detiene. Ahora se puede definir formalmente la idea intuitiva de fundamentación: una fbf está fundamentada si posee un valor de verdad en el punto fijo menor, de lo contrario es no fundamentada. El punto fijo menor o "mínimo" es el primer punto en el que el conjunto de oraciones verdaderas (falsas) es el mismo que el conjunto de oraciones verdaderas (falsas) del nivel precedente. Todas las oraciones paradójicas son no fundamentadas, pero no todas las oraciones no fundamentadas

son paradójicas; oración paradójica es aquella a la que no se le puede asignar consistentemente un valor de verdad en *ningún* punto fijo. Esto proporciona alguna explicación de por qué "Esta oración es verdadera" parece participar algo de la singularidad de "Esta oración es falsa" y, sin embargo, a diferencia de la oración del mentiroso, es consistente. Se *puede* dar un valor de verdad a "Esta oración es verdadera", pero sólo *arbitrariamente; no se puede* dar un valor de verdad consistentemente a "Esta oración es falsa". Este esquema tiene en cuenta también el "riesgo" de las asignaciones de verdad: pues el carácter paradójico de una oración puede ser o bien intrínseco (como sería en "Esta oración es falsa") o empírico (como sería en "La oración citada en la pág. 171, líneas 12-3, es falsa").

Observé arriba que la relajación del requisito de que "verdadero" tiene que ser completamente definido, la admisión de lagunas de valor de verdad, proporcionó a la idea de Kripke cierta analogía también con propuestas que, como la de Bochvar, sostienen que las paradojas semánticas se evitan recurriendo a la lógica trivalente. Esto plantea la cuestión de cómo evita Kripke las críticas hechas antes de la solución de Bochvar. Kripke mismo recalca que no considera como un desafío a la lógica clásica el uso que él hace de las reglas "trivalentes" de evaluación de Kleene. Si el uso de matrices trivalentes conlleva necesariamente tal desafío, es una cuestión difícil sobre la que hablaré más en el cap. 11, § 3; de momento, admitiré la afirmación de Kripke de que sus propuestas son compatibles con el conservadurismo lógico. Pero ¿qué pasa entonces con el mentiroso reforzado?

Kripke no se enfrenta directamente con este asunto, pero es posible calcular lo que diría sobre ello. Dice que las nociones de "fundamentación" y "paradójico", a diferencia del concepto de verdad, no pertenecen a su jerarquía de niveles del lenguaje. (Considérese de nuevo la representación intuitiva de un aprendiz que posee el concepto de verdad que se le ha explicado. Sus instrucciones no le proporcionan ningún medio para asignar un valor de verdad a una oración no fundamentada como "Esta oración es verdadera"; pero él no puede concluir que "Esta oración es verdadera" no es verdadera, pues sus instrucciones le indican que puede negar que una oración es verdadera sólo si tiene derecho a negar esa oración.) Ahora bien, si "paradójico" pertenece, no a la jerarquía de niveles del lenguaje, sino al metalenguaje de esa jerarquía, entonces Kripke puede ponerle en un aprieto al "mentiroso reforzado", "Esta oración es o falsa o paradójica", de una forma muy semejante a lo que Tarski hace con el mentiroso. Sin embargo, esto puede producir alguna insatisfacción: pues *es* un poco desilusionador encontrar que la innovación que el enfoque de Kripke hace del mentiroso tiene que comprometerse con el rechazo neotarskiano del mentiroso reforzado. (*¿Es* indiferente que le ahorquen a uno por una oveja o por un cordero?)

Vale la pena hacer un resumen de los puntos principales de comparación y contraste entre el enfoque de Kripke, la teoría de los tipos de Russell y la jerarquía de lenguajes de Tarski:

RUSSELL	TARSKI	KRIPKE
solución formal		
jerarquía de órdenes de proposiciones	jerarquía de lenguajes (problemas con niveles transfinitos)	jerarquía de niveles del lenguaje (con niveles límites)
ambigüedad sistemática de "verdadero" y "falso"	distintos predicados de verdad y falsedad en cada nivel	único predicado de verdad unívoco, con aplicación extendida hasta el punto mínimo fijo
"verdadero" y "falso" completamente definidos	"verdadero" y "falso" completamente definidos	"verdadero" y "falso" sólo parcialmente definidos
"Esta oración es falsa" carece de significado	"Esta oración es falsa en O" falsa en M	"Esta oración es falsa" ni verdadera ni falsa.
base racional		
P.C.V.	(relativización al lenguaje de "verdadero")	fundamentación

3 PARADOJA SIN "FALSO"; ALGUNAS OBSERVACIONES SOBRE LA TEORÍA DE LA REDUNDANCIA DE LA VERDAD DE TARSKI; Y EL P.C.V. DE NUEVO

Me temo que no podré ofrecer, en conclusión, una nueva solución a las paradojas. El propósito de esta sección es más bien modesto: cumplir la promesa (págs. 153, 157) de comentar la consecuencia, referente a las paradojas, de la teoría de la redundancia de la verdad con su resistencia a la idea de verdad como predicado metalingüístico; no obstante, una consecuencia de las consideraciones que esta investigación pone de relieve servirá de apoyo para una propuesta que, como argumentaré, mantiene afinidades con el P.C.V.

Una de las razones por las que Tarski se niega a aceptar el tratamiento del entrecomillado como una función y así negar que la verdad se pueda definir generalizando el esquema (T) para obtener "(p)('p' es verdadero sii p)" era, si se recuerda (pág. 126), que, con las funciones de entrecomillar, la paradoja continuaría incluso sin el uso de los predicados "verdadero" y "falso". (Y el requisito de la apertura semántica de Tarski resultaría, por supuesto, impotente para vencer a la paradoja generada sin predicados semánticos.) El argumento de Tarski procede así:

> Sea "c" la abreviatura de "la oración numerada con 1".
> Ahora considérese la oración:
> 1. $(p)(c = \text{"}p\text{"} \rightarrow -p)$
> Se puede establecer empíricamente que:
> 2. $c = \text{"}(p)(c = \text{'}p\text{'} \rightarrow -p)\text{"}$
> y, por tanto, suponiendo que:
> 3. $(p)(q)(\text{"}p\text{"} = \text{"}q\text{"} \rightarrow p \equiv q)$

"mediante leyes elementales de la lógica derivamos fácilmente una contradicción" (1931, pág. 162)[3]. Obsérvese que estamos ante una paradoja que surge, no intrínsecamente de la naturaleza de un enunciado único, sino extrínsecamente, debido a que, como diría Kripke, los hechos resultan mal. El diagnóstico de Tarski es que las funciones de entrecomillar son la raíz de la dificultad y no se deben permitir. Como respuesta, algunos autores han sugerido que, en vez de rechazar por completo las funciones de entrecomillar, se deben imponer ciertas restricciones sobre las mismas; Binkley, por ejemplo, sugiere una regla de "no mezclar" (1970) que impide a un mismo cuantificador ligar a la vez variables que figuran dentro y fuera de las comillas y, por tanto, rechaza 1 de arriba. Pero ni el diagnóstico de Tarski, ni este tipo de respuesta pueden considerarse totalmente correctos; porque se puede derivar una paradoja análoga sin el uso de comillas:

> Sea "§" un operador que forma un término a partir de una oración; se podría leer, por ejemplo, "el enunciado que...". Sea "c" la abreviatura de "el enunciado constituido por la oración numerada con 1".

Considérese ahora la oración:

1. $(p)(c = \S p \rightarrow -p)$

[3] Tarski no da la derivación, pero presumiblemente vendría a ser del modo siguiente. De 1, si $c = \text{"}(p)(c = \text{'}p\text{'} \rightarrow -p)\text{"}$, entonces $-(p)(c = \text{'}p\text{'} \rightarrow -p)$, por tanto, dado 2, $-(p)(c = \text{'}p\text{'} \rightarrow -p)$; de aquí que, por RAA, -1. Si -1, entonces $(\exists p)(c = \text{'}p\text{'} \,\&\, p)$. Supongamos, por ejemplo, que $c = \text{'}q\text{'} \,\&\, q$; entonces $\text{'}q\text{'} = \text{"}(p)(c = \text{'}p\text{'} \rightarrow -p)\text{"}$ ya que ambas $= c$, de aquí que, por 3, $q \equiv (p)(c = \text{'}p\text{'} \rightarrow -p)$. Pero q; por tanto, "$(p)(c = \text{'}p\text{'} \rightarrow -p)$", i.e., 1. En consecuencia, $1 \,\&\, -1$.

Se puede establecer empíricamente que:

$$2.\ c = \S\,(p)\,(c = \S\,p \rightarrow -p)$$

y se deriva una contradicción como antes[+]. Ahora bien, se puede intentar de nuevo imponer restricciones a los operadores que forman términos como "§"; por ejemplo, siguiendo el modelo de Harman, 1971, uno puede establecer la regla de que si "p" pertenece a L, "§p" debe pertenecer, no a L, sino al metalenguaje de L. Pero este tipo de maniobra —aparte de su carácter desagradablemente *ad hoc*— parece nuevamente que no llega al núcleo del problema; porque se puede derivar una paradoja análoga sin el uso de "§"; habría que hacer que "c" fuera la abreviatura de la oración numerada con 1 (en vez de "la oración numerada con 1", "c" es ahora la abreviatura de una oración, no de un término):

$$1.\ (p)\,((c \equiv p) \rightarrow -p)$$

de modo que, en virtud de la abreviatura,

$$2.\ (c \equiv (p)\,((c \equiv p) \rightarrow -p)$$

y una vez más sería derivable una contradicción.

Esto no debe sorprender demasiado. Pues, como ha mostrado la investigación de la teoría de la redundancia (págs. 153-7), el efecto de un predicado de verdad puede conseguirse usando cuantificadores (proposicionales) de segundo orden; y añadiendo la negación produce el efecto de "falso". Por tanto, era de esperar el hecho de que se puede derivar una paradoja del tipo de la del mentiroso sin uso explícito de predicados semánticos con tal de que se disponga de cuantificadores proposicionales y negación.

Pero, ¿cómo se pueden evitar paradojas de este tipo? Supongamos que los cuantificadores proposicionales son interpretados sustitucionalmente —como recomendé en el cap. 4, § 3—. En la interpretación sustitucional, una fórmula cuantificada, A, de la forma $(v)\,\Phi\,(v)$, es verdadera, sólo en el caso de que todas sus instancias de sustitución, $\Phi(s)$, sean verdaderas. Puesto que en el caso que es-

[+] Se requieren algunos comentarios sobre la moraleja que se debe sacar de las comillas. Tarski sostiene (y Quine está de acuerdo) que el resultado de encerrar una expresión entre comillas es una expresión que denota la expresión encerrada, pero *de la cual la expresión encerrada no es parte genuinamente*. La idea de que las comillas son una especie de "bloque lógico", de que "perro" no es parte de " "perro" ", conduce a consecuencias verdaderamente curiosas y es totalmente contraintuitiva (cfr. Anscombe, 1957). Por eso, es un alivio descubrir que el fracaso del diagnóstico tarskiano de la paradoja nos deja libres para tratar las comillas como una función; cfr. Belnap y Grover, 1973; Haack, 1975, para una discusión más detallada.

tamos considerando el cuantificador liga letras oracionales, los *sustituyentes* para v serán fbfs y, por consiguiente, ellos mismos pueden contener cuantificadores. Ahora bien, las condiciones usuales de adecuación definicional requieren que solamente se permitan *sustituyentes* que contengan menos cuantificadores que A; de otro modo, se produciría la ineliminabilidad (véase Marcus, 1972, y cfr. Grover, 1973). Esta interpretación no es en modo alguno *ad hoc*, ya que solamente es un caso especial de condiciones totalmente ordinarias de las definiciones; pero, al mismo tiempo, es suficiente para bloquear el argumento paradójico donde la fbf que sustituyera "p" en "$(p)\,((c \equiv p) \rightarrow -p)$" es "$(p)\,((c \equiv p) \rightarrow -p)$".

No sería del todo caprichoso, pienso, percibir afinidades entre esta idea y la teoría de los tipos con su jerarquía de proposiciones ordenadas según qué cuantificadores proposicionales ocurran en ellas; como tampoco lo sería ver afinidades entre la motivación de la restricción a los *sustituyentes* por variables oracionales y el P.C.V. El argumento de Russell de por qué una proposición "acerca de todas las proposiciones" no puede ella misma ser miembro de esa totalidad es que "origina" una nueva proposición que no pertenecía previamente a esa totalidad, lo cual no es convincente porque se supone lo que se pretendía demostrar, i.e., que la proposición acerca de todas las proposiciones no es previamente miembro de esa totalidad: Ryle y Mackie, sin embargo, propugnan a favor del P.C.V. que las violaciones del mismo conducen a una "autodependencia viciosa" que produce la ineliminabilidad. Y, finalmente, el hecho de que las paradojas puedan ser generadas sin predicados semánticos podría parecer sugerir que, después de todo, podría haber *algo* en el presentimiento que Russell tenía de que las paradojas no deberían manejarse en grupos distintos de acuerdo con los predicados semánticos o de teoría de conjuntos que ocurriesen esencialmente en ellas, sino que deberán manejarse todas juntas como resultado de una falacia.

9
Lógica y lógicas

> ... puesto que nunca sabe uno cuáles serán las líneas del progreso, es siempre muy temerario condenar lo que no está totalmente a la moda del momento.
>
> RUSSELL, 1906, citado en Rescher, 1974

1 LÓGICA "CLÁSICA" Y "NO CLÁSICA"

Existe un número considerable de sistemas lógicos formales. De hecho, desde que se formuló el aparato lógico "clásico" siempre ha habido autores que han insistido en que se debería mejorar, modificar o reemplazar. De la historia del condicional material se puede tomar un ejemplo instructivo; "la implicación material", anticipada por los estoicos, fue formalizada por Frege en 1879 y por Russell y Whitehead en 1910, y dotada de una semántica adecuada por Post en 1921 y por Wittgenstein en 1922. Sin embargo, a principios de 1880 MacColl había insistido en la necesidad de un condicional más estricto; "la implicación estricta" fue formalizada por Lewis en 1918; y, después, el descontento con sus pretensiones de representar el entrañamiento condujo a la introducción de "la implicación relevante" (véase cap. 10, § 7).

Mi propósito actual es conseguir alguna perspectiva dentro de la gran variedad de sistemas lógicos, abordar cuestiones tales como la de qué forma se relacionan entre sí dichos sistemas, si se debe elegir entre ellos y, si es así, cómo. Mi estrategia consistirá en considerar los diversos modos en que ha sido modificado el aparato lógico estándar y las diversas presiones que han motivado las modificaciones llevadas a cabo. Sin embargo, se debería establecer una nota inicial de precaución: esta estrategia de considerar el aparato lógico "no estándar" en contraste con el "estándar" lleva consigo el peligro de inducir a una actitud demasiado conservadora respecto a las innovaciones lógicas (Wolf, 1977, expresa bien la cuestión al recordar que "la posesión es nueve décimas de la ley"). Espero que el conocimiento de dicho peligro pueda por sí mismo ayudar a evitarlo en alguna medida. Y será también provechoso tener presente que "la lógica clásica" de hoy fue ella misma en su día una "innovación lógica". Kant, después de todo, sostuvo (1800) que la lógica era una ciencia completa, acabada en lo esencial, en la obra de Aristóteles; el siglo siguiente, sin embargo, contempló el desarrollo de nuevas técnicas lógicas más fuertes y rigurosas con los trabajos de Boole, Peirce, Frege y Russell. Recuérdese también que Frege había supuesto confiadamente que los principios de su sistema lógico eran autoevidentes hasta que Russell mostró que eran inconsistentes.

2 RESPUESTAS AL APREMIO A CAMBIAR EL FORMALISMO ESTÁNDAR

Las presiones para cambiar el cálculo estándar bivalente de oraciones y de predicados de primer orden han provenido de las preocupaciones acerca de la aparente inadecuación del aparato estándar para representar los diversos tipos de argumento informal y acerca de la interpretación y aplicación de dicho aparato. Las reacciones a tales presiones han sido muy variadas; primeramente esbozaré y luego ilustraré algunas de las respuestas más comunes:

1. Los argumentos informales a los que no se aplica adecuadamente el aparato estándar pueden excluirse del ámbito de la lógica. Por ejemplo, el apremio por una "lógica de la carencia de significado" puede oponer resistencia en virtud de que las oraciones carentes de significado se encuentran sencillamente fuera de la propia esfera de la formalización lógica. A ésta la llamaré la respuesta de la *delimitación del ámbito de la lógica*.
2. Se pueden admitir los argumentos informales problemáticos como pertenecientes al campo de la lógica y mantener el aparato estándar; pero haciendo reajustes de modo que los argumentos informales difíciles sean representados en el formalismo. Por ejemplo, la teoría de las descripciones de Russell propone que las oraciones que contengan descripciones definidas sean representadas no de la forma obvia, como "*Fa*", sino como fórmulas cuantificadas existencialmente. A ésta la llamaré la estrategia de la *nueva paráfrasis*. (Puesto que Russell comenta que la forma gramatical de tales oraciones oculta su forma lógica, en 1974 la denominé la estrategia de la *forma engañosa*. Pero he preferido que no parezca

que suscribo su punto de vista de que cada argumento posee "una forma lógica" única).

3. Una tercera respuesta admite, como la segunda, los argumentos problemáticos como pertenecientes al campo de la lógica y mantiene el aparato estándar sin ningún cambio a nivel sintáctico; sin embargo, la interpretación de dicho aparato se modifica de tal manera que las locuciones informales inicialmente recalcitrantes son, después de todo, adecuadamente representadas. Por ejemplo, respecto al aparente compromiso ontológico del cálculo de predicados se puede responder con la propuesta de que los cuantificadores se interpretan sustitucionalmente y se permiten los términos vacíos como *sustituyentes bona fide*, de tal manera que se asegura la neutralidad ontológica. A ésta la llamaré la respuesta de la *innovación semántica*.
4. El aparato estándar puede extenderse para obtener un formalismo aplicable a argumentos informales que antes eran inaccesibles al tratamiento formal. Por ejemplo, se pueden añadir nuevos operadores —operadores temporales o modales, por ejemplo— y axiomas/reglas que los gobiernen; o pueden extenderse las operaciones estándars para cubrir nuevos ítems —por ejemplo, oraciones imperativas o interrogativas—. A ésta la llamaré la respuesta de la *lógica extendida*.
5. Alternativamente, se puede restringir el aparato estándar en el sentido de que, mientras permanece el mismo vocabulario, sus axiomas/reglas de inferencia son restringidos de tal modo que los teoremas/inferencias clásicos pierden su validez. Por ejemplo, la preocupación por evitar anomalías en la mecánica cuántica ha conducido a propuestas en las que ya no se dan ciertos principios "clásicos", como la ley distributiva, por ejemplo. A ésta la llamaré la respuesta de la *lógica restringida*; el resultado es una "lógica divergente" (Haack, 1974).

A veces se proponen nuevas formulaciones que extienden y restringen a la vez la lógica clásica —añaden nuevos operadores y nuevos principios para gobernarlos, pero al mismo tiempo restringen los principios que gobiernan los antiguos operadores. Serían un ejemplo de esto "las lógicas de la relevancia" que introducen un nuevo condicional, mientras que al mismo tiempo rechazan algunas leyes clásicas, tales como el *modus ponens* para el condicional material.

He distinguido 4 y 5 de 2 y 3 debido a que los dos primeros llevan consigo modificaciones en el nivel sintáctico, mientras que los dos últimos dejan la sintaxis estándar intacta. Pero, por supuesto, una extensión o restricción de la sintaxis estándar requeriría a su vez una modificación semántica, de modo que se suministra una interpretación que verifique el conjunto extendido o restringido de teoremas/inferencias. En realidad, las restricciones de la lógica han sido motivadas muy a menudo por consideraciones semánticas —como, por ejemplo, los retos a la suposición de que toda oración dentro del campo de la lógica debe ser o verdadera o falsa condujeron al desarrollo de las lógicas plurivalentes que característicamente carecen de teoremas clásicos como "$p \vee -p$".

Como es de esperar, las extensiones son propuestas más habitualmente como respuesta a una supuesta *inadecuación* y las restricciones como respuesta a una supuesta *incorrección* en el formalismo estándar.

6. Las innovaciones en el formalismo lógico van a veces acompañadas por —y están a veces motivadas por— innovaciones en el nivel de los conceptos metalógicos. Por ejemplo, los intuicionistas (que propone una restricción del aparato estándar) lo hacen en parte porque recusan el concepto de verdad presupuesto en la lógica clásica; los lógicos de la relevancia recusan la concepción clásica de validez. A ésta la llamaré la *recusación de los metaconceptos clásicos*.
7. Finalmente —y, por decirlo así, como la inversa de la primera respuesta— hay retos a la concepción estándar del ámbito y aspiraciones de la lógica. Éstos van con bastante frecuencia asociados con los retos a los metaconceptos clásicos, como en 6. Por ejemplo, los intuicionistas no sólo restringen el cálculo de oraciones clásico de tal manera que, por ejemplo, "$p \vee -p$" deja de ser un teorema, y no sólo ofrecen una alternativa a la concepción clásica de la verdad; mantienen también un punto radicalmente diferente del de la mayoría de los lógicos clásicos en cuanto al papel de la lógica a la que consideran como secundaria respecto de la matemática, más bien que como un razonamiento que sirve de base a todo tipo de materias. A ésta la llamaré la respuesta de *la revisión del ámbito de la lógica* (Un intuicionista, sin embargo, consideraría que el lógico clásico hace revisión del ámbito de la lógica.)

Hablando de un modo aproximado, creo que hay razón para considerar a estas respuestas gradualmente más radicales. Pero eso *sólo* es hablando en forma aproximada. Por ejemplo, aunque usualmente se piensa que una reinterpretación del aparato estándar es más conservadora que una extensión del mismo, lo cierto es que hay un sentido en el cual el conservadurismo de 3 es *nominal* —quiero decir que el sistema solamente *parece* el mismo y, una vez que ha sido reinterpretado, el resultado difiere muy poco de la introducción del nuevo simbolismo. Vale la pena observar, por ejemplo, que algunos, en atención a la claridad, insisten en que usemos para los cuantificadores sustitucionales una notación diferente de la de los cuantificadores objetuales. He llamado la atención sobre el modo en que retos bastante serios a los metaconceptos clásicos o a las concepciones clásicas de la finalidad de la formalización pueden estar con frecuencia a la base de propuestas para extender o restringir el formalismo clásico; en vista de ello, no es nada sorprendente, como veremos, que tales sistemas no hayan sido a veces considerados realmente como *lógica* por los conservadores.

Más adelante (cap. 12) utilizaré las distinciones hechas aquí para intentar comprender los problemas epistemológicos surgidos de la existencia de una pluralidad de lógicas. Por el momento, sin embargo, mi preocupación fundamental es proporcionar algún tipo de marco para examinar esa pluralidad. Las estrategias 1-7 no son exclusivas (ni, probablemente, exhaustivas); es de cierto interés observar que algunos problemas, por ejemplo, los problemas originados de la posibilidad de términos singulares no denotativos, han motivado varias de estas estrategias: Strawson propone que las oraciones que contienen tales términos se excluyan del ámbito de la lógica, Russell propone que se proporcione una nueva traducción que revele su forma lógica real, Hintikka que se cree una lógica restringida.

Puesto que no me es posible considerar todos los problemas surgidos del hecho de elegir entre estas estrategias, en lugar de eso realizaré un examen más detenido de dos ejemplos que ilustran muy bien algunos de los problemas. Comenzaré con el problema de cómo manejar formalmente el *tiempo*.

3 Estudio del primer caso: la lógica del discurso temporal

Los pioneros de la lógica formal moderna fueron motivados principalmente por el deseo de representar los argumentos matemáticos de forma rigurosa. En consecuencia, y debido a la irrelevancia de las consideraciones sobre el tiempo respecto a la (in)validez de los argumentos matemáticos, pudieron ignorar en gran parte el hecho de que en los argumentos informales sobre asuntos no matemáticos el tiempo es a veces crucial.

Mientras que este problema es con bastante frecuencia descartado —junto con problemas afines sobre expresiones indexicales— con el comentario de que, al representar argumentos informales en forma simbólica, se debe cuidar que el tiempo permanezca constante a través del argumento (una especie de versión casual de la respuesta del no ítem), algunos autores han realizado intentos más serios para tener en cuenta el tiempo. Y han sido propuestas dos estrategias completamente distintas: Quine insiste en que el discurso

temporal se represente dentro del aparato estándar interpretando las variables del cálculo de predicados como fluctuando no sobre individuos que perduran espacio-temporalmente, sino sobre "épocas"; Prior insiste en que el discurso temporal se acomode mediante una extensión del aparato estándar añadiendo operadores temporales.

Así que, por una parte, Quine propone que se trate el problema mediante una innovación semántica mientras que, por otra parte, Prior propone una lógica extendida. La otra diferencia entre las dos estrategias es importante: aunque ambas son intentos de acomodar las consideraciones del tiempo, el enfoque de Prior lo lleva a cabo considerando seriamente el tiempo, en tanto que el enfoque de Quine intenta conseguir el mismo fin en un formalismo sin tiempo. Y una consecuencia de ello es que Quine necesita hacer reajustes en la forma en que el discurso informal temporal es formalmente representado así como en la forma de reajustarse el formalismo. Es decir, que su enfoque combina la innovación semántica con la estrategia de la nueva paráfrasis.

El enfoque de Quine (1960a, § 36; sus ideas han sido desarrolladas más detalladamente en Lacey, 1971, en quien también me apoyaré) consiste en representar lo que es lógicamente relevante en el discurso temporal de los argumentos informales dentro del formalismo lógico estándar. Aunque Quine admite la relevancia del tiempo para la validez de los argumentos informales, en realidad lo considera como no esencial, como un reflejo de la propensión del lenguaje ordinario hacia la perspectiva temporal del hablante. Por tanto, propone que los verbos con tiempos sean reemplazados por verbos sin tiempos con "cualificadores temporales", tales como "ahora", "entonces", "antes de *t*", "en *t*" y "después de *t*". Las variables "*t*", "*u*", ..., etc., se construyen para fluctuar sobre lo que Quine denomina "épocas", que son períodos espaciotemporales de cualquier duración elegida, por ejemplo, de una hora, o un día. Una época, explica Quine, es un "trozo del mundo material de cuatro dimensiones espacialmente exhaustivo y perpendicular al eje del tiempo" (1960a, pág. 172). La referencia a individuos espaciotemporales, como personas, es reemplazada por la referencia a "trozos temporales" de individuos, como una persona a través de un tiempo dado. Por tanto, las oraciones temporales ordinarias quedan reformuladas de la siguiente manera:

María es viuda	María *es* viuda ahora
Jorge se casó con María	$(\exists t)$ (*t es* antes de ahora y Jorge en *t se casa* con María en *t*)
Jorge se casará con María	$(\exists t)$ (*t es* después de ahora y Jorge en *t* se *casa* con María en *t*)

Las convenciones notacionales son que los verbos sin tiempo han de reescribirse en la forma temporal del presente, pero en cursiva;

la variable "t" ha de fluctuar sobre épocas; "Jorge en t" y "María en t" se refieren a trozos temporales de los individuos espaciotemporales Jorge y María.

Por supuesto, "ahora" conserva el carácter indexical del lenguaje ordinario; pero eventualmente Quine eliminará esto también por medio de términos singulares que denotan épocas. Así, "ahora" será reemplazado por la fecha apropiada, y se eliminará el último vestigio del discurso temporal, como:

María es viuda	María *es* viuda el 12 de marzo de 1977

El resultado es que las oraciones temporales cuyo valor de verdad varía con el tiempo son sustituidas por las que Quine llama *oraciones eternas* cuyo valor de verdad permanece constante. (Las oraciones eternas, por supuesto, son la respuesta de Quine a la supuesta necesidad de las proposiciones a las que, en vista de su carácter intensional, Quine no quiere admitir.)

Debería quedar claro por ahora que la propuesta de Quine exige considerables desviaciones de las locuciones del lenguaje ordinario, así como considerables innovaciones en la interpretación de las variables, términos singulares y predicados del cálculo de predicados. Sin embargo, Quine consideraría su propuesta como *conservadora* en un sentido importante; porque su propósito es permitir la representación del discurso temporal dentro de un formalismo *extensional*. Es por esto por lo que Quine —que considera la extensionalidad como la piedra de toque de la inteligibilidad— concede tanta importancia al mantenimiento de la sintaxis estándar.

Sin embargo, Quine piensa que su propuesta posee además otra virtud: su consonancia con la física moderna. Pues, mientras el discurso temporal ordinario *distingue* el tiempo, las representaciones de Quine tratan a la dimensión temporal completamente igual que a las tres dimensiones espaciales. Las partes temporales de una cosa son tratadas de la misma manera que sus partes espaciales (tema éste que Quine desarrolla (pág. 171) argumentando cómo su enfoque ilumina el problema de la identidad personal: ¿por qué deberíamos esperar que las partes temporales de una persona sean iguales, si sus partes espaciales, por ejemplo, su cabeza y sus pies, no lo son?). Los descubrimientos de Einstein, comenta Quine, "no dejan ninguna alternativa razonable para tratar el tiempo como igual al espacio" (pág. 172).

El enfoque de Prior (véase 1957, 1967, 1968) es diferente e interesante. Acomoda las consideraciones temporales, no mediante el reajuste de las locuciones temporales del lenguaje ordinario para que se adecúen a un simbolismo extensional sin tiempo, sino mediante la extensión del simbolismo estándar para acomodar las locuciones temporales. Prior parte de un cálculo de oraciones usual en el que, sin embargo, las letras oracionales se interpretan como representativas uniformemente de oraciones en el tiempo presente. (Y, consecuentemente, de ítems vulnerables para cambiar de valor de verdad en contraste con las oraciones sin tiempo y eternas de Quine.) Después, él enriquece el simbolismo con los operadores temporales "F" y "P", que son operadores para formar oraciones sobre oraciones cambiando el primero una oración de tiempo presente en una oración de tiempo futuro y el segundo cambia una oración de tiempo presente en una oración de tiempo pasado. Prior lee "F" como "Será el caso que..." y "P" como "Solía ser el caso que...". Los tiempos compuestos se construyen mediante iteración de estos operadores. Por ejemplo, si "p" es "Jorge se casa con María", tenemos:

Jorge se casó con María	Pp
Jorge se casará con María	Fp
Jorge se habrá casado con María	FPp

Los operadores temporales no son extensionales; el valor de verdad de "Fp" o de "Pp" no depende solamente del valor de verdad de "p".

Se suministran axiomas para gobernar a los nuevos operadores. De hecho, Prior ofrece conjuntos alternativos de axiomas, sugiriendo que cada uno se ajusta a puntos de vista metafísicos rivales sobre el tiempo, tales como el de si el tiempo tiene un principio y/o un fin, si es lineal o circular, si el determinismo es verdadero, etc. (véase Prior, 1968).

Prior observa que los operadores temporales, en lugar de tomarse como primitivos, podrían ser definidos en términos de cuantificación sobre instantes de tiempo; por ejemplo, "Será el caso que p" sería "Para algún tiempo t después de ahora, p en t"[1]. Tal vez esto reduzca en cierto grado el contraste con el enfoque de Quine. Pero "los instantes" son temporales, no espaciotemporales como las "épocas" de Quine. Y Prior dice (1968, pág. 118) que de todos modos prefiere considerar a los operadores temporales como primitivos y a los instantes de tiempo como "meras construcciones lógicas a partir de hechos temporales".

Así pues, el enfoque de Prior logra la simplicidad de la paráfrasis de los argumentos informales en el simbolismo formal, pero al mismo tiempo incrementa la complejidad del formalismo que re-

[1] Las lógicas temporales de Prior han sido constituidas en estrecha relación con los sistemas modales de C. I. Lewis (cfr. cap. 10); y la definibilidad de los operadores temporales mediante cuantificadores sobre instantes corresponde en las semánticas usuales de dichas lógicas modales a la explicación de la necesidad (posibilidad) como verdad en todos (algunos) los mundos posibles.

quiere particularmente la pérdida de la extensionalidad. Y contrasta también con el enfoque de Quine en el nivel metafísico; pues, aunque se ofrezcan conjuntos alternativos de axiomas para elegir entre ellos según el punto de vista que uno adopte sobre el tiempo, la sintaxis misma del sistema está de acuerdo con la concepción "newtoniana" del tiempo como totalmente diferente del espacio. Los principales puntos de contraste entre los dos enfoques están resumidos en la tabla 2.

TABLA 2

Enfoque de Quine	*Enfoque de Prior*
estrategias de la innovación semántica y de la nueva paráfrasis	lógica extendida
elimina el tiempo	introduce operadores temporales
oraciones eternas, ningún cambio del valor de verdad	oraciones temporales, cambio del valor de verdad
formalismo extensional	formalismo intensional
"reglamentación" rígida y sustancial de los argumentos informales	conformidad con el lenguaje ordinario
en consonancia con la teoría de la relatividad	newtoniano en espíritu
ontología del mundo espacio-temporal de 4 dimensiones	ontología de objetos que ocupan el espacio y duran a través del tiempo

He señalado que el tratamiento de Prior está más en consonancia con un punto de vista que considere el tiempo como categóricamente diferente del espacio, y el de Quine con un punto de vista que considere el tiempo como igual al espacio. No es sorprendente, por tanto, que se haya sugerido a veces que hay razones metafísicas para preferir un enfoque a otro[2]. Como ya he relatado, Quine

[2] Cfr. MacTaggart, 1908, donde se traza una distinción entre la "serie *A*", en la que los eventos están ordenados según si son pasados, presentes o futuros, y la "se-

piensa que la ciencia moderna "no deja ninguna alternativa" razonable a su enfoque. Geach, por otra parte, argumenta (1965) que la ontología que ofrece Quine de las épocas y de los objetos espacio-temporales de 4 dimensiones es defectuosa porque entraña la negación del cambio. Pero esto es falso; el enfoque de Quine admite el cambio en efecto, es precisamente para representar lo que ordinariamente se denominaría cambio en un objeto que perdura en el tiempo como la diferencia entre los trozos de tiempo anteriores y posteriores de ese objeto —como, por ejemplo, mi pelo se está volviendo gris se representaría por una diferencia en el color del pelo de mis trozos de tiempo anteriores y posteriores.

Mi preocupación actual, sin embargo, no se dirige a estos problemas metafísicos, sino hacia algunas cuestiones metodológicas originadas por la elección de estrategias.

En general, como en el caso presente, parece razonable esperar que el precio de ceñirse (como Quine) a un simbolismo austero será una pérdida de la naturalidad de la paráfrasis de los argumentos informales. (Para decirlo en términos russellianos: cuanto menos formas lógicas disponibles haya, más formas gramaticales tendrán que ser diagnosticadas como "engañosas".) Si se concede gran significación a algún grado de austeridad —en el caso de Quine, a la extensionalidad— en el formalismo, habrá que aceptar una divergencia del lenguaje natural. Si se concede gran significación a la conformidad con las formas del lenguaje natural —como hace Geach— se necesitará un formalismo más rico. Por lo que a mí se refiere, admito la conveniencia tanto de la austeridad del simbolismo (después de todo, parte de la finalidad de la formalización es sistematizar, tener relativamente pocas reglas para cubrir relativamente muchos casos) como la simplicidad de la paráfrasis (porque otra parte de la finalidad de la formalización es proporcionar una técnica para evaluar los argumentos informales); temo que sea precisamente un hecho de la vida lógica el que éstos sean *desiderata* en competencia.

Un factor que a veces puede ayudar a decidir tal competencia es que un sacrificio de austeridad de formalismo o de simplicidad de la paráfrasis se justificará mejor cuanto más amplio sea el conjunto de ventajas que mediante él se obtienen. Por ejemplo, se podría esperar que el formalismo equipado para enfrentarse con el discurso temporal pudiera también ser capaz de representar el discurso sobre la acción y el discurso sobre la causación —y claramente, si en este caso sólo tuviese éxito un enfoque, eso sería razón

rie *B*", en la cual están ordenados como anteriores, simultáneos o posteriores entre sí. El enfoque de Prior acentúa la primera, el de Quine la segunda. Véase también Strawson, 1969, para una defensa de la primera, y Whitehead, 1919, para una defensa de la postura metafísica segunda.

suficiente para prefirirlo. (Véase Lacey, 1971, para una discusión relevante; y cfr. Davidson, 1968a, donde se argumenta que para representar enunciados causales y de acción se necesita cuantificar sobre eventos. Recuérdese (pág. 147) que Davidson, lo mismo que Quine, está comprometido a limitarse a un formalismo extensional.)

Quine apela al carácter de las teorías físicas corrientes para sostener su enfoque; Geach insiste a favor de Prior en que es completamente impropio reajustar la lógica para adaptarla a la ciencia. Las cosas aquí están muy enredadas. La actitud de Geach se deriva en parte del hecho de que aparentemente considera la teoría de la relatividad como incoherente, ya que conlleva la negación de lo que él considera como una "diferencia de categoría" entre el espacio y el tiempo. Y su convicción de que *hay* una tal diferencia de categoría deriva a su vez de nuestros conceptos ordinarios de espacio y tiempo como formando parte de nuestro discurso temporal ordinario. Aquellos que, como yo, admiten que los desarrollos de la física pueden muy bien llevar a la revisión conceptual, se opondrán a este fácil diagnóstico de la teoría de la relatividad como "conceptualmente confusa".

Pero, completamente aparte de la cuestión de la coherencia o, en otro caso, de la física relativista, hay un asunto más profundo en discusión. Quine aspira expresamente en su elección del formalismo lógico a "un lenguaje adecuado para la ciencia" y, por así decirlo, considera la lógica como continua a la ciencia; Geach considera la lógica como autónoma de la ciencia y en realidad anterior a ella. La historia de la lógica ofrece cierto apoyo para este último punto de vista; por ejemplo, la lógica ideada por Frege y Russell, a diferencia de la silogística de Aristóteles, puede expresar relaciones así como propiedades; y es precisamente por esta superioridad de poder expresivo por lo que la lógica moderna es capaz —como la lógica aristotélica no lo era— de representar tipos de argumentos esenciales para la matemática moderna. Es preciso, sin embargo, distinguir la cuestión del poder expresivo de la lógica de la cuestión de su contenido doctrinal; quiero decir que, mientras parece irreprochable modificar el poder expresivo de un formalismo con el fin de habilitarlo para que exprese estilos de argumentos característicos de la ciencia, es un asunto más serio el abandonar una supuesta ley de la lógica (como, por ejemplo, los lógicos cuánticos que instan a que se abandone la ley distributiva) debido a los desarrollos de la ciencia. Esto sugiere que, epistemológicamente hablando, las extensiones de la lógica son menos radicales que sus restricciones: punto éste al que volveré en el cap. 12.

4 ESTUDIO DEL SEGUNDO CASO: PRECISIÓN VERSUS "LÓGICA VAGA"

Una gran parte del discurso informal en cierta medida es vago. Y por ello surge la cuestión de si los lógicos deberían tener en cuenta este hecho y, si es que sí, ¿cómo?

El primer aspecto a tomar en consideración es que una razón importante para construir sistemas formales de lógica es proporcionar cánones *precisos* de validez —una mayor ventaja de la lógica formal sobre el argumento informal no reglamentado rígidamente es su mayor rigor y exactitud. En vista de ello, no resulta sorprendente que Frege y Russell considerasen la vaguedad como un defecto de los lenguajes naturales y que debería eliminarse del lenguaje formal aceptable. (Y no cabe duda de que también es relevante aquí, en cuanto a su negligencia respecto de las consideraciones temporales, que ellos se dedicaron principalmente a la formalización del argumento matemático.)

Esto sugiere, quizás, que sería apropiado sencillamente excluir las oraciones vagas como inelegibles para el tratamiento lógico. Pero pienso que esta estrategia es demasiado tosca, porque está claro que las oraciones vagas pueden ocurrir en argumentos informales sin amenazar su validez. Hay aquí un contraste importante con el caso de las oraciones carentes de significado. Un argumento debe estar compuesto de oraciones significativas: una cadena de símbolos sin significado no *sería* un argumento, y una secuencia de oraciones significativas con una cadena carente de significado interpuesta, si fuese considerada en todo como un argumento, sería válida o inválida independientemente de la cadena carente de significado. Por tanto, es completamente razonable excluir las oraciones carentes de significado del ámbito de la lógica: "las lógicas de la carencia de significado" (por ejemplo, Halldén, 1949; Routley, 1966, 1969) no son, a mi entender, ni necesarias ni deseables[3]. Pero una oración vaga *puede* desempeñar un papel genuino en un argumento ("A Juan le gustan las chicas capaces; María es una chica capaz e inteligente; por tanto, a Juan le gustará María"); y, consecuentemente, los lógicos deben tomar más en serio la vaguedad.

Sin embargo, las oraciones vagas parecen presentar ciertas dificultades para la aplicación del aparato lógico estándar. Se supone que los sistemas lógicos formales son relevantes para la valoración de los argumentos informales; pero los sistemas lógicos clásicos, en los que toda fbf es o verdadera o falsa, no parecen apropiados para la valoración de los argumentos informales con premisas y/o

[3] No pretendo negar que puede haber algunas cuestiones filosóficas interesantes sobre el carácter y fuentes de la carencia de significado (considérese, por ejemplo, el papel que desempeña la susodicha carencia de significado producida por los "errores de categorías" en Ryle, 1949).

conclusiones que, debido a su vaguedad, duda uno en llamarles o bien definitivamente verdaderas o bien definitivamente falsas. Una vez que el problema ha sido planteado de esta manera, parece haber dos enfoques naturales para su solución: depurar los argumentos informales vagos antes de someterlos a la valoración de los estándars de la lógica clásica bivalente, o idear algún sistema lógico formal alternativo que se aplique a ellos más directamente.

El primer enfoque exige que los argumentos informales sean reglamentados rígidamente de modo que se pueda usar el aparato lógico estándar. (El procedimiento vendría a ser bastante análogo a las acomodaciones regularmente hechas para dar cuenta de las discrepancias entre las conectivas veritativo-funcionales y sus lecturas en castellano.) Carnap propone (1950, cap. 1) lo que él llama un programa de precisión: antes de la formalización, lo vago deberá sustituirse por lo preciso, por ejemplo, los predicados cualitativos por los comparativos o cuantitativos, de tal manera que (usualmente, pero no invariablemente) los términos precisos correspondan en extensión en todos los casos claros y centrales con los vagos a los que ellos sustituyen, pero que tengan también una aplicación bien definida en los casos que eran dudosos para los términos vagos. Esta propuesta contiene elementos de las estrategias 1 y 3 señaladas en el § 2: los argumentos informales son depurados antes de recibir la representación formal (estrategia 2), pero de tal modo que los argumentos reglamentados rígidamente eviten siempre la vaguedad de los originales (recomendación de la estrategia 1).

Algunos autores (por ejemplo, Russell, 1923, y Black, 1937) han insistido en que los lenguajes naturales son completamente vagos; y, por supuesto, si esto fuera así, el programa de Carnap no podría llevarse a cabo. Sin embargo, no se han ofrecido argumentos muy convincentes de por qué la precisión es imposible en principio (cfr. Haack, 1974, cap. 6) y procederé dando por supuesto que la precisión es factible.

Pero concediendo que la precisión sea posible, ¿es deseable? Cierto apoyo en favor de una estrategia diferente —de alterar la lógica clásica para adaptarla al argumento informal, en lugar de adaptar el argumento informal a la lógica clásica— se ha derivado de la creencia de que la especie de refinamiento sucesivo de los conceptos científicos subrayados por Carnap puede resultar que sea de restringida aplicabilidad y de complejidad poco manejable. En realidad, es significativo que el autor responsable de las propuestas que más han influido en pro de una lógica revisada de la vaguedad sea un ingeniero eléctrico cuyas primeras investigaciones (Zadeh, 1963, 1964) habían sido dedicadas a refinar conceptos tales como "estático" y "adaptativo", pero que eventualmepte terminó sacando la conclusión (Zadeh, 1972) de que " 'el pensamiento vago', después de todo, no puede ser deplorable si posibilita la solución

de problemas que son excesivamente complejos para un análisis preciso". No es nueva la idea de que el aumento de la precisión puede que no sea una pura bendición; Duhem señaló (1904, páginas 178-9) que los enunciados de la física teórica, precisamente porque son más precisos, son menos ciertos y más difíciles de confirmar que los enunciados más vagos del sentido común. Popper (1961, 1976) ha sugerido también que la precisión puede ser un "ideal falso".

¿Cuál es la alternativa a la precisión? Pues bien, si los argumentos informales no van a ser reglamentados rígidamente de modo que pueda aplicarse el aparato lógico clásico, quizás el aparato lógico se pueda modificar de tal modo que pueda aplicarse a los argumentos informales no reglamentados rígidamente. Se ha sugerido, por ejemplo, que una lógica trivalente sería más conveniente que la lógica clásica bivalente (Körner, 1966); la idea es que la dificultad con predicados vagos como "alto" radica en que hay casos dudosos, i.e., casos cuyo predicado no es claramente verdadero ni claramente falso, y que este problema pueda solucionarse admitiendo una tercera categoría distinta de "verdadero" y "falso" para acomodar los casos dudosos. Pero esto de ninguna manera soluciona satisfactoriamente el problema; pues se requiere trazar una línea clara entre los casos dudosos y los casos centrales verdaderos o falsos. Con todo, ciertamente el insistir en que a determinada altura un hombre deja de ser un caso dudoso y pasa a ser definitivamente alto o el insistir en que a determinada altura un hombre deja de ser no alto y pasa a ser definitivamente alto impone una precisión artificial.

Zadeh recomienda también que se adopte una lógica no estándar, pero su "lógica vaga" representa una desviación mucho más radical de la clásica. En primer lugar, esbozaré concisamente las características formales más destacables de la lógica vaga. (Para más detalles, el lector puede consultar a Zadeh, 1975, y el estudio de Gaines, 1976.) La lógica no estándar de Zadeh está ideada sobre la base de una teoría no estándar de conjuntos, una teoría "vaga" de conjuntos. Mientras que en la teoría clásica de conjuntos, un objeto o es o no es miembro de un conjunto dado, en la teoría vaga de conjuntos la pertenencia es cuestión de grado; el grado de pertenencia de un objeto en un conjunto vago se representa por medio de un número real entre 0 y 1, donde 0 denota *nula* pertenencia y 1 pertenencia *completa*. (Un conjunto vago constará, por tanto, de todos los objetos que pertenezcan a él en cualquier grado, y dos conjuntos vagos serán idénticos si pertenecen a ellos los mismos objetos en el mismo grado.) Ahora bien, la teoría vaga de conjuntos puede utilizarse para caracterizar semánticamente a la lógica no estándar; como valores de letras oracionales tenemos, en lugar de los dos valores clásicos, un número innumerable de valores del intervalo [0, 1] y las conectivas de oraciones pueden asociarse con las operaciones de teoría de conjuntos en la forma usual (por ejemplo, la negación con la complementación de conjuntos, la implicación con la inclusión de conjuntos, etc.). El resultado es una lógica de un número innumerable de valores. El carácter exacto de esta lógica dependerá de la caracterización de las operaciones de la teoría vaga de conjuntos; un conjunto totalmente natural de suposiciones produce la extensión de un número innumerable de valores de la lógica trivalente de Lukasiewicz (pág. 231). La lógica vaga es construida sobre la base de una u otra lógica de un número innumerable de valores. Hay, por tanto, una familia de lógicas vagas, cada una con su propia lógica básica. El innumerable número de valores de verdad de la lógica básica es reemplazado por el enumerable número de valores de verdad vagos, que son subconjuntos vagos del conjunto de valores de la lógica básica caracterizados como:

> *verdadero, falso, no verdadero, muy verdadero, no muy verdadero, más o menos verdadero, bastante verdadero, no muy verdadero y no muy falso...*
> (Zadeh, 1975, pág. 410)

Verdadero es definido como un subconjunto vago especificado del conjunto de valores de la lógica básica, y luego se definen los otros valores de verdad lingüísticos; *muy verdadero*, por ejemplo, es *verdadero*^{2}; si el grado de verdad 0,8 pertenece a *verdadero* con el grado 0,7, pertenece a *muy verdadero* con el grado 0,49.

Lo que esto significa en el nivel intuitivo viene a ser algo así como lo siguiente. Un predicado vago se toma para determinar, no un conjunto clásico, sino un conjunto vago; por ejemplo, una persona *a* puede ser alta en cierto grado. Si, por ejemplo, *a* pertenece con el grado 0,3 al conjunto de personas altas, entonces la oración "*a* es alto" recibiría en la lógica básica el valor 0,3 ("*x* es alto" es verdadero con el grado *n* si $x \in alto$ con el grado *n*). Pero, según Zadeh, "verdadero" es en sí mismo vago y, por tanto, recibe análogo tratamiento: el grado de verdad que tiene "*p*" puede ser muy bajo, bastante alto, no muy alto..., etc. Los valores de verdad lingüísticos de la lógica vaga pueden considerarse como correspondientes en la lógica básica a los grados de verdad bastante bajo ("*no muy verdadero*"), bastante alto ("*muy verdadero*"), no muy alto ("*más o menos verdadero*"). Por tanto, volviendo al ejemplo, si $a \in alto$ con el grado 0,3, de modo que "*a* es alto" tiene el valor 0,3 en la lógica básica, tendrá, por ejemplo, el valor de *no muy verdadero* en la lógica vaga, ya que su grado de verdad es bastante bajo.

Brevemente, se podría considerar a la lógica vaga como el resultado de dos estadios de "vaguedad": el paso de la lógica bivalente a la lógica de un número innumerable de valores como el resultado de admitir grados de pertenencia a conjuntos denotados por predicados del lenguaje objeto, y el paso a un número enumerable de valores de verdad vagos como el resultado de tratar el predicado metalingüístico "verdadero" como vago en sí mismo. El término "lógica vaga" es a veces utilizado también por las lógicas básicas no estándar; pero he seguido el propio uso más restringido de Zadeh, en el cual la "lógica vaga" denota una familia de sistemas con valores de verdad vagos. Y, según Zadeh, el segundo estadio de "vaguedad" tiene consecuencias radicales. Entre las más notables —por no decir alarmantes— están éstas. Sucede que el conjunto de valores de verdad de la lógica vaga no queda encerrado bajo las operaciones de negación, conjunción, disyunción e implicación: por ejemplo, la conjunción de dos oraciones que poseen cada una un valor de verdad lingüístico en ese conjunto puede en sí misma no poseer tal valor. Por tanto, la lógica vaga tiene "valores de verdad vagos..., tablas de verdad imprecisas... y... reglas de inferencia cuya validez es aproximada en vez de exacta" (1975, pág. 407). Por consiguiente, afirma Zadeh, en la lógica vaga asuntos tan tradicionales como la axiomatización, procedimientos de pruebas, consistencia y completud son solamente "periféricos" (Zadeh y Bellman, 1976, pág. 151). Resumiendo, la lógica vaga no es precisamente una lógica para manejar argumentos en los cuales ocurren esencialmente términos vagos; *en sí misma* es imprecisa. Es por esta razón por lo que dije que la propuesta de Zadeh es mucho más radical que cualquier otra previamente discutida; pues desafía a las ideas profundamente atrincheradas sobre los objetivos y métodos característicos de la lógica. Para los pioneros de la lógica formal, una gran parte de la finalidad de la formalización era que únicamente así se podría tener esperanza de poseer cánones *precisos* de razonamiento válido. Zadeh propone que la lógica transija con la vaguedad.

Nos enfrentamos aquí con un ejemplo notable de la estrategia 7, un desafío radical a la concepción del ámbito y finalidad de la lógica formal. De hecho, hemos visto que las respuestas a la vaguedad han tomado todas las posturas, desde las más conservadoras (los intentos de excluir totalmente las oraciones vagas del ámbito de la lógica) pasando por las moderadamente innovadoras (las propuestas de una lógica trivalente de la vaguedad) hasta la más radical (la propuesta de que la lógica abandone sus aspiraciones a la precisión).

La precisión es ciertamente un *desideratum* demasiado central e importante de la formalización para abandonarlo fácilmente. Y en el momento actual pienso que está justificado preguntarse si se puede esperar que los beneficios superen a los costos. Obviamente, la adopción de una lógica vaga desembocaría en una pérdida muy seria con respecto a la simplicidad (el propio Zadeh admite que la lógica vaga es en algunos aspectos mucho menos simple incluso que su lógica básica estándar); y si el lector recuerda que la razón que da Zadeh para preferir hacer lógica imprecisa en lugar de argumentos informales precisos es que estos últimos son propensos a introducir una complejidad difícil de manejar, entonces se sentirá probablemente inclinado a dudar aún más de si la cosa vale la pena. Por otra parte, ni siquiera está claro que la lógica vaga evite la imposición artificial de la precisión. En la lógica básica, aunque no esté uno obligado a insistir en que, por ejemplo, Jack debe ser o definitivamente alto o definitivamente no alto, ni que debe ser o definitivamente alto o definitivamente no alto o definitivamente dudoso, sí estará obligado a insistir en que es alto en el grado 0,7 o en el grado 0,8 o..., etc.; y en la lógica vaga resultante estará obligado a insistir en que, si "Jack es alto" es verdadero en el grado 0,8, se consideraría como *muy verdadero* o solamente como *verdadero pero no muy verdadero* o..., etc. Zadeh propone que *verdadero* se defina como:

$$\textit{verdadero} = 0{,}3/0{,}6 + 0{,}5/0{,}7 + 0{,}7/0{,}8 + 0{.}9/0{.}9 + 1/1$$

i.e., como el conjunto vago en el que el grado de verdad 0,6 pertenece al grado 0,3; 0,7 al grado 0,5; 0,8 al grado 0,7; 0,9 al grado 0,9 y 1 al grado 1 (1975, pág. 411); ¿no es esto una imposición artificial de la precisión? Es difícil evitar la sospecha de que el programa de Zadeh solamente aporta beneficios dudosos y a costos excesivos.

Post scriptum: grados de verdad

El segundo estadio de vaguedad de Zadeh —la extensión de la teoria vaga de conjuntos para "verdadero" y "falso"— se basa en la idea de que la verdad es un asunto de grado y se refleja en su lista de valores de verdad lingüísticos, donde modificadores adverbiales tales como "no muy" y "más o menos" (que él llama "cercos") se agregan a "verdadero" y "falso". Pero la lista que presenta Zadeh de valores de verdad lingüísticos es extremadamente rara: por ejemplo, aunque "muy verdadero" y "más o menos verdadero" parecen aceptables, "bastante verdadero", "ligeramente verdadero" y, por la misma razón, "no muy verdadero" me parecen completamente raros. Esto me lleva a considerar un poco más detenidamente la evidencia lingüística.

Entre los modificadores adverbiales que *se aplican* a "verdadero" están "bastante" y "muy". Ahora bien, "bastante" y "muy" se aplican a predicados de grado, i.e., predicados que denotan propiedades expresadas en grados (bastante alto, pesado, inteligente..., muy alto, pesado, inteligente...), los cuales indican posesión de la propiedad en grado modesto o considerable respectivamente. Y, al parecer, Zadeh piensa que análogamente "bastante verdadero" indica la posesión de un grado modesto de verdad, y "muy verdadero" la posesión de un alto grado de verdad. Pero, mientras "bastante alto (pesado, inteligente)" puede ser aproximadamente equivalente a "más bien (netamente) alto (pesado, inteligente)", "bastante verdadero" ciertamente no significa algo así como "más bien verdadero" o "netamente verdadero". Pues "más bien" y "netamente", lo mismo que otros adverbios que modifican típicamente a los adjetivos de grado, no se aplican precisamente a "verdadero" (sigo la costumbre que tienen los lingüistas de marcar con un asterisco las locuciones inaceptables):

* bastante verdadero
* netamente verdadero
* algo verdadero
* ligeramente verdadero
* extremadamente verdadero

De hecho, "bastante verdadero" puede ser aproximadamente equivalente a "perfectamente verdadero" o "absolutamente verdadero" y (mucho más lejos de contrastar con ello) a "muy verdadero". Además, cuando "bastante" (o "más bien" o "netamente") va junto con un predicado de grado, como "bastante (más bien, netamente) alto (pesado, inteligente)", no puede ir precedido de "no" ("no bastante alto" es inaceptable); mientras que cuando va junto con un predicado absoluto, como en "bastante listo" (ready), sí puede ir prededido de "no" ("no bastante listo"). El comportamiento de "bastante" y "muy" con "verdadero", lejos de apoyar la hipótesis de que "verdadero" es un predicado de grado, indica que es un predicado absoluto.

Pero ¿qué decir sobre otros modificadores adverbiales que se aplican a "verdadero", tales como "enteramente", "completamente", "sustancialmente", "en gran manera", "en parte", "más o menos", "aproximadamente", "esencialmente", "no estrictamente", "no exactamente"..., etc.? Conjeturo que puede ser posible explicar tales locuciones sin tratar la verdad como cuestión de grado; más o menos, se puede esperar algo así como de que " 'p' es enteramente verdadera sii la totalidad de 'p' es verdadera", " 'p' es en parte verdadera sii parte de 'p' es verdadera", " 'p' es aproximadamente verdadera sii 'aproximadamente p' es verdadera"..., etc. Prestaré nuevamente atención a estos temas en el cap. 11, § 3.

10
Lógica modal

1 Verdad necesaria

La lógica modal tiene por fin representar argumentos que involucran esencialmente los conceptos de necesidad y posibilidad. Por tanto, no vendrá mal hacer algunos comentarios preliminares sobre la idea de necesidad. Existe una larga tradición filosófica de distinguir entre verdades *necesarias* y verdades *contingentes*. La distinción se explica a menudo de las siguientes formas: verdad necesaria es aquella que no puede ser de otra manera, verdad contingente es aquella que puede ser de otra manera; o, la negación de una verdad necesaria es imposible o contradictoria, la negación de una verdad contingente es posible o consistente; o, una verdad necesaria es verdadera en todos los mundos posibles (págs. 213 y ss.), una verdad contingente es verdadera en la realidad, pero no en todos los mundos posibles. Evidentemente, tales explicaciones no resultan del todo aclaratorias en vista de sus "(no) puede ser de otra manera", "(im)posible", "mundo posible". Por ello, la distinción se introduce a veces más bien por medio de ejemplos: en un libro reciente (Plantinga, 1974, pág. 1), "7 + 5 = 12", "Si todos los hombres son mortales y Sócrates es hombre, entonces Sócrates es mortal" y "Si una cosa es roja, es coloreada" se ofrecen como ejemplos de verdades necesarias, y "El promedio de lluvia en Los Angeles viene a ser aproximadamente de 12 pulgadas" como ejemplo de verdad contingente.

La distinción entre verdades necesarias y contingentes es *metafísica;* y se debería distinguir de la distinción *epistemológica* entre verdades *a priori* y verdades *a posteriori*. Una verdad *a priori* es aquella que puede conocerse independientemente de la experiencia, una verdad *a posteriori* es aquella que no puede conocerse independientemente de la experiencia. Estas distinciones —metafísica y epistemológica— son ciertamente diferentes. Pero es discutible si coinciden en extensión, es decir, si todas y solamente las verdades

necesarias son *a priori* y todas y solamente las verdades contingentes son *a posteriori*. Las opiniones sobre esta cuestión varían: Kant pensó que había verdades contingentes *a priori;* los positivistas lógicos insistían en la coexistencia de lo necesario con lo *a priori* y de lo contingente con lo *a posteriori;* Kripke ha hecho hincapié recientemente (1972) en que, después de todo, hay verdades contingentes *a priori* (y necesarias *a posteriori*). No entraré en este asunto aquí, donde la verdad necesaria es la principal preocupación; resultará en algún modo relevante cuando en el cap. 12 § 3, me ocupe del status epistemológico de la lógica.

Dentro de las verdades necesarias también es tradicional distinguir entre verdades *físicamente* necesarias (verdades que físicamente no pueden ser de otra manera, cuyas negaciones son físicamente imposibles, verdaderas en todos los mundos físicamente posibles) y verdades *lógicamente* necesarias (verdades que lógicamente no pueden ser de otra manera, cuyas negaciones son lógicamente imposibles, verdaderas en todos los mundos lógicamente posibles). A veces la necesidad física se explica mediante la necesidad lógica como compatibilidad lógica con las leyes de la naturaleza. O, recurriendo de nuevo a ejemplos, tenemos que: "Dos cuerpos se atraen mutuamente con una fuerza proporcional a sus masas" puede servir como ejemplo de verdad físicamente necesaria y "Si dos cuerpos cualesquiera se atraen mutuamente con una fuerza proporcional a sus masas, entonces dos cuerpos cualesquiera se atraen mutuamente con una fuerza proporcional a sus masas" como ejemplo de verdad lógicamente necesaria. Algunos filósofos se muestran escépticos ante tal distinción; véase, por ejemplo, Kneale, 1962a; Molnar, 1969, y cfr. Quine, "Necessity", en 1966a. Por supuesto, que la cuestión de si *hay* verdades físicamente necesarias suscita cuestiones importantes en la filosofía de la ciencia. Pero las lógicas modales fueron diseñadas primordialmente con el fin de representar la necesidad y posibilidad lógica más bien que la física, y es por eso que por lo que solamente hago alusión y no doy respuestas a las interesantes cuestiones surgidas de la idea de necesidad física.

Se ha pensado a veces que la distinción entre verdades lógicamente necesarias y verdades lógicamente contingentes se basa a su vez en la distinción entre verdades *analíticas* y *sintéticas*. "Analítico" y su opuesto "sintético" han sido definidos de varias maneras: Kant definió la verdad analítica como un concepto cuyo predicado está incluido en el concepto de su sujeto o —discutiblemente no equivalente— como aquella verdad cuya negación es contradictoria; Frege definió la verdad analítica o bien como una verdad lógica o bien como una verdad reducible a una verdad lógica mediante definiciones en términos puramente lógicos (así, el logicismo es la tesis de que las verdades de la aritmética son analíticas en este sentido). Más recientemente, las verdades analíticas han sido caracterizadas como "verdaderas únicamente en virtud de su significado", las verdades sintéticas como "verdaderas en virtud de los hechos"; considerándose a las verdades de la lógica como una subclase, verdadera en virtud del significado de las constantes lógicas, de la clase más amplia de las verdades en virtud del significado. (Hintikka, 1973, informa sobre la historia de "analítico". Obsérvese el cambio característico habido desde la explicación cuasi-psicológica de Kant en términos de los *conceptos* involucrados en los *juicios* a las caracterizaciones lingüísticas más recientes en términos de los *significados* de las palabras que componen las *oraciones*.)

Se piensa que la analiticidad explica los *fundamentos* de la verdad necesaria, es decir, lo que hace a una verdad necesaria verdadera necesariamente. Así, se ha supuesto que las distinciones necesario/contingente y analítico/sintético coinciden: la idea es que una verdad analítica, que es verdadera únicamente en virtud de su significado, no podría ser de otra manera más que verdadera y, por tanto, es necesaria [1].

Ahora Quine es escéptico respecto a la distinción analítico/sintético [2]; y su escepticismo es, como veremos, una de las razones para oponerse a la lógica modal.

La crítica de la analiticidad en "Dos dogmas" se dirige principalmente al segundo disyunto de una explicación "fregeana" primitiva de la analiticidad, como:

---

[1] Pero las palabras pueden cambiar de significado; y si lo hacen, ¿no podrían las oraciones que antes eran analíticas convertirse en sintéticas o falsas? Los defensores de la analiticidad podrían replicar que, aunque una misma oración pueda en un tiempo expresar una verdad analítica y en otro tiempo una verdad sintética o tal vez una falsedad, la *proposición* expresada originariamente por la oración permanece analítica aunque la oración deje de expresarlo.

[2] Nota histórica: Quine no rechazó siempre la distinción analítico/sintético; en 1947 usó el concepto de analiticidad, aunque comenta en una nota de pie de página que Goodman había instado al escepticismo respecto de la misma. Y antes de que apareciese "Dos dogmas", Morton White (1950) ya había atacado la distinción analítico/sintético como un "dualismo insostenible". Sin embargo, fue el ataque de Quine el que ejerció más influencia. Hacia 1960a el escepticismo de Quine acerca de la sinonimia fue reforzado con su tesis de la *indeterminación de la traducción:* la tesis de que las nociones significativas no son precisamente, como él postulaba antes, oscuras o de contenido empírico dudoso, sino demostrablemente indeterminadas. No obstante, Quine, en 1973, aunque permanecía oficialmente escéptico acerca del significado, ofreció su propia concepción *ersatz* de analiticidad: en este sentido, un enunciado es analítico si cada uno en la comunidad lingüística aprende que es verdadero aprendiendo a comprenderlo; obsérvese que esta concepción es a la vez característicamente genética y social. De momento me ocuparé sólo de las opiniones de Quine claramente escépticas, esto es, del período comprendido entre 1951 y 1973; se remite al lector a Haack, 1977a, donde he argumentado que la nueva concepción de analiticidad de Quine es proclive a hundirse en la concepción tradicional de la verdad en virtud del significado.

A es analítica sii o bien:

(i) *A* es una verdad lógica

o

(ii) *A* es reducible a una verdad lógica mediante la sustitución de sinónimos por sinónimos

Sería conveniente llamar a la clase de enunciados pertenecientes a (i) o (ii) *analíticos en sentido amplio* y a los que pertenecen a (ii) *analíticos en sentido restringido;* según esta terminología, analiticidad en sentido amplio es la verdad lógica más analiticidad en sentido restringido. El esquema que Quine rechaza está representado en la fig. 5; su crítica va encaminada hacia el concepto de analiticidad en sentido restringido.

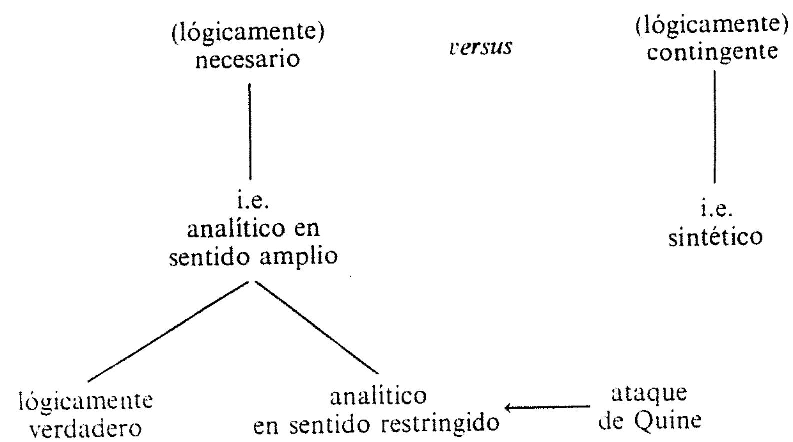


Fig. 5

Específicamente, la estrategia de Quine consiste en argumentar que no puede darse ninguna explicación satisfactoria de la segunda cláusula o de la concepción de sinonimia en la cual descansa. Las explicaciones que se han dado, afirma él, o no logran precisamente caracterizar todas las supuestas verdades analíticas (por ejemplo, la explicación que da Carnap de la analiticidad como verdad en todas las descripciones de estados, argumenta Quine, se aplica solamente a las verdades lógicas y no a las verdades analíticas en sentido restringido que podrían encuadrarse bajo la cláusula (ii)) o dependen abierta o encubiertamente de la comprensión de alguna otra noción intensional no más clara que la de analiticidad misma. Si, por ejemplo, la cláusula (ii) se explica en términos de sustitución a base de definiciones, ello implica una apelación indirecta a la sinonimia sobre la que se basan las definiciones; además, tampoco puede explicarse la sinonimia como sustitutibilidad en todos los contextos *salva veritate* (es decir, sin cambio de valor de verdad), a no ser que se tengan en cuenta contextos tales como "Necesariamente...". En pocas palabras, las explicaciones de la analiticidad nunca pueden evadirse de un "círculo intensional" de conceptos que no son más claros que lo que se está explicando (véase la fig. 6).

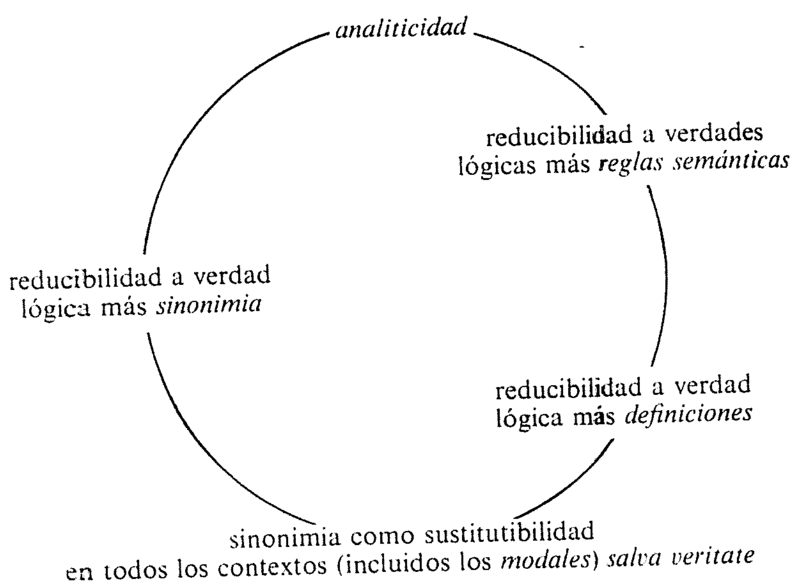


Fig. 6

No es éste el lugar para una discusión exhaustiva del argumento de "Dos dogmas" (lo trataré más detenidamente en el cap. 12, § 3); mi propósito actual consiste más bien en destacar algunos puntos que son especialmente relevantes para la actitud que Quine adopta respecto a la lógica modal.

Primero: el ataque de Quine, aunque con acierto, va dirigido solamente a la analiticidad en sentido restringido: las verdades lógicas, calificadas como analíticas en la cláusula (i), no se ven afectadas. La aversión de Quine hacia el concepto de analiticidad en sentido restringido no es extensible al concepto de verdad lógica. Esto resultará relevante para la discusión (págs. 207, 218) de si en las lógicas modales usuales se puede considerar al operador de ne-

cesidad como representación de la verdad lógica o si debe corresponder a una concepción más amplia de la necesidad.

Quine caracteriza la verdad lógica como "un enunciado que es verdadero y permanece verdadero bajo todas las reinterpretaciones de sus componentes independientemente de las partículas lógicas" (1951, págs. 22-3, y cfr. 1970, cap. 4). Aquí, como en otros lugares, Quine descuida la distinción entre la idea relativa al sistema de la verdad lógica de una fbf de un lenguaje formal y la idea extrasistemática de la verdad lógica de un enunciado de un lenguaje natural. Sugerí (págs. 34-35) que la idea extrasistemática de verdad lógica viene a ser inicialmente sólo una idea no demasiado precisa de un enunciado que es trivialmente verdadero. Sin embargo, si esta idea se depurase de la misma forma que la idea de argumento válido como aquel cuyas premisas no pueden ser verdaderas y su conclusión falsa es depurada por consideración de que un argumento es válido si no hay ningún argumento de la misma forma con premisas verdaderas y conclusión falsa, el resultado sería la idea de un enunciado tal que ningún otro enunciado de la misma forma es falso: lo cual se aproxima mucho a la caracterización de Quine[3].

Segundo: la objeción de Quine a la analiticidad en sentido restringido se apoya en el fondo sobre la idea de que no se puede dar ninguna explicación de la misma a no ser mediante otros términos del "círculo intensional" y que todos estos términos no son claros. Esto resultará relevante para la discusión (págs. 214-15, 218-19) de si, en la interpretación de los sistemas modales, es realista la postura de esperar una explicación no modal de los términos modales.

Sin embargo, me ocuparé primeramente de la caracterización sintáctica de las lógicas modales.

2 SISTEMAS MODALES

Extensiones de la lógica clásica

Un sistema es una extensión de otro si comparte su vocabulario y posee los mismos teoremas e inferencias válidas que involucran solamente el vocabulario compartido, pero posee también un vocabulario adicional y teoremas y/o inferencias válidas adicionales

[3] Strawson ha argumentado (1957) contra Quine que la explicación de la verdad lógica, como la explicación de la analiticidad en sentido restringido, requiere una apelación a la sinonimia. ¿Cómo puede uno saber, se pregunta, que "si él está enfermo, entonces él está enfermo" es lógicamente verdadero, a menos que se tenga la seguridad de que "él está enfermo" significa lo mismo en cada ocurrencia? Pienso que la réplica sería decir que donde puede ser necesario apelar al significado es más bien respecto a la cuestión de "si él está enfermo, entonces él está enfermo" está formalmente representado de forma apropiada por "$p \rightarrow p$".

que involucran esencialmente ese vocabulario; una "lógica extendida" es un sistema que es una extensión de la lógica clásica (capítulo 1, § 2; cap. 9, § 1). Las extensiones de la lógica clásica están frecuentemente motivadas por la creencia de que los cálculos estándar de oraciones y de predicados, aunque resulten irrefutables, no son totalmente adecuados: sus teoremas son lógicamente verdaderos y sus secuentes válidos son preservadores de la verdad, pero hay otras verdades lógicas y/o argumentos válidos que implican operaciones para las que no existe vocabulario y que ni siquiera pueden ser expresadas.

La *lógica modal* añade al vocabulario clásico los operadores uniposicionales "L", que se lee "necesariamente", y "M", que se lee "posiblemente", y el operador biposicional "$\prec$", que se lee "implica estrictamente". (Se han construido otras lógicas extendidas bastante afines a la lógica modal, tales como la *lógica epistémica*, que añade los operadores "K", que se lee "x sabe que", y "B", que se lee "x cree que"; la *lógica deóntica*, que añade los operadores "O", que se lee "debe ser el caso que", y "P", que se lee "está permitido que"; y la *lógica temporal* (cap. 9, § 3).)

Observaciones históricas

La lógica de las oraciones modales fue tratada por Aristóteles y los lógicos medievales; en el presente siglo, Hugh MacColl (1880, 1906) contribuyó con propuestas formales y filosóficas. Pero el desarrollo formal ininterrumpido llegó en el presente siglo como consecuencia de los desarrollos llevados a cabo por Frege y Russell del cálculo no modal de oraciones. Las primeras axiomatizaciones de la lógica modal de oraciones fueron ofrecidas por Lewis en 1918. La extensión a la lógica modal de predicados tuvo lugar con Marcus en 1946.

El motivo principal que impulsó a Lewis a desarrollar las lógicas modales fue la insatisfacción respecto a la noción de implicación material —noción central en la lógica de la *Begriffsschift* y de los *Principia Mathematica*. Puesto que "p" implica materialmente "q" si o "p" es falso o "q" es verdadero, tenemos entonces los teoremas llamados "paradojas" de la implicación material:

$$p \rightarrow (q \rightarrow p)$$
$$-p \rightarrow (p \rightarrow q)$$
$$(p \rightarrow q) \vee (q \rightarrow p)$$

Lewis sostiene que la implicación material de la lógica clásica es totalmente inadecuada para la noción intuitiva de implicación, que requiere no solamente que "p" no *sea* verdadero y "q" falso, sino

que "p" no *pueda* ser verdadero y "q" falso. En consecuencia, propuso que se debería introducir en la lógica de los *Principia* un nuevo operador para la implicación *estricta*, que podría definirse como la necesidad de la implicación material.

Un esbozo formal

Sólo es preciso añadir un operador modal como primitivo al vocabulario de la lógica clásica; con "L" ("necesariamente") como primitivo, "M" ("posiblemente") se define normalmente como:

$$MA = \text{df.} \quad -L-A$$

"$\prec$" como:

$$A \prec B = \text{df.} \quad L(A \rightarrow B)$$

O bien "M" puede tomarse como primitivo y "LA" definirse como "$-M-A$". En las lógicas modales corrientes, las reglas de formación admiten "LA" como una fbf siempre que "A" sea una fbf; esto, por supuesto, tiene en cuenta las *modalidades iteradas*, tales como "LMp" o "LLp".

Ahora, no hay una, sino toda una gama de lógicas modales que difieren entre sí en cuanto al número de axiomas que ellas admiten para regir los operadores modales. Esbozaré alguno de los sistemas más conocidos siguiendo un orden creciente del número de axiomas.

S0·5, que es uno de los sistemas modales más débiles, resulta de la adición de los axiomas:

1. $Lp \rightarrow p$

y

2. $L(p \rightarrow q) \rightarrow (Lp \rightarrow Lq)$

junto con la regla:

(R) Si A es un teorema del cálculo de oraciones, entonces $\vdash_{S0\cdot 5} LA$

El sistema T resulta de la extensión de (R) a:

(RN) Si $\vdash_T A$, entonces $\vdash_T LA$

(de modo que $\vdash LA$ no sólo cuando A es un teorema del cálculo de oraciones (como en (R)), sino también cuando es uno de los axiomas 1 y 2 añadidos, etc.; otra diferencia con (R) es que (RN) es iterable, de modo que se obtiene $\vdash LLA$, $\vdash LLLA$, etc.).

El sistema S4 resulta de T mediante la adición del axioma:

3. $Lp \rightarrow LLp$

y el sistema S5 resulta de S4 mediante la adición de:

4. $Mp \rightarrow LMp$

Existen también otros sistemas modales que son unos más débiles y otros más potentes. El carácter exacto de la lógica modal cuantificada puede también diferir de acuerdo con algunas variaciones en la presentación del cálculo elemental de predicados. (Para más detalles, consúltese Hughes y Cresswell, 1968.)

Relaciones entre los sistemas modales

La proliferación de sistemas modales plantea inmediatamente la cuestión de si estamos obligados a elegir entre ellos, esto es, si cada sistema —y en ese caso, a lo sumo un solo sistema logra el éxito— aspira a captar precisamente las verdades lógicas/inferencias válidas que involucran *la* noción de necesidad, o si tal vez cada sistema —y entonces, todos pueden lograr el éxito— aspira a captar *un* sentido de "necesario". Anticipando una idea que luego discutiré con más detalle (cap. 12, § 1), esta es la cuestión de si los diversos sistemas modales deberían considerarse como *rivales* entre sí. Lemmon ha abogado (1959) en favor de un enfoque tolerante y pluralista; sostiene que cada uno de los sistemas modales puede ser considerado como la formalización de una idea diferente de necesidad: por ejemplo, S0·5 sería la formalización de la idea de tautologicidad, S5 de la idea de analiticidad. Otros, sin embargo (por ejemplo, Cargile, 1972), dudan de la viabilidad de las interpretaciones de Lemmon. Entre los que creen que hay una lógica modal *correcta*, los sistemas que más frecuentemente parecen gozar de su favor son los más potentes S4 y S5.

3 Críticas a la lógica modal

Sin embargo, son más profundas las dudas que la discusión acerca de si hay una lógica modal correcta o de cuál es la lógica modal correcta. Pues la factibilidad e incluso la inteligibilidad de todo el quehacer de la lógica modal han sido cuestionadas. Ha sido Quine quien más las ha recusado durante mucho tiempo (pero cfr. también, por ejemplo, las críticas de Bergmann en 1960).

Las objeciones de Quine son tres: que la motivación para el desarrollo de los sistemas modales se basa en una confusión; que, de todos modos, las lógicas modales no se necesitan para ninguno de los legítimos propósitos de la formalización; y que la interpretación de las lógicas modales presenta dificultades insuperables. A la base de estas objeciones está desde luego la profunda actitud escéptica de Quine en torno al concepto de analiticidad. Es sobre este fondo de escepticismo acerca del status de las nociones modales como se deben ver las objeciones de Quine a las lógicas modales.

La lógica modal "fue concebida en pecado"

Se necesita "$-\!\!3$" o "implica estrictamente", argumentaba Lewis, debido a la excesiva debilidad de "$\rightarrow$" o "implica materialmente". Actualmente Quine señala que "implica materialmente" es de todos modos una lectura gramaticalmente incorrecta de "$\rightarrow$"; pues "$\rightarrow$" es un operador para formar oraciones sobre oraciones, mientras que "implica materialmente" es un predicado biposicional. De este modo, la lógica modal fue "concebida en pecado", el pecado de confundir *uso* (como en "$p \rightarrow q$") y *mención* (como en " 'p' implica materialmente 'q' "). Parece que Lewis efectivamente cayó en esta confusión inducido sin duda por el mal ejemplo de Russell; pero también está muy claro que este delito gramatical no necesita viciar el quehacer de la lógica modal (y cfr. Belnap, 1974, para los argumentos de por qué la divergencia gramatical puede ser positivamente deseable en las innovaciones lógicas). La gramática, insiste Quine, deplora la lectura de "$\rightarrow$" como "implica materialmente"; sin embargo, se da una relación entre dos oraciones "p" y "q" justamente cuando "$p \rightarrow q$", relación débil que con toda propiedad gramatical puede considerarse como "implicación material"; y la lógica modal formaliza otra relación que se da entre dos oraciones "p" y "q", relación más fuerte que puede ser considerada como "implicación estricta".

No se necesita la lógica modal

Como ya he explicado, las lógicas modales son extensiones de la lógica clásica; Quine sugiere (por ejemplo, en 1960a, § 41) que no se necesitan tales extensiones. Desde luego, se plantea el problema de "¿para qué se necesitan?". Quine sostiene que el propósito de la formalización o, como él dice, la "rígida reglamentación" del argumento informal consiste en lograr un lenguaje preciso "adecuado para la ciencia"; y piensa que para los propósitos de la ciencia no se requieren las nociones modales.

La suposición de que el propósito de la formalización es conseguir un lenguaje "adecuado para la ciencia" puede ser discutida, aun cuando Quine, pienso, entienda "ciencia" en sentido muy amplio, incluyendo la matemática así como la física, química, biología, psicología y ciencias sociales, y el discurso cognitivo del sentido común así como la tarea oficial de los científicos profesionales. Ciertamente, algunos lógicos consideran también como parte de su tarea inventar un lenguaje adecuado para representar argumentos, por ejemplo, en el discurso moral (cfr. Smiley, 1963) o en el discurso ficticio (cfr. cap. 4, § 5). La afirmación de que las nociones modales no son esenciales para el discurso científico es también discutida. Resulta particularmente difícil conseguir una perspectiva que no esté distorsionada respecto de este asunto, puesto que el mismo Quine está inclinado —bastante naturalmente— a aplicar cánones severos cuando considera la tesis de que el discurso científico requiere la concepción amplia de necesidad que él rechaza de todos modos. En otras palabras, la pretensión de Quine de que estos conceptos no se necesitan y su afirmación de que son vacíos actúan recíprocamente sin duda alguna.

Un ejemplo: disposiciones y el condicional subjuntivo

Por consiguiente, la mejor manera de comprender lo que aquí está en discusión será considerar detalladamente un caso en el que se sostiene, pero donde Quine niega, que ciertas locuciones son (i) esenciales para el discurso científico y (ii) inexplicables excepto en términos modales. Una familia de locuciones que está al parecer muy profundamente afincada en el lenguaje de la ciencia es el idioma de *disposiciones* y su íntima relación, el *condicional subjuntivo*. Decir, por ejemplo, que x es soluble en el agua es decir que si x fuera puesto en agua, x se disolvería. El condicional material del cálculo de oraciones veritativo-funcional es inadecuado para representar el condicional subjuntivo, pues "$A \rightarrow B$" es verdadero si "A" es falso, mientras que no se supone que "x es soluble" o "si x fuera puesto en agua, x se disolvería" tenga que ser verdadero precisamente porque x no ha sido puesto en agua. Algunos autores creen que una representación formal adecuada del condicional subjuntivo requiere el aparato modal y de modo especial apelar a posibilidades. Stalnaker, 1968, y D. K. Lewis, 1973, han ofrecido análisis modales de los condicionales subjuntivos. Quine, por supuesto, mantiene un punto de vista poco favorable a tales propuestas pre-

cisamente porque usan tal aparato modal; pero también ha sostenido, lo cual es más importante para la presente discusión, que los condicionales subjuntivos se pueden acomodar sin ese aparato modal. Al mismo tiempo, parece que Quine admite que los términos disposicionales deben ser parte de un lenguaje de la ciencia y ofrece un análisis extensional de ellos: por ejemplo, "x es soluble" era explicado a tenor de "$(\exists y)$ (x tiene una estructura interna semejante a y que ha sido colocado en agua y se ha disuelto)". Algunas veces, observa Quine, como en el caso de la solubilidad, se conoce la estructura relevante; otras veces, como en el caso de la irritabilidad, la referencia a una estructura interna no es más que un "pagaré" (véase 1960a, § 46). Sin embargo, se ha argumentado que esta explicación no permite la, por supuesto, genuina posibilidad de que todas las cosas de un cierto tipo tengan una disposición dada, y hasta ahora ninguna de ellas lo ha manifestado, como, quizás, todas las centrales de energía nuclear que tienen una disposición a explotar en ciertas circunstancias, aunque las medidas de seguridad han evitado que no ocurriesen tales circunstancias y por eso ninguna de ellas ha llegado a explotar jamás (Mellor, 1974). De todos modos, posteriormente Quine parece sugerir que, después de todo, los términos disposicionales no pertenecen realmente al lenguaje de la ciencia; son esenciales sólo mientras la empresa de la ciencia sea incompleta, pero pueden abandonarse una vez conocidas las estructuras relevantes. Se podría pensar que el intento de Quine por excluir los lenguajes disposicionales de un "lenguaje teorético rígidamente reglamentado" hubiera sido más convincente de no haber ido precedido por un intento abandonado de incluirlos de una manera extensional; y puede uno sentirse realmente disconforme con la apelación a una ciencia acabada en contraste con una ciencia en desarrollo, puesto que tal distinción resulta especialmente inadecuada con el enfoque usualmente pragmático de Quine en filosofía de la ciencia.

Como he observado, la convicción de Quine acerca de lo que nosotros podemos realizar sin las nociones modales descansa también en su creencia de que la interpretación de la lógica modal se encuentra tan acosada por las dificultades que el empleo de tal aparato no es realmente útil en modo alguno. Por tanto, es el momento de considerar estas dificultades.

La interpretación de la lógica modal está llena de dificultades

Estas críticas se encierran en dos grupos: las dificultades que Quine encuentra en la interpretación de la lógica modal de oraciones y las dificultades adicionales que encuentra en la interpretación de la lógica modal de predicados.

En 1953b, Quine distingue lo que él llama los "tres grados de involucración modal", a saber:

(i) el uso de "necesario" como predicado de oraciones. Aquí "L" se aplicaría a nombres de oraciones o sería citado como en "L'2 + 2 = 4' "; se podría leer "... es necesariamente verdadero" y tendríamos una fuerte analogía con "... es verdadero" de la teoría de Tarski, donde es tratado como un predicado de oraciones;

(ii) el uso de "necesariamente" o "es necesario que" como un operador de formación de oraciones sobre oraciones, como en "$L(2 + 2 = 4)$", donde "L" es tratado como análogo sintácticamente a "es verdadero que...";

(iii) el uso de "necesariamente" como un operador tanto de oraciones cerradas tal como en (ii), como de oraciones abiertas tal como en "$L(2 + 2 = x)$" y de su generalización existencial, "$(\exists x)\ L(2 + 2 = x)$".

La lógica modal de oraciones requerirá a lo sumo el segundo grado de involucración modal, mientras que la lógica modal de predicados requiere el tercero. En los "Tres grados" está claro que Quine considera (i) y (ii), aunque de ningún modo exentos de problemas, como preferibles por lo menos a (iii); y esto guarda correspondencia con su concepción de la lógica modal de oraciones, aunque de ningún modo exenta de problemas, como preferible por lo menos a la lógica modal de predicados.

Dificultades en la interpretación de la lógica modal de oraciones

En las lógicas modales de oraciones del tipo convencional, "L" y "M" son operadores de oraciones como en el grado (ii). Sin embargo, si nos fijamos al menos en los operadores modales únicos, podemos considerar, por ejemplo, a "$L(2 + 2 = 4)$" como una variante sintáctica de "L'2 + 2 = 4' ", tal como en el grado (i).

Desde que Quine no está de acuerdo con la noción de analiticidad, no se siente tan siquiera satisfecho con el uso de "necesario" como predicado oracional. Sin embargo, admite los conceptos de teorematicidad y su contrapartida semántica, la verdad lógica, de modo que la interpretación de "LA" como " 'A' es lógicamente verdadero (un teorema)" es válida para él. Pero este tipo de interpretación tiene en cuenta solamente un fragmento de las lógicas modales oracionales usuales, dado que abandona el status de las modalidades iteradas en cuestión. Sugiere, sin embargo, una interesante línea de pensamiento: que si "LA" se interpretase como "'A' es un teorema (fórmula válida) de L", donde L es una teoría

formal, entonces "LLA" se podría interpretar como " 'LA' es un teorema de M", donde M es el metalenguaje de L. En otras palabras, los operadores modales iterados no serían unívocos, sino que cada uno se referiría a la teorematicidad o verdad lógica de una teoría de la jerarquía de teorías. De acuerdo con esto se han inventado las lógicas modales —motivadas por las diversas consideraciones que acabamos de discutir (Priest, 1976). Sin embargo, las lógicas modales usuales con sus operadores modales iterados unívocamente interpretados no son susceptibles de este tipo de enfoque.

Dificultades en la interpretación de la lógica modal de predicados[+]

Si la adjunción de operadores modales a la lógica de oraciones es dudosa, la mezcla de operadores modales y cuantificadores, argumenta Quine, es desastrosa.

Las dificultades de Quine con la lógica modal cuantificada derivan fundamentalmente de la intersección de sus puntos de vista sobre los cuantificadores y sus puntos de vista sobre la modalidad. Según Quine (cap. 4, § 2), puesto que los términos singulares son eliminables, el compromiso ontológico recae sobre los cuantificadores; ellos constituyen el artificio básico mediante el cual hablamos *acerca de las cosas*. Por otra parte, Quine considera que las locuciones modales hablan no directamente acerca de las cosas, sino acerca de *nuestras formas de hablar sobre las cosas*: la "necesidad", observa él, "reside en la forma en que nosotros decimos las cosas y no en las cosas sobre las que nosotros hablamos" (1953b, pág. 174). Dicho de otra manera, Quine piensa que la modalidad en tanto que inteligible es *de dicto* y no *de re*; "necesario" y "posible" son predicados de oraciones, no de cosas extralingüísticas: " '2 + 2 = 4' es necesario" (el grado (i)) es comprensible, pero "2 + 2 necesariamente = 4" (el grado (iii)) no lo es (cfr. Plantinga, 1974, cap. 1, § 2, y cap. 2). Dados sus puntos de vista sobre los papeles contrapuestos de los cuantificadores y operadores modales, el tema principal de las críticas de Quine a la lógica modal cuantificada no debería sorprendernos: cuando los cuantificadores y operadores modales se combinan, resulta desesperadamente oscuro sobre *lo* que estamos hablando.

Algunas de las dificultades aparecen en el comportamiento anómalo de los términos singulares en el alcance de los operadores modales. Los operadores modales, como dice Quine, son *referencialmente opacos* (o intensionales); la *sustitutividad* (ley de Leibniz)

[+] Una presentación y discución útil de estas críticas se halla en Follesdal, 1969; hay una réplica de Quine en el mismo volumen.

falla en los contextos modales, es decir: dentro del alcance de un operador modal, sustituir un término singular por otro que denote precisamente el mismo objeto puede cambiar el valor de verdad de la oración resultante; por tanto, los términos singulares dentro del alcance del operador modal *no son puramente referenciales*, esto es, no sirven solamente para identificar sus referentes. (Respecto a la opacidad referencial, argumenta Quine, los operadores modales son igual que las comillas u operadores epistémicos.) Por ejemplo (Quine, 1943, 1947, 1953b):

$$9 = \text{número de los planetas}$$

es verdadera, aunque la sustitución sobre la base de esta identidad en la oración verdadera:

$$L(9 > 7)$$

genere la oración presumiblemente falsa:

$$L(\text{el número de los planetas es} > 7)$$

Sin embargo, dado que Quine no permite ni tan siquiera significado fundamental a los términos singulares, los cuales, después de todo, piensa él que pueden y deberían ser eliminados, el núcleo de su objeción yace en el comportamiento anómalo de los cuantificadores y variables ligadas que se encuentran dentro del alcance de los operadores modales.

En el cálculo no modal de predicados:

$$(\exists x)(x > 7)$$

se sigue por generalización existencial de:

$$9 > 7$$

y, análogamente, en el cálculo modal de predicados

$$(\exists x)\, L(x > 7)$$

se sigue de:

$$L(9 > 7)$$

Pero Quine no puede aceptar que *haya algo que sea necesariamente mayor que* 7 ("$(\exists x)\, L(x > 7)$"); el "algo", argumenta él, no puede ser el número 9, pues ése es el número de los planetas y el número

de los planetas no es necesariamente, sino sólo contingentemente, mayor que 7. *Ser necesariamente mayor que* 7, insiste Quine, no puede ser una propiedad de un número; es solamente que el que un tal número sea mayor que 7 se sigue necesariamente si se especifica en ciertas formas (por ejemplo, como el número 9 o como la suma de 5 y 4), pero no si se especifica en otras ciertas formas (por ejemplo, como el número de los planetas). Si el cálculo modal de predicados nos exige aceptar que el número 9 posee la propiedad de ser necesariamente mayor que 7, ello se debe al *esencialismo*, la tesis de que las cosas tienen algunas de sus propiedades necesaria o esencialmente. Pero, según Quine, el esencialismo es una "jungla metafísica" (1953b, pág. 174) para la cual el "enigma" es la única respuesta apropiada (1960a, pág. 199).

Quine admite que las dificultades discutidas hasta aquí podrían evitarse si estuviéramos dispuestos a poner restricciones suficientemente severas sobre el universo de discurso, y de modo expreso a admitir sólo objetos tales que las dos condiciones cualesquiera que los especifican sean necesariamente equivalentes, esto es:

$$\text{C: } ((y)(Fy \equiv y = x) \,\&\, (y)(Gy \equiv y = x)) \rightarrow L(y)(Fy \equiv Gy)^{5}$$

La condición C restaura la sustitutividad: es decir, dado C:

$$(x)(y)((x = y \,\&\, Fx) \rightarrow Fy)$$

Sin embargo, Quine señala (1953b, págs. 155-6) que la sustitutividad junto con la presumiblemente verdadera:

$$L(x = x)$$

produce la consecuencia de que:

$$(x)(y)(x = y \rightarrow L(x = y))$$

esto es, que todas las identidades son necesarias. Y esto, piensa Quine, es dudoso. (Por ejemplo, algunos defensores de la teoría

[5] Quine señala (1953a, págs. 152-3, *contra* Church, 1943) que restringiendo el universo del discurso a objetos intensionales, por ejemplo, conceptos numéricos en vez de números, no sería suficiente para restaurar la sustitutividad. Pues si a es un tal objeto intensional y p una oración que es verdadera pero no necesariamente verdadera, entonces:

$$a = (\iota x)(p \,\&\, x = a)$$

Pero esta identidad no es analítica y sus dos lados no son intercambiables en contextos modales *salva veritate*.

fisicalista han mantenido que la identidad que ellos propugnan entre mente y cerebro es contingente y no necesaria; es como, por ejemplo, la identidad entre el rayo y las descargas eléctricas de la atmósfera. Suponen que tales identidades contingentes son lugares comunes en la ciencia.) Esto está más bien estrechamente relacionado con el primer problema discutido, con "$L\,(\ldots = x)$" reemplazando a "$L\,(\ldots > 7)$". En realidad, la paradoja de la "Estrella de la mañana" es otra versión bien conocida del problema original; necesariamente (presumiblemente) la Estrella de la mañana = la Estrella de la mañana; pero, aunque la Estrella de la mañana = la Estrella de la tarde, no es necesario, sino contingente, que la Estrella de la mañana = la Estrella de la tarde.

Quine argumenta que las ulteriores consecuencias de imponer la condición C son de todos modos aún peores: el colapso de las distinciones modales, pues con la condición C se puede probar que:

$$p \rightarrow Lp$$

lo cual, a tenor del axioma "$Lp \rightarrow p$" significa que $Lp \equiv p$, de modo que "L" es redundante[6].

Así, ésta es la estrategia de la crítica de Quine: la adjunción de los operadores modales induce a una parte de los términos singulares y variables ligadas a un comportamiento anómalo; estas dificultades pueden evitarse mediante una restricción sobre el universo del discurso, pero a costa del colapso de las distinciones modales. Por supuesto, el colapso de las distinciones modales no podría ser tolerado por los defensores de la lógica modal; por tanto, la cuestión es si ellos pueden evitar o explicar lo que Quine ve como "mal comportamiento" por parte de los términos singulares y de las variables ligadas en los contextos modales. De una manera u otra, sus respuestas consisten, como se podría esperar, en afirmar que lo que Quine considera que son consecuencias falsas (o, quizás, dudosamente inteligibles) de la lógica modal cuantificada, de hecho, cuando se entienden apropiadamente, son verdaderas. Por ejemplo, se defienden las modalidades *de re* y el esencialismo (por ejemplo, Plantinga, 1974) y se argumenta que todas las identidades son en realidad necesarias (Marcus, 1962; Kripke, 1972).

No puedo examinar todas las réplicas hechas a la crítica de Quine y me limitaré a un par de ellas que sirven muy bien para ilustrar el asunto en discusión. Varios escritores (por ejemplo, Smullyan, 1948; Fitch, 1949) argumentan que se puede mostrar que el aparente fallo de la sustitutividad en los contextos modales es *meramente* aparente una vez que se ponga el adecuado cuidado acerca de la distinción entre los nombres y las descripciones. Smullyan argumenta así:

[6] El argumento discurre: sea p cualquier oración verdadera y F sea "$p \;\&\; y = x$" y G sea "$y = x$"; entonces, de C se sigue que $L(y)\,(p \;\&\; y = x \equiv y = x)$, por consiguiente, en particular, $L(p \;\&\; x = x \equiv x = x)$ y, por tanto, Lp.

$$9 = \text{el número de los planetas}$$

no es un simple enunciado de identidad cuyos dos términos son nombres *bona fide*, sino más bien tiene la forma:

$$9 = (\iota x)\, Fx$$

y la oración

$$L\,(\text{el número de los planetas} > 7)$$

que Quine considera que es sencillamente falsa, es ambigua; teniendo en cuenta el alcance dado (pág. 86) a la descripción definida, puede entenderse o bien como:

El número de los planetas es necesariamente > 7

o bien como

Es necesario que el número de los planetas sea > 7.

De éstas, argumenta Smullyan, la primera se sigue de "$L\,(9 > 7)$" y "9 = el número de los planetas", pero esto es correcto porque es verdadera; mientras que la segunda es falsa, pero también esto es correcto, porque no se sigue.

La distinción de Smullyan bloquea el argumento original de Quine de una manera muy clara. Sin embargo, su solución requiere que se acepte la verdad de "El número de los planetas es necesariamente mayor que 7" que, cuando se elimina la descripción definida, tiene la forma:

$$(\exists x)\,((y)\,(y \text{ números de los planetas} \equiv x = y) \;\&\; L\,(x > 7))$$

Pero Quine, sin duda, se opondría a esto por cuanto que un cuantificador (el inicial "$(\exists x)$") liga una variable (la "x" en "$x > 7$") dentro de un contexto modal; después de todo, éste es justamente un ejemplo de comportamiento anómalo de las variables ligadas en contextos modales. La solución de Smullyan sería, a los ojos de Quine, inaceptablemente esencialista.

Sin embargo, Marcus niega (1962) que *haya* realmente comportamiento anómalo por parte de los cuantificadores en contextos modales. Las dificultades de Quine proceden de que lee el cuantificador *objetualmente*, como "Hay al menos un objeto, x, tal que x es necesariamente mayor que 7" y luego pregunta qué podría ser ese objeto. Marcus, en cambio, propone que se lea el cuantificador sustitucionalmente, como "Alguna instancia de sustitución de '$L\ (x > 7)$' es verdadera"; y esto, argumenta él, es sencillamente verdadero, ya que "$L\ (9 > 7)$", por ejemplo, es una instancia de sustitución verdadera.

Pero Quine, por supuesto, rechaza la interpretación sustitucional de los cuantificadores. Además, asimila los nombres propios a las descripciones definidas eliminables contextualmente. Por eso, las actitudes de Quine ante los cuantificadores y términos singulares son tales que (i) borran la distinción de la que depende la respuesta de Smullyan y (ii) suponen una prioridad de los cuantificadores sobre los términos singulares que es directamente opuesta a la interpretación sustitucional de la cuantificación hecha por Marcus. El debate discurre de la siguiente forma: se replica a las críticas de Quine refutando las premisas sobre las que descansan. Quine piensa que los cuantificadores hablan acerca de las cosas; según la interpretación sustitucional, los cuantificadores hablan acerca del hablar sobre las cosas. Quine piensa que la modalidad es hablar acerca del hablar sobre las cosas; según el esencialismo, la modalidad es hablar acerca de las cosas.

Los puntos de vista de Quine sobre la cuantificación y la necesidad no son sacrosantos, por supuesto —en efecto, ya he expresado algunas reservas respecto a ellos—. Pero esto no servirá para producir una tendencia del debate entre Quine y los defensores de la lógica modal que degenere en una aserción y contra-aserción menos desagradable, especialmente en vista del hecho de que los puntos de vista rivales sobre los nombres, por ejemplo, tienden a ser defendidos apelando a intuiciones "esencialistas" (por ejemplo, Kripke, 1972; Plantinga, 1974). ¿Qué perspectivas hay de una solución más independiente?

4 Semántica para las lógicas modales

Las críticas de Quine a la lógica modal tienen el sentido, no de que no es *formalmente* factible, sino de que su *interpretación* involucra serias dificultades filosóficas. Estas críticas deberían contemplarse a la luz del hecho de que la lógica modal fue, en sus inicios, desarrollada sintácticamente mediante la introducción de un nuevo vocabulario modal, reglas de formación y axiomas; y que permaneció durante largo tiempo después de su desarrollo sintáctico sin disponer de semántica alguna. Sin embargo, después de la publicación de la crítica de Quine, en los años 1940 y 1950, se desarrolló una semántica formal para la lógica modal (Kanger, 1957a, b;

Kripke, 1963; Hintikka, 1969); es decir, se inventó un modelo formal —comparable, por ejemplo, a las tablas semánticas de verdad para la lógica no modal de oraciones. Y es bastante comprensible que algunos hayan pensado que esto resuelve la cuestión de la interpretabilidad de la lógica modal y muestra que los temores de Quine han sido innecesarios. Resulta, como veremos, que esto está muy lejos de ser obvio.

Semántica formal —un esbozo

Una estructura de modelo es un triplo ordenado $\langle G, K, R\rangle$, donde K es un conjunto del que G es un miembro y en el que R es una relación: para T, R ha de ser una relación reflexiva; para S4, reflexiva y transitiva; para S5, reflexiva, transitiva y simétrica. Una estructura de modelo cuantificada es un par ordenado del cual el primer miembro es una estructura de modelo como la ya descrita, y el segundo una función $\Psi(w)$ que asigna a cada w en K un conjunto de individuos. En cada miembro w de K se especifican las condiciones para la evaluación de las fórmulas; y entonces esta construcción de teoría de conjuntos suministra una definición de "fórmula válida" para cada uno de los sistemas tratados: una fórmula A es válida en un sistema S sii la evaluación de A es verdadera para todo w de K en la estructura de modelo cuantificada.

Hasta aquí se ha suministrado una construcción de teoría de conjuntos en términos en los que se puede definir la validez y establecer la consistencia de los sistemas modales. Sin embargo, se necesita algo más para establecer que estos sistemas, además de ser formalmente factibles, posean una pretensión plausible de representar el razonamiento modal, razonamiento en el que juegan un papel esencial las nociones de necesidad y posibilidad. Kripke sugiere que intuitivamente K podría concebirse como un conjunto de *mundos posibles* $w_1 \ldots w_n$, G como el *mundo real*, R como la relación de accesibilidad que se da entre w_1 y w_2 cuando w_1 es posible respecto a w_2, y $\Psi(w_i)$ como el *conjunto de individuos* que existen en el mundo posible w_i. Según esta interpretación, la semántica formal nos dice, por ejemplo, que "LA" es verdadera solamente en el caso de que "A" sea verdadera en todos los mundos posibles y "MA" solamente en el caso de que "A" sea verdadera en algún mundo posible, de modo que se puede aceptar con cierta plausibilidad que "L" corresponde a "necesariamente" y "M" a "posiblemente".

Semántica "pura" y "depravada"

Distinguí (cap. 3, § 2) cuatro aspectos relevantes para la comprensión de la lógica no modal de oraciones ordinaria; la distinción se aplica igualmente a la lógica modal. Tenemos:

(i)	(ii)	(iii)	(iv)
sintaxis del lenguaje formal	lecturas informales de (i)	semántica formal para (i) ("semántica pura")	explicación informal de (iii) ("semántica depravada")

En el caso del cálculo de oraciones, la semántica formal (iii) suministra una construcción matemática en la que se asigna uno de los valores de verdad, *v* o *f*, a las fbfs del cálculo, y en cuyos términos se define la validez (semántica) y resultan probadas la consistencia y la completud. Se dice, sin embargo, que para toda semántica formal, el cálculo podría ser una notación que representa los circuitos lógicos, donde *v* representa el paso de corriente ("on") y *f* representa la ausencia de paso de corriente ("off"). (En realidad (cap. 1, § 2), las interpretaciones de este tipo de los cálculos bivalente y plurivalente son factibles y útiles.) Pero la afirmación de que el cálculo *sea una lógica de oraciones* que represente los argumentos cuya validez depende de la estructura oracional molecular, depende de la comprensión que se tenga de la semántica formal de tal manera que *v* representa la verdad y *f* la falsedad; en otras palabras, depende de una explicación informal de la semántica formal —nivel (iv).

Las cuestiones que quiero plantear ahora conciernen al status de la semántica depravada. En primer lugar, ¿la necesitamos? Bueno, ya he insistido en que la semántica pura por sí misma no es suficiente: para justificar la afirmación de que un sistema formal *sea una lógica modal* (lógica de oraciones) parece esencial alguna explicación intuitiva de la semántica formal que conexione dicha construcción de teoría de conjuntos con las ideas de necesidad y posibilidad (verdad y falsedad). Al hacer hincapié en este punto de vista me opongo, por supuesto, a una concepción puramente formalista según la cual la lógica consistiría en un puro formalismo no interpretado (pero confróntese con Curry, 1951; hay también una discusión pertinente formulada en términos menos familiares en Derrida, 1973). En segundo lugar, ¿con qué grado de seriedad debemos tomar la semántica depravada? Se ha sugerido que es conveniente considerar la explicación intuitiva dada de la semántica formal como una imagen o metáfora, un artificio heurístico para hacer un poco más aceptable la semántica pura. Pero creo que necesitamos tomar la explicación intuitiva con algo más de seriedad que esto y considerar que la explicación de los mundos posibles y de sus habitantes posibles aspira a la verdad literal (a ser, en memorable expresión de Plantinga, "la sobria verdad metafísica"). Esto es claro para el caso no modal; después de todo, la explica-

ción de "*v*" y "*f*" como "verdadero" y "falso" escasamente ha de rechazarse como meramente metafórica.

Una tercera cuestión metodológica que surge en este punto es, me temo, tan difícil como importante. Tal vez el mejor modo de introducirla sea mediante la crítica que D. K. Lewis hace a las explicaciones de los mundos posibles en términos de consistencia. Lewis argumenta (1973, cap. 4) que tales explicaciones tienden a ser censurablemente circulares. Supongamos que se dice que *w* es un mundo posible solamente si hay una descripción consistente de *w*; si esto significa solamente si hay una descripción de *w* que es *posiblemente verdadera*, no logra explicar "posible" de una manera adecuadamente independiente. Además, Lewis dice que su propia explicación realista de mundos posibles es explicativa y no circular; de hecho, él propone que se use como una prueba de los principios modales discutidos, tal como el principio de S4. Sin embargo, los críticos han puesto de manifiesto que la explicación de Lewis no es más afortunada a este respecto que la alternativa que él critica (Richards, 1975; Haack, 1977a).

Pero, según lo veo ahora, hay que hacer una pregunta más profunda: ¿estamos autorizados para requerir, como hace Lewis, que la explicación intuitiva de la que disponemos en el nivel de la semántica depravada proporcione una respuesta informal explicativa no circular de la semántica formal? Pienso que lo que se requiere es que la semántica depravada se dé en términos que sean, por así decirlo, epistemológicamente independientes de las lecturas de los operadores modales, de modo que podamos decir *si hay un mundo posible en el que A* independientemente de lo que creamos acerca *de si posiblemente A*. Pero, ¿es factible esto? Pueden surgir sospechas inicialmente del hecho de que la explicación de Lewis, como sus rivales, parecen no lograr el requisito epistemológico. Tales sospechas pueden ser confirmadas en cierto grado por las siguientes consideraciones. Los operadores sintácticos de un sistema lógico formal se dan tanto en las lecturas del lenguaje natural como en la semántica formal, que luego tiene que ser a su vez "interpretada". En esta fase, creo que una nueva interpretación *formal* únicamente demoraría el resultado; se necesita, como señalé antes, una explicación *informal*. Pero ahora la explicación informal (a la que llamaré "parloteo" (patter)) o estará estrechamente relacionada con las lecturas en el lenguaje natural de los operadores del sistema o no. Si no lo está, probablemente tendremos que considerar el parloteo como un tanto inapropiado (considérese, por ejemplo, la sugerencia de que *w* sea un mundo posible precisamente en el caso de que sea un país del hemisferio sur; entonces, ¿por qué "*L*", i.e., "verdadero en todos los mundos posibles", debería leerse "necesariamente"?). Pero si el parloteo está cercano a las lecturas, es propenso a violar el requisito de la independencia epistemológica. Es pedir demasiado el que ni "necesario", ni "posible", ni ninguno de sus equivalentes aparezcan en el parloteo; las explicaciones del significado deben terminar en algún lugar. Naturalmente, esto no es decir que no valga la pena dar el parloteo que *amplíe* las lecturas originales; después de todo, puede uno servirse para comprender alguna cosa de la misma cosa dicha de otra manera.

Enfoques de los mundos posibles

Es notable que incluso entre aquellos que aceptan en serio los mundos posibles existe un desacuerdo acerca de qué tipo de cosas sean los mundos posibles. Pueden distinguirse al menos tres enfoques:

(i) el enfoque lingüístico, que interpreta el hablar acerca de los mundos posibles como hablar acerca de conjuntos de oraciones máximamente consistentes (por ejemplo, Hintikka, 1969) y en los que la consistencia podría entenderse o sintáctica o semánticamente

(ii) el enfoque conceptualista, que interpreta el hablar acerca de los mundos posibles como hablar acerca de los diferentes modos en que podríamos concebir el mundo (véase Kripke, 1972)

(iii) el enfoque realista, que considera el hablar acerca de los mundos posibles en su significado literal como hablar acerca de entidades abstractas reales completamente independientes de nuestro lenguaje y pensamiento (véase D. K. Lewis, 1973, cap. 4)[7].

Enfoques de los individuos posibles: identidad transmundana

Por más que se interpreten los mundos posibles, se necesita dar alguna explicación de cuándo se consideran idénticos los individuos posibles en los diferentes mundos posibles; pues las condiciones de verdad de oraciones tales como:

$(\exists x)\ M\,(Fx)$ "hay un x que es posiblemente F"

[7] Los diferentes enfoques guardan una analogía muy fuerte en la filosofía de la matemática con los puntos de vista formalista, intuicionista y logicista sobre el status de los números.

o

$M(Fa)$ ("a es posiblemente F")

serán "en el mundo real hay algún individuo que en algún mundo posible es F" y "en algún mundo posible a es F", y así parece requerir que sea capaz uno de identificar un individuo como idéntico en los diferentes mundos posibles. Considérese, por ejemplo, una oración tal como "Sócrates podría haber sido carpintero"; sus condiciones de verdad vendrían dadas como "Hay un mundo posible en el que Sócrates es carpintero". Pero, ¿qué es lo que determina *qué* individuo es *Sócrates* en otro mundo posible? Supongamos, por ejemplo, que en w_n hay dos individuos posibles, uno justamente igual que Sócrates, pero que es zapatero en vez de filósofo, y otro justamente igual que Sócrates, pero que es carpintero en vez de filósofo; ¿cuál ha de identificarse con el Sócrates real? (véase Chisholm, 1967).

Actualmente el problema de la identidad transmundana se ha confirmado como notablemente espinoso, y existe aquí un desacuerdo considerable acerca de cuál es el mejor modo de abordarlo. Las alternativas parecen ser:

(1) Algunas propiedades del individuo se consideran esenciales para que sea tal individuo, y el criterio para que un individuo sea el mismo individuo en otro mundo posible es que posea esas propiedades. (Éste es el modelo "red" del capítulo 5, § 2.)

(2) El peso del problema es trasladado de los predicados a los nombres. Así, Kripke niega que los nombres propios de individuos sean equivalentes en sentido a cualquier conjunto de descripciones de sus *denotata*, y elude la cuestión de cuántos de dichos conjuntos de descripciones tendría que satisfacer un individuo en otro mundo posible para ser idéntico, por ejemplo, a Sócrates en este mundo. Los nombres propios son *designadores rígidos* que denotan el mismo individuo en todos los mundos posibles; la respuesta correcta a la cuestión de qué individuo es Sócrates en otro mundo posible es simplemente, "Sócrates", *ese* individuo. (Éste es el modelo "arpón".)

(3) Se rechazan los términos de la dificultad original. Para que tenga sentido decir que los individuos son uno y el mismo en diferentes mundos, se niega que sea necesario proporcionar los criterios mediante los cuales se pueda identificar qué individuo es el mismo en otro mundo que un individuo dado en este mundo. Según los defensores de tal enfoque (por ejemplo, Plantinga, 1964, cap. 6), es imposible e indeseable

exigir (cfr. cap. 4, § 2) que se dé el requisito de los "criterios de identidad". Después de todo, observa Plantinga, tiene sentido decir que Georg Cantor fue en otro tiempo un niño precoz, aun cuando fuéramos completamente incapaces de "localizar" o "identificar" a ese niño, o de especificar qué propiedades debe tener un individuo específicamente para ser el niño Cantor.

(4) Otros rechazan los términos del problema original no porque consideren demasiado riguroso el requisito de que se den los criterios de identidad si ha de ser significativo para identificar a los individuos a través de los mundos posibles, sino porque *niegan* que el mismo individuo pueda existir en diferentes mundos posibles, de modo que no existe el problema. Leibniz, el creador de la metafísica de los mundos posibles, pensaba que cada individuo existe solamente en un mundo posible. D. K. Lewis, 1968, adopta esta línea, pero la desarrolla con lo que él denomina la "teoría de las copias". Según esta teoría, cada individuo existe solamente en un mundo posible, pero tiene copias en otros mundos posibles (no necesariamente en todos los otros mundos posibles, y quizás más de una en algún mundo posible); y la verdad de aserciones tales como "Sócrates podría haber sido carpintero" depende ahora, no de si hay un mundo posible en el que Sócrates es carpintero, sino de si hay un mundo posible en el que una copia de Sócrates es carpintero.

¿Se confirman las dudas de Quine?

Las dudas de Quine sobre la lógica modal son anteriores al desarrollo de las semánticas de los mundos posibles; sin embargo, claramente Quine no cree que este desarrollo justifique la confianza sobre la factibilidad filosófica de la lógica modal como opuesta a la puramente formal (véase, por ejemplo, Quine, 1976). Sugeriré en lo que sigue que los problemas filosóficos surgidos de la metafísica de los mundos posibles y de sus habitantes posibles vienen a iluminar y en alguna medida a confirmar las primitivas dudas de Quine (y que las reservas acerca de los puntos de vista sobre la modalidad y la cuantificación en las que se apoyan las críticas originales de Quine pueden ser hasta cierto punto eludidas).

(i) En primer lugar, Quine ha sugerido que si la modalidad —como él sostiene— fuese entendida como un concepto esencialmente metalingüístico, los sistemas modales estándar no serían apropiados. Montague, 1963, investiga detalladamente las restricciones que impondría un tratamiento sintáctico de la modalidad, concluyendo que "virtualmente toda la lógica modal, incluso el sistema débil S1, debe ser sacrificada" (pág. 294). Esto significa además que pocas son las esperanzas de interpretar los sistemas modales convencionales mediante la comprensión de los mundos posibles en sentido sintáctico, como se sugirió en el primer enfoque de los mundos posibles.

(ii) En segundo lugar, Quine duda de si podría darse una explicación de las locuciones modales que no vuelva eventualmente a requerir una comprensión precisamente de las ideas que pretende explicar. Pienso que la negativa de Quine a contentarse con una explicación en términos del "círculo intensional" (analiticidad-sinonimia-definición-regla semántica) de "Dos dogmas", puede verse como bastante fuertemente análoga a una instancia del requisito de la independencia epistemológica, tratado más arriba. La única explicación de los mundos posibles que muestra mayores esperanzas de encontrar este requisito es una explicación lingüística puramente sintáctica, tal que w es un mundo posible solamente si hay una descripción consistente de él, donde "consistente" se entiende en términos puramente sintácticos como "ninguna fórmula de la forma 'A & $-A$' es derivable". Pero tal explicación —como he observado en (i)— lleva a una concepción de la necesidad más débil que la formalizada por los sistemas modales usuales. (Es pertinente aquí el hecho de que el escepticismo de Quine sobre la analiticidad no se extiende a la verdad lógica.) Las explicaciones rivales de mundos posibles parecen ser todas propensas a violar el requisito de independencia: el enfoque lingüístico semántico porque "consistente" es explicado como "posiblemente verdadero"; el realista porque (como, a pesar del hábito de Lewis de hablar de otros mundos posibles como si fueran lugares tan distantes como Australia o Marte, no podemos visitar otros mundos posibles ni, para emplear una de las expresiones de Kaplan, tampoco disponemos de un "instrumento de inspección a lo Julio Verne" mediante el cual poder observarlos) no da ninguna prueba de qué mundos son posibles; el conceptualista porque de alguien que afirme que puede imaginar un mundo en el que A, hay el riesgo de que se diga que él ha descrito equivocadamente lo que imagina si A es inconsistente. Sin embargo, sospecho que el requisito de independencia epistemológica puede ser inaceptablemente riguroso; y, si es así, entonces, en cierto modo, se justifican aquellos críticos de Quine (por ejemplo, Grice y Strawson, 1956) que comentaban que, en "Dos dogmas", él pedía lo imposible y se quejaba cuando no lo conseguía.

(iii) En tercer lugar, Quine encontró desesperadamente oscuro lo que cuantifican las lógicas modales cuantificadas. Pienso que ahora los problemas acerca de la identidad transmundana de los individuos posibles pueden verse razonablemente en cuanto que confirman algunas de las sospechas de Quine a este respecto. Por lo que se refiere a las "soluciones" esbozadas arriba, (4) significa más bien renunciar a resolver el problema en vez de solucionarlo, (3) depende del rechazo del requisito de que solamente se cuantifiquen los ítems para los que se puedan dar las condiciones de identidad adecuadas, (2) depende de la distinción entre nombres y descripciones, la cual rechaza Quine, y (1) parece requerir una forma de esencialismo —esencialismo no sobre *tipos* de cosas, sino sobre *individuos* (cfr. Parsons, 1969)—. Es decir, si se aceptan los puntos de vista de Quine acerca de la cuantificación, entonces el problema de la identidad transmundana de los individuos es solucionable únicamente a costa del esencialismo (individual). Por supuesto, esto nos coloca ante las opciones de rechazar los puntos de vista de Quine sobre la cuantificación o aceptar el esencialismo, además de la opción que recomienda Quine de abandonar la empresa de la lógica modal.

5 PERSPECTIVAS

Para dejar totalmente claro que las siguientes observaciones deberían considerarse bajo el espíritu de una crítica al estilo de Strawson sobre las insuficiencias de los lenguajes formales para con las sutilezas de la lengua inglesa, yo diría que la formalización involucra inevitablemente cierto grado de simplificación; es un propósito legítimo de las lógicas modales aspirar a representar lo que es vital para el razonamiento sobre la posibilidad y necesidad mientras se ignoran las características no esenciales del discurso modal en el lenguaje ordinario. Sin embargo, en vista de la carga metafísica que llevan las lógicas modales, pienso que puede resultar provechoso hacer una nueva consideración del argumento informal que ellas se proponen formalizar.

Hay muchas características del discurso modal en nuestro idioma para las cuales estas lógicas modales son completamente insensibles. Por ejemplo, "posible" toma modificadores como: Es perfectamente (totalmente, enteramente, claramente, remotamente, muy escasamente ...) posible que...; es una posibilidad clara (remota, real...) que... Algunas de estas locuciones sugieren una conexión con "probablemente" (por ejemplo, "Es muy posible que llegue tarde" parece aproximarse en el significado a "Es posible, pero altamente improbable, que llegue tarde"). "Necesariamente" no toma los mismos modificadores —y puede por sí mismo ocasionar algunas dudas acerca de "$Mp \equiv -L-p$"—, pero puede calificarse de otras formas, tal como: Es absolutamente necesario (completamente esencial...) para...

Esta característica puede resultar lógicamente significativa. Pero me siento más molesta por la falta de atención de los lógicos respecto de algunas otras características del discurso modal en nuestro idioma[8] que parecen tener *prima facie* una importancia bastante grande para la relevancia lógica. En nuestro idioma necesitamos prestar atención (i) al tiempo del operador modal y (ii) al tiempo y modo del verbo en la oración en que está contenido. Las lógicas modales convencionales son totalmente insensibles al tiempo y al modo; sin embargo, parece haber una diferencia si leemos:

$$M\,(\exists x)\,(Fx)$$

por ejemplo, como:

Es posible que haya un F (Puede haber un F)

o como:

Es posible que hubiese un F (Podría haber (habido) un F)

o si leemos:

$$M\,(Fa)$$

como:

Es posible que a sea F (a puede ser F)

o como:

Es posible que a fuese F (a podría ser (haber sido) F).

O considérese también la diferencia entre:

Es posible que haya tenido un accidente

dicho cuando la persona a quien se espera tarda y no ha llegado todavía, y:

Era posible que hubiese tenido un accidente

cuando la persona a quien se esperaba ha llegado tarde y se sabe que el retraso ha sido debido a un atasco de tráfico; o el significado del tiempo del operador modal en:

Era posible $\left\{\begin{array}{l}\text{que el gobierno salvase la libra}\\ \text{para el gobierno salvar la libra}\end{array}\right.$
pero no logró actuar a tiempo.

[8] Sería una cuestión pertinente saber si son compartidas por otros lenguajes.

Es sabido que los filósofos se encuentran incapaces para ponerse de acuerdo sobre los valores de verdad de las fórmulas de la lógica modal, especialmente respecto de las que conllevan operadores modales iterados. Esto sorprende poco en vista del hecho de que, sin prestar atención a las consideraciones del tiempo y modo, tenemos dificultad en comprender incluso los enunciados modales con operadores modales simples. Por eso, supongo que podría resultar provechoso para los lógicos tratar de idear sistemas modales que se fundamenten en un aparato básico en el que puedan representarse el tiempo y el modo. Es obvio, sin embargo, que existirá peligro en construir sobre las lógicas modales actualmente disponibles —pues aquellos sistemas han sido ellos mismos construidos por analogía con las lógicas modales convencionales cuyas insuficiencias motivaron esta sugerencia en primer lugar. Y se necesitará algo más que una ingeniosidad meramente formal, incluso en combinación con la sensibilidad respecto a aquellas complejidades del discurso modal no formalizado que parecen ser inferencialmente relevantes; por ejemplo, la interacción entre la modalidad y el tiempo puede plantear cuestiones metafísicas sobre el determinismo. Y entonces, es precisamente este tipo de interdependencia entre formalismo, argumento informal y argumento filosófico lo que hace interesante a la filosofía de la lógica.

6 La implicación de nuevo: un post scríptum sobre la "lógica de la relevancia"

Las "paradojas" de la implicación estricta

Una motivación importante para el desarrollo de la lógica modal fue, como hemos visto, la introducción de una relación de implicación más fuerte desviando así el impacto de las "paradojas" de la implicación material. Y, por supuesto, la implicación estricta es ciertamente más fuerte que la implicación material (puesto que $A \prec B \equiv L(A \to B)$); sin embargo, continúa discutible el hecho de si se conoce enteramente la necesidad por la que se introdujo. Pues la implicación estricta posee sus propias paradojas dado que en los sistemas modales usuales tenemos los teoremas:

$$Lp \to (q \prec p)$$
$$L - p \to (p \prec q)$$

i.e., una proposición necesaria está implicada estrictamente por cualquier proposición y una proposición imposible implica estrictamente a cualquier proposición. No es difícil ver cómo ocurre esto: pues una proposición implica estrictamente a otra sólo en el caso

de que sea imposible que la primera sea verdadera y la segunda sea falsa; y, por tanto, en particular, si es imposible que la primera sea verdadera o si es imposible que la segunda sea falsa.

El mismo Lewis sostenía que estas consecuencias, aunque sorprendentes, debían aceptarse como verdaderas. (En esto ha sido seguido, por ejemplo, por Kneale, 1945-6; Popper, 1947; Bennett, 1954.) Él lo consideraba apropiado todavía para identificar la implicación estricta como la contrapartida formal de la idea intuitiva de "implicación" o "entrañamiento". Pues el entrañamiento, decía, es la conversa de la deducibilidad (A entraña B sii hay una deducción válida de B a partir de A); y las "paradojas" son verdades acerca del entrañamiento dado que, argumentaba él, hay una deducción válida de cualquier conclusión necesaria a partir de una premisa arbitraria y una deducción válida de una conclusión arbitraria a partir de una premisa imposible: por ejemplo, en el último caso discurre de la siguiente manera:

(1)	$p \,\&\, -p$	[premisa imposible]
(2)	p	de (1)
(3)	$p \vee q$	de (2)
(4)	$-p$	de (1)
(5)	q	de (3) y (4) [conclusión arbitraria]

(Por supuesto, este argumento es válido en la lógica oracional estándar. Recuérdese que en los sistemas lógicos estándars la consistencia tiene una importancia tan primordial debido a que de una contradicción se sigue cualquier cosa.) El reto de Lewis a los críticos de la implicación estricta consiste en decir qué paso de este argumento, o de su gemelo para la otra "paradoja", habría posiblemente que rechazar.

Sin embargo, otros autores consideran las "paradojas" de la implicación estricta tan escandalosas como Lewis consideraba las "paradojas" de la implicación material. Han dicho de ellas que están "tan completamente desprovistas de racionalidad [que son] una *reductio ad absurdum* de cualquier punto de vista que las implique" (Nelson, 1933, pág. 271), que son "ultrajantes" (Duncan-Jones, 1935, pág. 78). Estos autores no admitirán que la implicación estricta represente adecuadamente la idea intuitiva de entrañamiento. Nelson, por ejemplo, argumenta que lo que se requiere para que A entrañe B no es sólo que sea imposible que A sea verdadera y B falsa, sino también que haya alguna "conexión de significados" entre A y B[9]. Sin embargo, la dificultad radica en especificar precisamente cuándo

[9] Puede tener algo más que significación histórica el hecho de que el mismo Lewis hiciese algunas propuestas similares en todo a las de Nelson en uno de sus primeros artículos (1912) atacando la noción de implicación de Russell.

hay una "conexión de significados" entre proposiciones y en justificar el rechazo de cualquier paso o pasos de las "pruebas" de las paradojas de la implicación estricta ofrecidas por Lewis que se considera que violan este requisito. Un nuevo problema es que las maniobras adoptadas para bloquear las "pruebas" de Lewis pueden quizás ramificarse en formas no esperadas y poco atractivas; por ejemplo, algunos críticos se encuentran obligados a negar la transitividad del entrañamiento. Sin sorpresa, quizás, se ha dudado (por ejemplo, Suppes, 1957) de si la idea de la conexión de significado o, más generalmente, de si la idea de la *relevancia* de una proposición con respecto a otra es susceptible de tratamiento formal. Los lógicos de la relevancia, sin embargo, piensan de otra manera.

Lógica de la relevancia

Como ocurre con la lógica modal, no hay solamente una, sino todo un rango de "lógicas de la relevancia". Me centraré en R, que es el sistema de la "implicación relevante" propuesto por Anderson y Belnap (1962a, b; 1975), y en E, que es la combinación de R con el sistema modal S4 para producir un sistema de "entrañamiento" (Anderson y Belnap, 1975); cfr. Smiley, 1959, para una explicación especialmente clara y útil de las alternativas.

Anderson y Belnap están de acuerdo en que el entrañamiento, como sostenía Lewis, es la conversa de la deducibilidad; insisten, no obstante, en que la concepción estándar de deducibilidad es defectuosa debido a que ignora las consideraciones de la relevancia. Los lógicos de la relevancia hacen hincapié en que *su* concepción de deducibilidad, y no la noción "oficial" de los lógicos clásicos, es la que cualquier sentido intuitivo y no contaminado exige para que un argumento sea válido:

> Un matemático publica un artículo sobre los espacios de Banach y... concluye con una conjetura. Como nota de pie de página referente a la conjetura escribe: Además de su intrínseco interés, esta conjetura tiene conexiones con otras partes de la matemática que podrían no aparecer inmediatamente para el lector. Por ejemplo, si la conjetura es verdadera, entonces el cálculo funcional de primer orden es completo; mientras que si es falsa, entonces implica que la última conjetura de Fermat es correcta. ... el editor se opone... "a pesar de lo que la mayoría de los lógicos dicen de nosotros, los estándares mantenidos por esta revista requieren que el antecedente de un enunciado de la forma 'si... entonces' debe ser *relevante* para la conclusión obtenida."
> ... *la suposición de que la relevancia es irrelavante para la validez se nos presenta como absurda y, por tanto, realizamos un intento para explicar la noción de relevancia de A con respecto a B.* (Anderson y Belnap, 1975, págs. 17-18; las cursivas finales son mías.)

B es deducible de *A*, por sus estándars, solamente si *se usa* genuinamente la derivación de *B* y no se hace simplemente *un rodeo vía A*. Por supuesto, la idea de que una premisa está siendo realmente usada necesita de explicación. Pero es bastante fácil proporcionar ejemplos del tipo de argumento que Anderson y Belnap describirían como "probar *B bajo* la suposición *A*", pero no "probar *B a partir de* la suposición *A*"; por ejemplo, en un sistema con "$p \rightarrow p$" como axioma, sería:

(1)	q	suposición
(2)	$p \rightarrow p$	axioma
(3)	$q \rightarrow (p \rightarrow p)$	(de (1) y (2) por el teorema de deducción: si $A \vdash B$ entonces $\vdash A \rightarrow B$)

Así lo que Anderson y Belnap proponen es, en primer lugar, poner restricciones apropiadas a la deducibilidad tales como la de que *B* sea deducible a partir de *A* solamente si *A* se usa en la derivación de *B*. Fogelin resume económicamente estas restricciones como "la regla de los negocios sin trampas". Después construyen un sistema de la "implicación relevante" de tal modo que *A* implique relevantemente a *B* sólo en el caso de que *B* sea deducible a partir de *A*, entendiendo "deducible" en el sentido de ellos. Los axiomas para la implicación relevante son (escribiré "$\Rightarrow$" para distinguirlo claramente de "$\rightarrow$" y "$\dashv$"):

1. $A \Rightarrow A$
2. $(A \Rightarrow B) \Rightarrow ((C \Rightarrow A) \Rightarrow (C \Rightarrow B))$
3. $(A \Rightarrow (B \Rightarrow C)) \Rightarrow (B \Rightarrow (A \Rightarrow C))$
4. $(A \Rightarrow (A \Rightarrow B)) \Rightarrow (A \Rightarrow B)$

(Éste es el "fragmento implicacional" de R, i.e., los axiomas que involucran solamente la implicación de R.) Sin embargo, ellos piensan que el entrañamiento requiere la necesidad así como la relevancia; por tanto, la conectiva que representa el entrañamiento sería restringida imponiendo, además de las restricciones a la deducibilidad que aseguran la relevancia, otras restricciones características de la implicación estricta tal como se especificó en S4. Se tiene como resultado los siguientes axiomas que constituyen el fragmento implicacional de E para el entrañamiento (por una obvia analogía escribiré " $\Rightarrow$ "):

1. $A \Rightarrow A$
2. $A \Rightarrow B \Rightarrow ((B \Rightarrow C) \Rightarrow (A \Rightarrow C))$
3. $A \Rightarrow B \Rightarrow (((A \Rightarrow B) \Rightarrow C) \Rightarrow C)$
4. $(A \Rightarrow (B \Rightarrow C)) \Rightarrow ((A \Rightarrow B) \Rightarrow (A \Rightarrow C))$

Finalmente, se obtiene el sistema completo E agregando los axiomas de las otras conectivas oracionales.

Queda por mostrar cómo responden Anderson y Belnap al reto de Lewis: ¿dónde está equivocada, según ellos, la "prueba" que Lewis hace de "q" a partir de "p & $-p$"? Por supuesto, no niegan la afirmación de Lewis de que "q" sea deducible a partir de "p & $-p$" mediante "alguna forma válida de inferencia" *en el sentido "oficial"* de "*válida*"; lo que ellos niegan es que "q" sea deducible a partir de "p & $-p$" mediante una forma válida de inferencia en el sentido en que ellos entienden por válida, i.e., ellos toman *el sentido real de "válida"*. Anderson y Belnap dirigen sus críticas al paso de "$p \vee q$" y "$-p$" a "q". (En la lógica clásica, este paso es claramente justificado al ser una instancia de lo que a veces se conoce como "silogismo disyuntivo".) El diagnóstico más detallado que ellos hacen de lo que hay erróneo en el argumento de Lewis transcurre de la siguiente manera (Anderson y Belnap, 1975, págs. 165-6). "O" tiene dos sentidos, el veritativo-funcional y el intensional; en este último sentido, pero no en el primero, la verdad de "$p \vee q$" requiere que los miembros de la disyunción sean relevantas entre sí. Ahora bien, argumentan ellos, el paso de "p" a "$p \vee q$" es válido sólo si " $\vee$ " se entiende veritativo-funcionalmente, mientras que el paso de "$p \vee q$" y "$-p$" a "q" es válido sólo si " $\vee$ " se entiende intensionalmente. Una vez más indicaré que por "válido" aquí ellos entienden, naturalmente, válido según *su* sentido; no niegan que si "$p \vee q$" (donde " $\vee$ " es veritativo-funcional) es verdadera y "$-p$" es verdadera, entonces necesariamente "q" es verdadera, sino que lo que niegan es que esto sea suficiente para mostrar que el argumento es válido.

Ahora bien, en la lógica clásica, como "$A \rightarrow B$" es equivalente a "$-A \vee B$", el silogismo disyuntivo (de "$-A$" y "$A \vee B$", se infiere "B") es equivalente al *modus ponens* (de "A" y "$A \rightarrow B$" se infiere "B"). Y efectivamente, como se podría esperar en vista de esta equivalencia, el *modus ponens* para la implicación material falla en E.

Según parece por ahora, los lógicos de la relevancia desafían a la lógica clásica en más de una forma.

(i) Más fundamentalmente, su reto se dirige a la concepción clásica de validez. Los lógicos han concebido la relevancia como irrelevante para la validez de un argumento; la irrelevancia, de cualquier manera que se la considere, tiende a ser relegada a la categoría de los defectos retóricos. Por consiguiente, los lógicos de la relevancia dan un sentido más estricto a la noción de que una proposición sea deducible de otra y, consecuentemente, a su conversa, la noción de que una proposición entrañe a otra.

(ii) Por tanto, los lógicos de la relevancia introducen la nueva conectiva de entrañamiento " $\Rightarrow$ " para extender el aparato lógico clásico.

(iii) Y finalmente, su diagnóstico de una "falacia de la relevancia" en el silogismo disyuntivo y, consecuentemente, en el *modus ponens* para la implicación material les lleva no sólo a añadir una nueva conectiva al aparato lógico clásico, sino también a rechazar ciertos principios de inferencia *para las conectivas clásicas.*

En el caso de la lógica de la relevancia tenemos un *reto a los metaconceptos clásicos* (estrategia 6 del cap. 9 § 2), es decir, una *extensión del aparato clásico* (estrategia 4) y, al mismo tiempo, una *restricción* del mismo (estrategia 5). De éstos, el reto al concepto clásico de validez es el más básico. ¿Cómo hay que valorar este reto? Es difícil negar que, en un nivel informal, la irrelevancia sea considerada como defecto del argumento. La cuestión es, más bien, si es considerada más apropiadamente como un defecto lógico o como un defecto retórico. La diferencia entre asunto lógico y retórico podría quizás señalarse de una manera aproximada y fácil mediante la acentuación del interés en la audiencia a la que se dirige el argumento; y teniendo esto en cuenta, la relevancia —concebida como una relación entre proposiciones— posee aparentemente la exigencia de pertenecer a la lógica. Creo que una razón importante de por qué los lógicos han tendido a no prestar atención a las consideraciones sobre la relevancia es que, a primera vista, no parece que sean muy fácilmente susceptibles de tratamiento formal. De modo interesante, el comentario que hace Schiller de que "la doctrina central de la lógica más corriente consiste todavía en un terminante rechazo de la Relevancia" (1930, pág. 75), citado con consentimiento por Anderson y Belnap, se propone como un argumento en contra de las pretensiones de la lógica formal, no como un pretexto para la formalización de la relevancia. Si esto es así, la empresa de la lógica de la relevancia sería, por decirlo así, justificada por el éxito (más bien como las tesis de Davidson para la aplicabilidad de los métodos de Tarski a los lenguajes naturales) —la posibilidad de formalización de la relevancia sería una razón para considerarla como materia de la lógica. Los esfuerzos de los lógicos de la relevancia han contribuido seguramente mucho a rechazar las sospechas de que la relevancia es desesperadamente recalcitrante al tratamiento formal. Puede haber razones, por supuesto, para poner reservas a las lógicas de la relevancia actualmente disponibles (y hay también rivalidad entre ellas) —algunos encuentran la construcción de E, realizada por Anderson y Belnap, en desacuerdo *ad hoc*, y otros, entre los que me encuentro yo misma, dudan de que las lógicas modales convencionales puedan sentirse del todo felices con la estrecha alianza de E con S4. (La sugerencia de Anderson y Belnap, pág. 28, de que la necesidad podría entenderse en términos de entrañamiento mejor que a la inversa es procedente si se considera que la lógica se ocupa primeramente de la validez y secundariamente de la verdad lógica). Y la lógica de la relevancia será inevitablemente más compleja que la lógica veritativo-funcional clásica; de modo que tenemos derecho a preguntar por las ventajas que podríamos esperar de ella.

Una razón por la que las paradojas semánticas y de la teoría de conjuntos se consideran tan catastróficas es que, como en la lógica clásica de una contradicción se sigue cualquier cosa, un sistema formal en el que sea derivable una paradoja carece de valor. Algunos autores han observado, sin embargo, que en el argumento informal los efectos de la contradicción no se consideran tan catastróficamente globales, sino localizados; y es bastante comprensible que algunos de ellos hayan esperado que un formalismo en el que una contradicción no entrañe ninguna fbf arbitraria, pudiese tener ventajas como una "lógica de la paradoja".

Pero el interés de la lógica de la relevancia no ha de restringirse a cuestiones de filosofía de la lógica (esto no debe ser motivo de sorpresa; recuérdense, después de todo, las cuestiones metafísicas para las que resulta pertinente la lógica temporal). Por ejemplo, veo ciertas esperanzas de un concepto de la relevancia en algunas cuestiones interesantes de epistemología. Considérese la idea de Quine, expresada en "Dos dogmas del empirismo", de que la unidad de verificación/falsificación debe ser la totalidad de la ciencia; Quine argumenta bastante persuasivamente que una sola oración no puede ser sometida aisladamente a contrastación empírica y concluye, en algo que podría razonablemente parecer más bien de orden breve, que es el todo de la ciencia lo que se enfrenta con "el tribunal de la experiencia". No sería demasiado caprichoso sospechar que el hecho de que de dos oraciones cualesquiera, o bien la primera implica materialmente a la segunda o bien que la segunda implica materialmente a la primera, pueda prestar un aire de inevitabilidad al dilema de que o bien una oración aislada o bien la totalidad de la ciencia, es la unidad de contrastación empírica; y es interesante especular con que una noción de implicación más fuerte podría hacer sitio a una tercera posibilidad, esto es, que una oración sea contrastada junto con aquellas otras oraciones que resultan relevantes para ella.

He observado más arriba que E conlleva una restricción a la lógica clásica: el *modus ponens* falla en la implicación material. Las lógicas plurivalentes, de las que me voy a ocupar ahora, conllevan también una restricción al aparato lógico clásico.

11
Lógica plurivalente

1 Sistemas plurivalentes

Restricciones de la lógica clásica: lógicas divergentes

Un sistema es divergente de otro si incorpora el vocabulario del primero pero tiene un conjunto diferente de teoremas/inferencias válidas; una "lógica divergente" es un sistema que difiere de la lógica clásica. (Un sistema puede involucrar *tanto* una extensión *como* una divergencia de la lógica clásica, si añade nuevo vocabulario y, por ende, nuevos teoremas/inferencias válidas, pero, al mismo tiempo, difiere de la lógica clásica en lo que respecta a teoremas/inferencias válidas que contienen esencialmente sólo el vocabulario incorporado. El sistema E, considerado en el cap. 10, § 6, sería un ejemplo.) Las lógicas plurivalentes son divergentes; si bien incorporan el vocabulario de la lógica clásica, carecen por norma común de ciertos teoremas de la misma, tales como la "ley de tercero excluido", "$p \vee -p$". (Algunas añaden además nuevo vocabulario, y por eso entran también en la categoría de extensiones.)

Las lógicas plurivalentes que consideraré en este capítulo han sido ideadas gracias a dos tipos principales de motivación; el interés puramente matemático de ofrecer alternativas a la semántica bivalente de la lógica clásica de oraciones; y —de un interés más filosófico— la insatisfacción respecto de la imposición clásica de una dicotomía exhaustiva en la verdad y la falsedad e, igualmente, la insatisfacción con ciertos teoremas o inferencias clásicas. El segundo tipo de motivación es característico —como señalé en el cap. 9, § 2— de las propuestas que propugnan restricciones de la lógica clásica.

Observaciones históricas

Las lógicas plurivalentes tienen una historia tan larga como las lógicas modales: Aristóteles expresa ya reservas sobre la bivalencia

(*De interpretatione*, ix); a comienzos del presente siglo, Hugh Mac Coll elaboró propuestas tanto formales como filosóficas. Pero, como en el caso de las lógicas modales, el impulso del desarrollo formal pormenorizado llegó como consecuencia del desarrollo formal de la lógica bivalente, específicamente, de la semántica de las tablas de verdad para la lógica de la *Begriffsschrift* y de los *Principia Mathematica*, iniciada por Post y Wittgenstein. Los primeros sistemas plurivalentes fueron inventados por Lukasiewicz y Post, 1921 (véase Rescher, 1969, cap. 1, para una discusión histórica detallada).

Sin embargo, en un sentido hay una diferencia importante entre el desarrollo de las lógicas modales y el desarrollo de las lógicas plurivalentes: mientras que en el primer caso la sintaxis (vocabulario, axiomas, reglas) se desarrolló primero y la semántica lo fue sólo considerablemente más tarde, en el segundo caso el desarrollo inicial fue de la semántica, mientras que el de la sintaxis tuvo lugar posteriormente con la creación por parte de Jaskowski, en 1934, de las axiomatizaciones de las lógicas plurivalentes. Es decir, las lógicas plurivalentes comenzaron con el desarrollo de tablas de verdad plurivalente; es de justicia decir, sin embargo, que la cuestión de la interpretación de los valores de estas matrices está todavía, en el mejor de los casos, contestada sólo parcialmente. El problema de la semántica depravada constituye una preocupación principal tanto aquí como en el capítulo precedente.

Esbozo formal

Recuérdese (cap. 3, § 1) que un sistema es n-valente si n es el menor número de valores que tiene cualquier matriz característica de ese sistema. Al hablar de lógicas "plurivalentes", me limitaré a las lógicas n-valentes donde $2 < n$ (por tanto, la lógica bivalente no es plurivalente; ella sigue el uso estándar).

Aunque hay solamente un sistema de lógica bivalente (en el sentido amplio de "sistema" explicado en el cap. 2, en que dos sistemas son iguales si, tenidas en cuenta las diferencias de notación, los símbolos primitivos y los axiomas/reglas, tienen exactamente los mismos teoremas/inferencias válidas), hay sistemas alternativos de lógica trivalente (etc.). Esto no debe sorprendernos en exceso; porque una vez que un sistema posee tres o más valores, son posibles obviamente decisiones alternativas acerca del valor que hay que asignar a las fórmulas compuestas.

Ofreceré solamente un esbozo de algunas de las lógicas plurivalentes más conocidas, y me centraré en los puntos formales que hacen referencia a los temas filosóficos que se plantearán ulteriormente. (Un tratamiento formal más detallado se puede encontrar en Rosser y Turquette, 1952; Ackerman, 1967; Rescher, 1969.) Mi

presentación será más bien semántica que sintáctica; esto no está solamente en consonancia con la historia de las lógicas plurivalentes, sino que también, creo, muestra las diferencias entre ellas de una manera más clara.

La lógica trivalente de Lukasiewicz (Lukasiewicz, 1920, 1930) se caracteriza por las matrices siguientes:

A	$-A$
$*v$	f
i	i
f	v

$A \backslash B$ ($A \& B$)	v	i	f
v	v	i	f
i	i	i	f
f	f	f	f

$A \backslash B$ ($A \rightarrow B$)	v	i	f
v	v	i	f
i	v	v	i
f	v	v	v

$A \backslash B$ ($A \vee B$)	v	i	f
v	v	v	v
i	v	i	i
f	v	i	f

(* indica el valor designado, esto es, el valor tal que las fbfs que uniformemente lo adquieren se consideran como tautologías.)

Inicialmente, Lukasiewicz pensó que el tercer valor, al que denominó "indeterminado" o "posible", debía atribuirse a los enunciados futuros contingentes, los cuales, siguiendo a Aristóteles, consideró que no podían ser ni verdaderos ni falsos. Ni la ley de tercero excluido, ni la ley de no contradicción son uniformemente designadas en estas matrices, de modo que ninguna de las dos en un teorema en L_3; "$p \vee -p$" y "$-(p \& -p)$" toman el valor i cuando "p" lo toma. En cambio, como la tabla de verdad de la implicación proporciona para "$A \rightarrow B$" el valor v, incluso cuando el antecedente y el consecuente toman i, la ley de identidad, "$p \rightarrow p$" es un teorema[1].

La lógica trivalente de Kleene (Kleene, 1952) difiere de la de Lukasiewicz con respecto a la implicación. Mientras Lukasiewicz,

[1] Esta lógica trivalente puede generalizarse. Si se representan los tres valores por los números (1, $\frac{1}{2}$, 0), entonces las matrices de Lukasiewicz caen bajo las reglas:

$$|-A| = 1 - |A|$$
$$|A \vee B| = \text{máx}\,\{|A|, |B|\}$$
$$|A \& B| = \text{mín}\,\{|A|, |B|\}$$
$$|A \rightarrow B| = \begin{cases} 1 & \text{si } |A| \leqslant |B| \\ 1 - |A| + |B| & \text{si } |A| > |B| \end{cases}$$

("$|A|$" significa "el valor de A") y estas reglas producen matrices para 4, 5 ... n, y un número infinito de valores. Cfr. pág. 190.

preocupado por salvar la ley de identidad, asigna v a $|A \to B|$ para $|A| = |B| = i$, Kleene tiene:

A \ B	v	i	f
v	v	i	f
i	v	i	i
f	v	v	v

($A \to B$)

A diferencia de Lukasiewicz, Kleene no consideró i como valor de verdad intermedio; más bien, se trataba de representar "indecidible" y asignárselo a oraciones matemáticas que, aunque verdaderas o falsas, no son ni demostrables, ni refutables. De este modo, las matrices de Kleene se construyen sobre el principio de que allí donde la verdad o la falsedad de un componente es suficiente para decidir la verdad o la falsedad de un compuesto, el compuesto tomaría ese valor a pesar de tener (un u) otros componentes indecidibles; en otro caso: el compuesto es en sí mismo indecidible.

El sistema trivalente de Bochvar (Bochvar, 1939) fue propuesto originariamente como una solución a las paradojas semánticas (cap. 8, § 2), y la interpretación que él concedió para el tercer valor fue "paradójico" o "carente de significado". Bochvar, de acuerdo con el principio de que una oración compuesta con un componente paradójico, es asimismo paradójica, ofreció matrices en las que el tercer valor es, por así decirlo, "infeccioso":

A	$-A$
v	f
i	i
f	v

A \ B ($A \& B$)	v	i	f
v	v	i	f
i	i	i	i
f	f	i	f

A \ B ($A \vee B$)	v	i	f
v	v	i	v
i	i	i	i
f	v	i	f

A \ B ($A \to B$)	v	i	f
v	v	i	f
i	i	i	i
f	v	i	v

(En estas matrices no habrá, desde luego, ninguna fbf que tome v para todas las asignaciones de sus componentes atómicos, pues i de la entrada siempre produce i en la salida. Ya que en cada tabla las entradas centrales exhiben i en todas partes.) Así Bochvar añade un "operador de aserción", que denominaré "V" porque parece significar algo así como "Es verdadero que":

A	VA
v	v
i	f
f	f

Esto le permite definir las conectivas "externas" del siguiente modo:

$$\begin{aligned} -A &= -VA \\ A \& B &= VA \& VB \\ A \vee B &= VA \vee VB \\ A \to B &= VA \to VB \end{aligned}$$

Por consiguiente, las matrices para las conectivas externas siempre poseen v o f en la salida: y de hecho sólo las tautologías bivalentes de la lógica clásica toman uniformemente v para todas las asignaciones de sus componentes. (Las matrices para las conectivas externas son, por decirlo así, tablas trivalentes para la lógica bivalente con i y f como tipos de falsedad.)

Todas las matrices consideradas hasta ahora son *normales* (terminología de Rescher): se asemejan a las matrices bivalentes familiares donde solamente se tiene en cuenta la entrada clásica —donde una fbf compuesta tiene solamente componenetes verdaderos o falsos, las matrices trivalentes proporcionan el mismo valor que daría la tabla clásica. (Es decir, las matrices trivalentes se parecen a las clásicas en lo que respecta a las entradas de ángulo.) Las lógicas plurivalentes de Post son una excepción a causa de su matriz "cíclica" para la negación:

A	$-A$
v	i
i	f
f	v

2 Motivaciones filosóficas

No podré considerar todos los argumentos que sus proponentes han ofrecido a favor de las lógicas plurivalentes, pero voy a limitarme a lo que, espero, sea un ejemplo razonablemente representativo.

Futuros contingentes

Lukasiewicz introduce su lógica trivalente mediante un argumento procedente de Aristóteles, en el sentido de que si uno no acepta que los enunciados sobre el futuro no son aún verdaderos ni falsos, se verá empujado hacia el fatalismo. (La interpretación que Lukasiewicz hace de Aristóteles es discutida, pero no es preciso que me ocupe aquí de dicha discusión; cfr. Haack, 1974, cap. 4, para una discusión relevante.) El argumento de Lukasiewicz transcurre de la siguiente forma. Supongamos que es verdadero ahora que yo estaré en Varsovia a mediodía del 21 de diciembre del año próximo; entonces no puedo no estar en Varsovia a mediodía del 21 de diciembre del año próximo, es decir, es necesario que esté en Varsovia a mediodía del 21 de diciembre del año próximo. Supongamos, por otra parte, que es falso ahora que yo estaré en Varsovia a mediodía del 21 de diciembre del año próximo; entonces no puedo estar en Varsovia a mediodía del 21 de diciembre del año próximo, es decir, es imposible que esté en Varsovia a mediodía del 21 de diciembre del año próximo. Por tanto, si ahora es o verdadero o falso que estaré en Varsovia después, es o necesario o imposible que esté en Varsovia después. La única manera de evitar esta conclusión fatalista, argumenta Lukasiewicz, es negar dicho tiempo futuro, y así los enunciados contingentes son verdaderos o falsos con anticipación al evento. La bivalencia, concluye, debe ser rechazada.

Si este argumento fuese válido, desde luego, daría ocasión al desacuerdo de si tomarlo como una prueba en favor del fatalismo o como una refutación de la bivalencia. (Todos los argumentos discurren, en cierto sentido, por ambos caminos; quiero decir que, dado un argumento a tenor de que B se sigue de A, se podría o bien aceptar la premisa y, por ende, la conclusión, o bien, rechazando la conclusión, rechazar la premisa también.) Sin embargo, como pienso que el argumento es inválido, no necesito detenerme en la cuestión de si el fatalismo es o no una conclusión tolerable. El argumento es inválido, a mi entender, porque depende de una falacia modal, la falacia de argumentar de:

> Necesariamente (si es ahora verdadero [falso] que estaré en Varsovia a mediodía del 21 de diciembre del año próximo, entonces [no] estaré en Varsovia a mediodía del 21 de diciembre del año próximo)

que es, por supuesto, verdadero, a:

> Si es ahora verdadero [falso] que estaré en Varsovia a mediodía del 21 de diciembre del año próximo, entonces necesariamente [no] estaré en Varsovia a mediodía del 21 de diciembre del año próximo

i.e., de argumentar de:

$$L(A \rightarrow B)$$

a:

$$A \rightarrow LB$$

(Si no resulta obvio que esto es una falacia, considérese este ejemplo que claramente *no* preserva la verdad: $L((p \& q) \rightarrow p)$, por tanto $(p \& q) \rightarrow Lp$.)

Si estoy en lo cierto en esto, el fatalismo no se sigue de la bivalencia, de modo que, aunque el fatalismo es una tesis inaceptable, no es necesario rechazar la bivalencia por ese motivo, y Lukasiewicz no ha dado una buena razón para adoptar la lógica trivalente.

Sin embargo, otros autores han presentado argumentos completamente diferentes a favor de la lógica de Lukasiewicz.

Mecánica cuántica

Reichenbach argumenta (1944; Putnam, 1957, apoya su propuesta) que la adopción de la lógica trivalente (la que él propone es exactamente igual que la de Lukasiewicz, excepto en que añade nuevos operadores para la negación e implicación) proporcionaría una solución a algunos problemas planteados por la mecánica cuántica. Su argumento tiene la siguiente estructura: si se usa la lógica clásica, la mecánica cuántica desemboca en consecuencias inaceptables, que él llama "anomalías causales" (aproximadamente, enunciados acerca de los fenómenos de la mecánica cuántica que contradicen las leyes de la física clásica para los objetos observables); pero estas anomalías causales pueden evitarse sin meterse con la mecánica cuántica ni con la física clásica, utilizando una lógica trivalente en vez de una bivalente. En resumen:

> la física clásica & la mecánica cuántica & la *lógica clásica* → anomalías causales
> la física clásica & la mecánica cuántica & la *lógica trivalente* → ninguna anomalía causal

Reichenbach, al igual que Lukasiewicz, denomina al tercer valor "indeterminado"; pero el tipo de enunciado al que pretende atribuirle este valor es completamente diferente del que pensaba Lukasiewicz. En pocas palabras, una de las peculiaridades de la mecánica cuántica es ésta: aunque es posible medir la posición de una partícula y también medir su momento, es imposible —esto se sigue de la teoría— medir simultáneamente la posición y el momento. Bohr y Heisenberg han sugerido que los enunciados que indican la posición y el momento de una partícula en un tiempo dado deben considerarse como carentes de significado o mal formados; Reichenbach prefiere admitir que son significativos (después de todo, cada componente, por separado, no ofrece problemas en absoluto), pero no son ni verdaderos ni falsos, sino indeterminados. El argumento de Reichenbach me plantea demasiadas cuestiones a discutir aquí; por ejemplo: ¿son genuinas las dificultades por las que Reichenbach pretende modificar la lógica? y, de todas formas, ¿es metodológicamente propio modificar la lógica para evitar las dificultades en física? (Pero cfr. Haack, 1974, cap. 9). Sin embargo, parece que apenas si hay dudas de que, aunque Reichenbach tiene razón en que se necesita una lógica no clásica, la lógica particular no clásica que él propone no responde a esa necesidad. Lo que motivó la adopción de una lógica no clásica fue el intento de evitar las anomalías causales sin hacer intervenir a la física (véase Reichenbach, 1944, págs. 159-160, 166); sin embargo, puesto que Reichenbach mantiene que todos los enunciados que indican simultáneamente la posición y el momento son indeterminados, asigna el valor "indeterminado" no sólo a los enunciados anómalos, sino también a ciertas leyes (por ejemplo, al principio de conservación de la energía; 1944, pág. 166).

Es dudoso, por tanto, que Reichenbach haya dado una buena razón para adoptar la lógica de Lukasiewicz. (Por supuesto, queda como posible el que los desarrollos en la mecánica cuántica necesiten realmente adoptar una lógica no clásica, quizás el sistema no veritativo-funcional desarrollado por Birkhoff y von Neumann en 1936; cfr. Putnam, 1969.)

Paradojas semánticas

La lógica trivalente de Bochvar fue propuesta para ofrecer una solución a las paradojas semánticas: "esta oración es falsa" es verdadera si falsa, y falsa si verdadera; la propuesta de Bochvar es que no se le asigne ni "verdadero" ni "falso", sino un tercer valor, "paradójico" o "carente de significado". He argumentado ya (capítulo 8, § 2) que este modo de enfocar las paradojas tiende a llevar de la sartén —la paradoja del mentiroso— al fuego —el mentiroso reforzado— ("esta oración es falsa o paradójica" es verdadera si es falsa o paradójica, es falsa o paradójica si verdadera). Pero lo mismo que con la lógica de Lukasiewicz, se han sugerido también otras razones que las dadas por su creador en favor de una lógica trivalente como la de Bochvar.

Carencia de significado

La "lógica del sin sentido" de Halldén (1949), por ejemplo, tiene matrices como las de las conectivas internas de Bochvar en las

que el tercer valor ("carencia de significado") infecta a cualquier compuesto con algún componente al que se le asigne dicho valor. Pero, de nuevo, esto no proporciona ninguna razón de enorme importancia en favor de la lógica de Bochvar. Porque, como he argumentado en el cap. 9, § 4, toda la empresa de la "lógica de la carencia de significado" me parece fundamentalmente mal concebida.

He comentado ya lo curioso que resulta el carácter "infeccioso" del tercer valor de Bochvar señalando que conlleva la consecuencia, tal vez, un tanto desilusionadora de que no hay ninguna fbf que use sólo las conectivas internas, que tome el valor "verdadero" para todas las asignaciones de sus componentes. Hay una propuesta, sin embargo, que proporciona a esto una base racional interesante.

Sentido sin denotación

Recuérdese (cap. 5, § 2) que Frege mantenía que la denotación/sentido de una expresión compuesta dependía de la denotación/sentido de sus componentes; y que, en consecuencia, a una oración que contiene un término singular que no tiene denotación, le falta el valor de verdad, y una oración compuesta en la que uno de sus componentes no tiene valor de verdad, carece de valor de verdad. El propio Frege prefirió, como hemos visto, asegurar que su lenguaje formal no permita ningún término sin denotación; no obstante, si se permitiesen tales términos, se necesitaría una lógica no clásica para manejarlos a la manera que requiere la teoría de Frege. Smiley sugiere (1960) que una lógica trivalente como la de Bochvar sería el sistema no clásico apropiado. La asignación del tercer valor a una fbf indica aquí, no que tiene un valor de verdad intermedio, sino que no tiene ningún valor de verdad en absoluto. Ahora bien, el hecho de que las matrices para las conectivas internas no asignen ningún valor de verdad a una fbf compuesta si algún componente carece de valor de verdad, corresponde al principio de Frege de que una expresión compuesta carece de denotación si algún componente carece de denotación. Y, con la ayuda del operador de aserción, puede definirse la concepción fregeana de la presuposición ("*A*" presupone "*B*" si "*A*" no es ni verdadero ni falso, a menos que "*B*" sea verdadero). Así pues, creo que esta propuesta logra óptimamente representar el sistema formal que resultaría de adoptar la teoría del sentido y denotación de Frege (compárese con la formalización de Woodruff (1970), que no satisface el principio fregeano de entrada sin valor de verdad/salida sin valor de verdad). Por supuesto, el que se acepte con el fin de suministrar, al mismo tiempo, un argumento en favor de la adopción de la lógica de Bochvar, depende de si se acepta la explicación que hace Frege de las expresiones sin denotación.

Oraciones indecidibles

La lógica trivalente de Kleene ha sido propuesta, como hemos visto, para dar cabida a los enunciados matemáticos indecidibles; el tercer valor representa "indecidible" y la asignación de ese valor a una fbf no se propone para indicar que no es verdadera ni falsa, sino sólo para indicar que no se puede decir qué es. En efecto, es precisamente debido a que Kleene considera que las fbfs indecidibles son verdaderas o falsas, por lo que él adopta el principio de que una fbf compuesta con un componente indecidible sería, decidible si los valores de los otros componentes bastasen para asegurar que la fórmula entera es o verdadera o falsa (por ejemplo, si $|p| = i$ y $|q| = v$, $|p \vee q| = v$). Así, mientras que la motivación filosófica para la lógica trivalente de Kleene parece correcta, lo que él propone parece ser menos radical, menos que un desafío a la lógica clásica bivalente, de lo que inicialmente aparenta (cfr. la insistencia de Kripke (1975) en que el uso que él hace de las reglas de evaluación de Kleene no supone ninguna amenaza para la lógica clásica; véase el cap. 8, § 2). Estas consideraciones suscitan algunas preguntas interesantes acerca de cómo puede esperarse que afecte a la teoría de la verdad la adopción de una lógica plurivalente.

3 Lógicas plurivalentes y valores de verdad

A menudo se ha supuesto sin sorpresa alguna que el uso de una lógica plurivalente llevaría consigo inevitablemente la afirmación de que hay más de dos valores de verdad: afirmación que —nuevamente, quizás sin sorpresa alguna— ha sido a veces el motivo principal de oposición a las lógicas plurivalentes. Pero, de hecho, pienso que está claro que una lógica plurivalente no necesita exigir la admisión de uno o más valores de verdad extra además de "verdadero" y "falso", y, en realidad, que ni tan siquiera necesita exigir el rechazo de la bivalencia.

El uso que Smiley ha hecho de la lógica trivalente de Bochvar ilustra el primer punto. La asignación del tercer valor a una fbf indica que tal fórmula no tiene ningún valor de verdad, no que tiene un tercer valor de verdad no estándar. (Si alguno se siente tentado a considerar a "ni verdadero ni falso" como un tercer valor de verdad parejo con "verdadero" y "falso", la observación de McCall (1970) de que nadie supone que "o verdadero o falso" es un tercer valor de verdad, puede ayudarle a fortalecer su resistencia.)

Además, a veces, los valores intermedios son entendidos no como nuevos valores de verdad, sino, por decirlo así, como variantes epistemológicas de "verdadero" y "falso". Prior sugiere interpretaciones de los valores de una lógica tetravalente de la siguiente manera:

1 = verdadero y puramente matemático (o verdadero y conocido que es verdadero)
2 = verdadero pero no puramente matemático (o verdadero pero no conocido que es verdadero)
3 = falso pero no puramente matemático (o falso pero no conocido que es falso)
4 = falso y puramente matemático (o falso y conocido que es falso)

Y estos ejemplos sirven para verificar también mi afirmación más fuerte de que el uso de un sistema plurivalente no necesita ni siquiera exigir la negación de la bivalencia. Porque esta interpretación supone que toda fbf es o verdadera o falsa.

Otro ejemplo donde la amenaza a la bivalencia resulta ser solamente aparente es éste: Michalski y otros, 1976, proponen una lógica dedocavalente que se dice que es útil para diseñar programas de computador que manejan material sobre las enfermedades de las plantas. Se le podría disculpar a uno que sintiese cierta perplejidad en este punto: ¿cómo debe uno dar sentido a los *diez* valores de verdad extra? Sin embargo, un examen más detenido revela que lo que sucede es bastante menos radical y bastante menos enigmático de lo que parecía al principio. La idea (simplifico, pero espero que no sea erróneamente) es que en vez de clasificar la información acerca de la apariencia de síntomas en forma evidente como, por ejemplo:

> Las manchas rojas aparecen primeramente en enero — falso
> Las manchas rojas aparecen primeramente en febrero — verdadero
> Las manchas rojas aparecen primeramente en marzo — falso
> etc.

se puede clasificar mucho más económicamente así:

> Las manchas rojas aparecen — valor 2

Los 12 valores equivalen, en efecto, a "verdadero en enero", "verdadero en febrero"..., etc. Obsérvese aquí que aquello a lo que le fueron asignados los dos valores clásicos de verdad ("Las manchas rojas aparecen primeramente en enero", etc.) y aquello a lo que le fueron asignados los 12 valores no clásicos ("Las manchas rojas aparecen") son diferentes.

Esto nos lleva a un punto importante y más general: que lo que parece a primera vista la asignación de un valor no estándar a un ítem estándar puede resultar mejor explicable como la asignación de un valor estándar a un ítem no estándar. Esto puede sugerir lo que hay de correcto en las repetidas críticas (por ejemplo, Lewy, 1956; Kneale y Kneale, 1962, págs. 51 y ss.) de que los proponentes de las lógicas plurivalentes están equivocados simplemente acerca de los portadores de verdad.

La interpretación sugerida para los sistemas plurivalentes de Post proporciona una ilustración interesante de este punto. La idea, brevemente, consiste en tomar las "letras oracionales" para representar *secuencias* de oraciones, y considerar que las asignaciones de valores a estas secuencias dependen de la proporción de sus miembros verdaderos respecto de sus miembros falsos (más exactamente: en la lógica n-valente, P representa un $(n - 1)$-tuplo, $\langle p_1, p_2 \ldots p_{n-1} \rangle$, de las oraciones regulares bivalentes, y P toma el valor i justamente cuando $i - 1$ de sus elementos son falsos). Esto sugiere que podría considerarse que las lógicas de Post ofrecen un análogo formal de la idea intuitiva de verdad parcial; así, una oración es parcialmente verdadera si es compleja y parte de la misma es verdadera (cfr. página 169; y Haack, 1974, págs. 62-4, para más discusión).

Lo que he estado sosteniendo hasta ahora es que las lógicas plurivalentes *no necesitan* exigir la admisión de valores de verdad intermedios, ni siquiera el rechazo de la bivalencia. Esto no quiere decir, por supuesto, que no planteen *nunca* este tipo de desafío al punto de vista clásico sobre la verdad. Por ejemplo, el uso de las matrices de Bochvar para representar la explicación fregeana de las expresiones carentes de denotación exige ciertamente la negación de la bivalencia; la asignación del tercer valor representa precisamente la idea de que la fórmula no es ni verdadera ni falsa. (Recuérdese (cap. 7, § 5) que la explicación clásica tarskiana de la verdad es bivalente y, en efecto, que el esquema (T) amenaza con excluir las teorías de la verdad no bivalentes.)

Se habrá observado, tal vez, que de los argumentos filosóficos presentados anteriormente en favor de la adopción de las lógicas plurivalentes, los más persuasivos son los que abogan por considerar los valores intermedios como variantes epistémicas de los valores de verdad clásicos (Kleene), como una asignación de los valores de verdad clásicos a ítems no clásicos (Post), o como una carencia de valor de verdad clásico (Smiley). Esto *puede* ser una coincidencia; pero quienes sospechan de la inteligibilidad de la idea de valores de verdad intermedios podrían encontrar en ello alguna confirmación de sus sospechas.

4 Lógicas divergentes no veritativo-funcionales

Las lógicas plurivalentes, por supuesto, como la lógica clásica, son veritativo-funcionales; el valor asignado a una fbf compuesta depende solamente de los valores asignados a sus componentes. (Las lógicas modales, por el contrario, no son vertitativo-funcionales; el valor de verdad de una fórmula modal no depende solamente del valor de verdad de sus componentes, y las lógicas modales estándar no tienen matrices características finitas.) La preferencia de los lógicos por las conectivas veritativo-funcionales (cfr. cap. 3, § 2) es comprensible puesto que las tablas de verdad proporcionan un procedimiento de decisión sencillo para las lógicas bivalentes y plurivalentes.

Sin embargo, al reflexionar sobre la motivación de las matrices de Kleene, pienso que la suposición de la veritativo-funcionalidad puede resultar cuestionable. Recuérdese que el argumento de Kleene de por qué $|p \vee q|$ sería v si $|p| = i$ y $|q| = v$ es que la verdad de *un* miembro de una disyunción es suficiente para determinar la verdad de la disyunción entera sin tener en cuenta el valor del otro miembro de la disyunción; es decir, que "$p \vee q$" *será verdadera si* "q" *lo es, independientemente de que* "p" *sea verdadero o falso*. Sin embargo, las matrices de Kleene asignan i a "$p \vee q$" cuando $|p| = |q| = i$; y así, en particular, asignan i a "$p \vee -p$" cuando $|p| = |-p| = i$. Pero se puede observar que, mientras "$p \vee q$" no puede considerarse verdadera independientemente de si "p" y "q" son verdaderos o falsos, "$p \vee -p$" *será verdadera independientemente de que* "p" *sea verdadero o falso*. Y esto sugiere que los principios de Kleene podrían justificar una asignación diferente a "$p \vee -p$" de la asignada a "$p \vee q$", cuando los dos miembros de la disyunción tomen i. Pero esto requeriría, por supuesto, una lógica no veritativo-funcional.

Superevaluaciones

Los "lenguajes superevaluacionales" no veritativo-funcionales de van Fraassen (véase 1966, 1968, 1969) parecen estar más en consonancia que las propias matrices trivalentes de Kleene con los principios en virtud de los que Kleene argumenta a favor de sus asignaciones. La idea, en resumen, es ésta: una *superevaluación* asigna a una fbf compuesta algún(os) componente(s) que carece(n) de valor de verdad, valor que todas las evaluaciones clásicas asignarían, si hay un único valor, y en otro caso no habría ningún valor. Dado que todas las evaluaciones clásicas —es decir, tanto las que asignan "verdadero" como las que asignan "falso" a "p" — asignarían "verdadera" a "$p \vee -p$", también la superevaluación atribuye "verdadera" a "$p \vee -p$". Sin embargo, puesto que la evaluación clásica, que asigna "falso" a "p" y a "q", asigna "falsa a "$p \vee q$", mientras que todas las demás evaluaciones clásicas asignan "verdadera" a "$p \vee q$", no hay un único valor asignado por todas las evaluaciones clásicas a "$p \vee q$", y la superevaluación no le atribuye *ningún* valor. No es difícil ver que las superevaluaciones atribuirán "verdadero" a todas las tautologías clásicas y "falso" a todas las contradicciones clásicas, pero no atribuirán ningún valor a las fórmulas contingentes; no obstante, aunque los sistemas de van Fraassen posean de este modo precisamente las mismas tautologías que la lógica clásica, difieren de la lógica clásica en lo que respecta a las inferencias que son aceptadas como válidas —por ejemplo, el "dilema disyuntivo" (si $A \vdash C$ y $B \vdash C$, entonces $A \vee B \vdash C$) falla— que es por lo que se les considera como divergentes.

Lógica intuicionista

Otra lógica divergente no veritativo-funcional que es de interés sustancialmente filosófico *y* formal es la lógica intuicionista de Heyting.

Los intuicionistas afirman (véase, por ejemplo, Brouwer, 1952; Heyting, 1966) que la lógica clásica es, en ciertos aspectos, incorrecta. Es importante, sin embargo, subrayar que su desacuerdo es más profundo que su rechazo de ciertas leyes clásicas. Pues, en principio, el punto de vista intuicionista acerca del ámbito y carácter de la lógica es completamente distinto; los intuicionistas consideran la lógica como secundaria a la matemática, como un conjunto de principios descubiertos *a posteriori* para gobernar el razonamiento matemático. Esto es, obviamente, un reto a la concepción "clásica" de la lógica entendida como el estudio de principios aplicables a todo razonamiento sin tener en cuenta el contenido, y considerada como la teoría más general y fundamental respecto de la cual incluso la matemática es secundaria. Pero esta concepción diferente de la lógica no explicaría por sí misma el desafío de los intuicionistas a ciertas leyes de la lógica clásica, si no fuese por el hecho de que los intuicionistas mantienen también un punto de vista caracterizado acerca de la naturaleza de la matemática; pues las leyes de la lógica clásica, son supuestos, desde luego, para gobernar todo razonamiento, incluyendo el razonamiento matemático clásico.

Según los intuicionistas, la matemática es esencialmente una actividad mental, y los números son entidades mentales (cfr. lo que denominé en el cap. 10, § 4, el punto de vista conceptualista sobre el carácter de los mundos posibles); similarmente, lo que quiere decir que hay un número con tal y tal propiedad es que un tal número es

constructible. El punto de vista constructivista, característicamente psicológico, que los intuicionistas mantienen acerca de la matemática, les conduce a la conclusión de que algunas partes de la matemática clásica —las que tratan de las totalidades completas e infinitas, por ejemplo— son inaceptables. Y de esta restricción de la matemática se sigue una restricción de la lógica; algunos principios de la lógica clásica no son, afirman los intuicionistas, universalmente válidos. Por ejemplo, argumenta Brouwer, hay contraejemplos a la ley de tercero excluido. Supongamos que no es posible ni construir un número con la propiedad F, ni probar que no puede haber un tal número. Entonces, según los cánones de los intuicionistas, no es verdadero que o hay un número que es F o no lo hay.

Obsérvese aquí una diferencia interesante con la actitud de Kleene. Kleene no acepta el hecho de que algún enunciado matemático es indecidible en principio como una razón para negar que es, no obstante, o verdadero o falso. Los intuicionistas, por el contrario, aceptan la idea de que puede haber un número que no pueda ser construido como una parte de una metafísica irremediablemente confusa (véase Heyting, 1966, pág. 4). La diferencia puede servir de modo útil para atraer la atención hacia el hecho de que la distinción entre lo que en el apartado anterior he llamado valores *epistemológicos* versus valores genuinos de *verdad* no puede ser totalmente neutral, sino que puede presuponer algunos supuestos discutibles acerca de las relaciones entre la metafísica y la epistemología.

Brouwer no ofreció un sistema formal de los principios lógicos intuicionistamente válidos debido a que consideraba la matemática como una actividad esencialmente mental y, en consecuencia, pensaba que el formalismo matemático y, *a fortiori*, el lógico eran relativamente poco importantes. Sin embargo, la lógica intuicionista fue formalizada por Heyting que ofrece estos axiomas:

1. $p \rightarrow (p \,\&\, p)$
2. $(p \,\&\, q) \rightarrow (q \,\&\, p)$
3. $(p \rightarrow q) \rightarrow ((p \,\&\, r) \rightarrow (q \,\&\, r))$
4. $((p \rightarrow q) \,\&\, (q \rightarrow r)) \rightarrow (p \rightarrow r)$
5. $q \rightarrow (p \rightarrow q)$
6. $(p \,\&\, (p \rightarrow q)) \rightarrow q$
7. $p \rightarrow (p \vee q)$
8. $(p \vee q) \rightarrow (q \vee p)$
9. $((p \rightarrow r) \,\&\, (q \rightarrow r)) \rightarrow ((p \vee q) \rightarrow r)$
10. $\neg p \rightarrow (p \rightarrow q)$
11. $((p \rightarrow q) \,\&\, (p \rightarrow \neg q)) \rightarrow \neg p$

("$\neg$" es el símbolo usual para la negación intuicionista.) Obsérvese que esta lista contiene los axiomas que gobiernan las conectivas

("&", "$\vee$", "$\rightarrow$", "$\neg$"); en la lógica intuicionista las conectivas no son interdefinibles, por tanto, deben tomarse todas como primitivas. Esto está relacionado, por supuesto, con el hecho de que la lógica intuicionista no posee ninguna matriz característica finita (cfr. los comentarios sobre la interdefinibilidad de las conectivas en el cap. 3, § 1). A la lógica de Heyting le faltan algunos teoremas clásicos; especialmente, ni "$p \vee \neg p$", ni "$\neg\neg p \rightarrow p$" son teoremas. Sin embargo, las dobles negaciones de todos los teoremas clásicos son válidas en la lógica intuicionista.

El sistema de Heyting, aunque sea el mejor confeccionado, no es el único sistema de lógica intuicionista: de hecho, la lógica de Johansson (1936), que no contiene el axioma décimo, es, razonablemente, una propuesta más apropiada para representar los principios lógicos aceptables según los cánones intuicionistas. Pero la lógica de Heyting tiene algunas afinidades inesperadas —afinidades que plantean cuestiones sobre la distinción entre la lógica divergente y la lógica extendida— con la lógica modal, que es por lo que ocupará mi atención en lo que resta de este apartado.

No hay muchas dudas de que los intuicionistas se consideran como modificadores de la corrección de ciertos teoremas de la lógica clásica. Esto justifica que ellos propongan una restricción de la lógica clásica, en la que los teoremas discutidos no valen. Sin embargo, el cálculo de Heyting, aunque a primera vista parece una divergencia de la lógica clásica, puede también ser interpretado como una extensión de la lógica clásica. Si consideramos la negación y la conjunción intuicionista como primitivas, y, en términos de las mismas, definimos la disyunción ($p \vee q$ = df. $\neg(\neg p \,\&\, \neg q)$), la implicación y la equivalencia de la manera usual clásica, entonces todos los teoremas clásicos pueden ser derivados en la lógica de Heyting; además, por supuesto, también son derivables todos los teoremas en la disyunción, implicación y equivalencia intuicionistas —que no son definibles en términos de la negación y conjunción intuicionistas. Y esto hace que la lógica intuicionista parezca más una extensión que una restricción de la lógica clásica. (Pero no todas las inferencias clásicas son preservadas por la traducción propuesta, no lo es, por ejemplo, el MPP, pues de $\neg\neg p \rightarrow p$ según la traducción, la validez de MPP implicaría $\neg\neg p \vdash p$.) Es también posible interpretar el cálculo de Heyting como una lógica modal; si:

$$
\begin{aligned}
m(A) &= LA \text{ (para oraciones atómicas)}\\
m(\neg A) &= L - m(A)\\
m(A \vee B) &= m(A) \vee m(B)\\
m(A \,\&\, B) &= m(A) \,\&\, m(B)\\
m(A \rightarrow B) &= L(m(A) \rightarrow m(B))
\end{aligned}
$$

("$m(A)$" significa "la traducción de A"; las conectivas de la parte izquierda son las intuicionistas y las de la derecha son las conectivas clásicas) es demostrable que una fbf es válida en el cálculo de Heyting sii su traducción es válida en S4 (McKinsey y Tarski, 1948; cfr. Fitting, 1969). De estas dos "traducciones" de la lógica de Heyting, la última parece algo más natural que la primera; pues, a veces, Brouwer y Heyting leen "$\neg$" como "es imposible que...", como, por ejemplo, cuando leen "$(\exists x)\, Fx \vee \neg (\exists x)\, Fx$" como "o existe un F, o se deriva una contradicción del supuesto de que exista un F". Pero, ¿cuál es exactamente el significado de la disponibilidad de estas traducciones?

Es natural esperar una correlación entre las propuestas de restricción de las lógicas clásicas y la idea de que es de alguna forma errónea, por una parte, y las propuestas de extensión de la lógica clásica y la idea de que es de alguna forma inadecuada, por otra. La idea es que una lógica restringida (divergente) excluye algunos teoremas/inferencias expresables totalmente en el vocabulario clásico, y de este modo conlleva la negación de que algunos teoremas clásicos/inferencias son realmente válidos. Pero ahora se puede ver que la cuestión de si una lógica no clásica tiene realmente "el mismo vocabulario" que la lógica clásica, no es tan simple como (quizás) parecía. Los intuicionistas se consideran como críticos de la lógica clásica, y el cálculo de Heyting lo ven como una restricción de la misma; la posibilidad de representar el cálculo de Heyting como una extensión de la lógica clásica plantea la cuestión de si las conectivas intuicionistas difieren en significado de sus "análogas" clásicas. Yo, por mi parte, me inclino a pensar que el hecho de que haya una forma de representar el cálculo de Heyting como una lógica extendida, exige prudencia respecto de la idea de que la crítica de los intuicionistas a la lógica clásica pueda ser completamente explicada como el resultado de discrepancias de significado. Pero el tema general acerca de la relevancia de las consideraciones del significado, en la distinción entre la lógica divergente y la lógica extendida, resultará de importancia para el argumento del próximo capítulo.

12

Algunas cuestiones metafísicas y epistemológicas acerca de la lógica

1 Cuestiones metafísicas

El propósito de este capítulo es abordar algunas de las cuestiones sobre el tema de la lógica que se suscitan debido a la existencia de una pluralidad de sistemas lógicos —pluralidad que se ha explorado en anteriores capítulos. Algunas de estas preguntas son metafísicas: por ejemplo, ¿hay tan sólo un sistema lógico correcto, o podría haber varios que fueran igualmente correctos?, y ¿qué podría significar "correcto" en este contexto?; ¿podríamos estar equivocados en lo que consideramos que son tales verdades? Empezaré con las preguntas metafísicas, ya que las respuestas a las preguntas epistemológicas, hasta cierto punto, es susceptible que dependan de las respuestas a ellas.

Monismo, pluralismo e instrumentalismo

Será útil empezar distinguiendo, a groso modo, tres tipos generales de respuesta a la pregunta de si hay únicamente un sistema correcto:

monismo:	hay tan sólo un sistema lógico correcto
pluralismo:	hay más de un sistema lógico correcto
instrumentalismo:	no hay ninguna lógica "correcta": la noción de corrección es inadecuada

Evidentemente esto necesita ser elaborado y depurado. En principio, con unos comentarios sobre la concepción de corrección que tanto el monismo como el pluralismo exigen: esta concepción depende de la distinción entre validez/verdad lógica extrasistemática y relativa al sistema; aproximadamente, un sistema lógico es correcto si los argumentos formales que son válidos en ese sistema se corresponden con los argumentos informales que son válidos en sentido extrasistemático, y las fórmulas bien formadas que son lógicamente verdaderas en el sistema se corresponden con los enunciados que son lógicamente verdaderos en sentido extrasistemático. El monista sostiene que hay un único sistema lógico que es correcto en este sentido; el pluralista, que hay varios.

Ahora la significatividad de la distinción entre *extensiones* de la lógica clásica y *divergencias* de ella puede apreciarse por completo. Prima facie por lo menos, el lógico modal, por ejemplo, parece que afirma que hay argumentos válidos/verdades lógicas que no pueden representarse en el vocabulario de la lógica clásica y, por tanto, no son argumentos válidos/verdades lógicas de ella; de modo que, aunque la lógica clásica es *correcta hasta donde llega*, no llega bastante lejos. En contraste, el que propone una lógica trivalente, parece afirmar que hay argumentos válidos/verdades lógicas de la lógica clásica cuyos análogos informales no son válidos/lógicamente verdaderos, de modo que la lógica clásica es *en realidad incorrecta*. (Esto explica, de forma más precisa, la idea, bosquejada por primera vez en el cap. 9, § 3, de que las lógicas divergentes plantean un reto a la lógica clásica más serio que las lógicas extendidas.)

Si las lógicas divergentes rivalizan con la lógica clásica, mientras que las lógicas extendidas la complementan, esto parece indicar que ante las primeras sería oportuna una actitud monista (se nos obliga a elegir entre sistemas clásicos y divergentes) y ante las últimas lo sería una actitud pluralista (se podrían aceptar ambas como correctas, la lógica clásica y una lógica extendida). Alternativamente, se podría considerar que la lógica clásica y las extensiones de ella (o, por otra parte, por supuesto, alguna lógica divergente y extensiones de ella) constituyen *juntas* la "lógica correcta". La cuestión es que la diferencia entre pluralismo que acepta la lógica clásica y sus extensiones (o la lógica divergente y sus extensiones) como sistemas lógicos, ambos correctos y un monismo que acepta tanto la lógica clásica como sus extensiones (o una lógica divergente y sus extensiones) como fragmentos las dos de *el* correcto sistema de la lógica, es sólo verbal.

De ahora en adelante concentraré mi atención en la elección entre lógicas clásicas y divergentes (cuestiones parecidas surgen acerca de la elección entre una lógica divergente y otra, y quizás entre una lógica modal y otra —véase pág. 202—; pero no las discutiré aquí), donde el problema entre monismo y pluralismo es sustancial. El monista entiende las lógicas clásicas y las desviadas en cuanto que hacen afirmaciones rivales acerca de qué formalismos representan correctamente a los argumentos válidos/verdades lógicas extrasistemáticamente; el pluralista, en resumen, sostiene que la aparente rivalidad es, de un modo u otro, *meramente* aparente. De hecho, hay varias versiones del pluralismo, modos diferentes de acabar con la aparente rivalidad. Algunos pluralistas comparten con los monistas el supuesto de que la lógica es aplicable al razonamiento sobre cualquier contenido; sin embargo, otros alegan que se pueden aplicar lógicas diferentes a razonamientos sobre diferentes contenidos. De este modo se puede distinguir entre versiones *globales* y *locales* del pluralismo[1]; consideraré primero la versión local.

Según el *pluralismo local*, diferentes sistemas lógicos son aplicables (es decir, correctos con respecto) a diferentes áreas del discurso; quizás, por ejemplo, la lógica clásica a los fenómenos macroscópicos, y la "lógica cuántica" (pág. 235) a los fenómenos microscópicos, tal como diferentes teorías físicas pueden ser más válidas para los fenómenos macroscópicos que para los microscópicos. El pluralismo local relativiza las ideas extrasistemáticas de validez y verdad lógica, y, por tanto, la idea de corrección de un sistema lógico, a un área específica del discurso; un argumento no es válido, sino válido en *d*.

El *pluralismo global* comparte el supuesto monista de que los principios lógicos pueden aplicarse sin tener en cuenta el contenido. Sin embargo, mientras el monista da por sentado que el lógico clásico y el divergente discrepan *en el mismo* sentido acerca de la validez/verdad lógica de *uno y el mismo* argumento/enunciado, el pluralista global niega que el lógico clásico y el divergente estén realmente usando "válido"/"lógicamente verdadero" en el mismo sentido, o si no, que realmente estén discrepando acerca de uno y el mismo argumento/enunciado. La primera idea se relaciona evidentemente con lo que llamé, en el cap. 9, § 2, el "reto a los metaconceptos clásicos"; la última con algunas ideas acerca de los significados de las conectivas que se discutieron en el cap. 3, § 2.

Aproximadamente, el pensamiento de la segunda versión del pluralismo global es éste: fbfs/argumentos, tipográficamente idénticos, no tienen el mismo significado en las lógicas clásicas y las divergentes y, por tanto, no pueden representar en ambas exactamente a los mismos enunciados/argumentos informales. Un argumento a favor de esta opinión es que el significado de las constantes lógicas depende completamente de los axiomas/reglas del sistema en el que ocurren; en consecuencia, cuando cierta fbf, por ejemplo, "$p \vee -p$", es lógicamente verdadera en un sistema y no en otro, entonces esas fórmulas bien formadas, aunque tipográficamente iguales, tienen

[1] El contraste entre la idea de Boole de la lógica como un cálculo, y la de Leibniz de la lógica como un lenguaje universal, discutida en van Heijenoort, 1967b, puede tener afinidades con la distinción entre planteamientos de la lógica locales y globales que por el momento estoy teniendo en cuenta.

diferentes significados en sistemas diferentes; ésta es la tesis de la *varianza del significado*[2]. De modo que lo que dice la fórmula fbf "$p \vee -p$" *es* lógicamente verdadero en lógica clásica, pero lo que exactamente la misma fbf dice en lógica trivalente (en la que "$\vee$" y "$-$", o quizás "p", tienen significados no clásicos) *no* es lógicamente verdadero; así ambas, la lógica clásica y la trivalente, son correctas. Desde este punto de vista el lógico divergente, muy semejante, por ejemplo, al lógico modal, no parece haber desafiado lo antiguo, sino que ha propuesto nuevos argumentos válidos/verdades lógicas —sólo se diferencia del lógico modal en el hábito desagradablemente desconcertante de utilizar los símbolos viejos para su nueva concepción (véase la discusión de la traducibilidad del intuicionismo a la lógica modal en el cap. 11, § 4).

La postura *instrumentalista* se deriva del rechazo de la idea de "corrección" de un sistema lógico, una idea aceptada tanto por monistas como por pluralistas. Según la opinión instrumentalista no tiene ningún sentido hablar de que un sistema lógico es "correcto" o "incorrecto", aunque se podría admitir que es adecuado hablar de que un sistema es más productivo, útil, conveniente..., etcétera, que otro (quizás para ciertos propósitos). El rechazo del concepto de corrección es susceptible de fundamentarse en un rechazo de las ideas extrasistemáticas de verdad y validez que esa concepción requiere; si sólo son inteligibles los conceptos de *verdad lógica en* L y *validez en* L, la pregunta de si las fbfs/argumentos, que son lógicamente verdaderas/válidos en L, se corresponden con los enunciados/argumentos que son extrasistemáticamente lógicamente verdaderos/válidos, simplemente no puede plantearse. Un instrumentalista sólo admitiría la pregunta "interna" de si un sistema lógico es seguro, es decir, si todos y solamente los teoremas/argumentos sintácticamente válidos *del sistema* son lógicamente verdaderos/válidos *en el sistema*.

Otra versión del instrumentalismo parece derivar de una negativa a aplicar a la lógica *cualquier* idea de verdad, incluso una idea relativa al sistema. La lógica, se argumenta, no ha de entenderse como un conjunto de *enunciados*, como una *teoría* que se ha de evaluar en tanto que verdadera o falsa; más bien la lógica ha de entenderse como un conjunto de *reglas* o *procedimientos*[3] a los que sim-

[2] Elijo esta expresión deliberadamente para recordar la tesis de Feyerabend de que los significados de los términos teoréticos de la ciencia, dependen de las teorías en las que ocurren de modo que no se presenta rivalidad entre las teorías científicas alternativas aparentemente en competencia (cfr. Feyerabend, 1963; y véase Haack, 1974, págs. 11-14, para una ulterior exploración de la analogía).

[3] Del mismo modo, la idea de que las "leyes" de la física no se han de considerar como enunciados verdaderos-o-falsos, sino como principios de inferencia, con frecuencia se entiende que es característica de una filosofía de la ciencia "instrumentalista"; véase, por ejemplo, Toulmin, 1953.

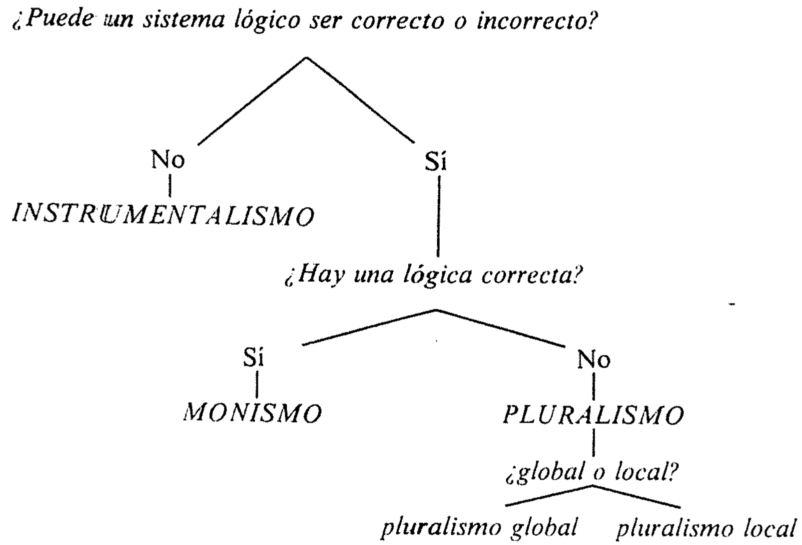


Fig. 7

plemente no se aplican los conceptos de verdad y falsedad. Sin embargo, la pregunta por la corrección, en esta opinión orientada a las reglas, aún se plantearía con respecto a la validez (¿se corresponden los argumentos válidos en **L** con los argumentos informales que son válidos extrasistemáticamente?), a no ser que también se rechace la concepción extrasistemática de validez. De modo que la versión inicial de la postura intrumentalista, que descansa sobre un rechazo de las ideas extrasistemáticas en cuya correspondencia suponen los monistas y los pluralistas que consiste la corrección, es la más fundamental. Estas alternativas se pueden resumir de forma apropiada como en la fig. 7.

He pretendido proyectar las alternativas de una manera tan sistemática como he podido, mejor que hacer una lista de las posiciones mantenidas por autores específicos. Pero de hecho es posible encontrar ejemplos de autores que mantengan cada una de las posiciones que he identificado. Quine parece dar por supuesto algo parecido a lo que he llamado la posición monista cuando, en la segunda mitad de "Dos dogmas" (1951), considera la cuestión (epistemológica) de la posibilidad de revisión de la lógica; en el capítulo 6 de *Filosofía de la lógica* (1970), sin embargo, parece optar por algo similar al pluralismo de la varianza del significado, utilizando argumentos bastante complejos de su teoría de la traducción para apoyar la declaración de que hay suficientes cambios de

significado para excluir la rivalidad; algunos lógicos cuánticos, más claramente Destouches-Février, 1951, pero también, probablemente, Putnam en 1969, apoyan un pluralismo local; el "relativismo" de Rescher (1969, cap. 3) parece estar bastante más cerca de lo que he llamado instrumentalismo, pero en 1977 parece que intenta combinar un instrumentalismo orientado a reglas con la aceptación de una noción extrasistemática de validez.

Resumen de temas

En todo caso, ahora está más claro que los temas más importantes son:

> *¿Tiene sentido hablar de un sistema lógico como correcto o incorrecto? ¿Hay concepciones extrasistemáticas de validez/verdad lógica mediante las cuales caracterizar qué es para una lógica el ser correcta?*

La postura instrumentalista se caracteriza por una respuesta negativa a estas preguntas, los monistas y los pluralistas responden a ellas positivamente. (También ahora podría quedar claro por qué hice notar que algunas cuestiones epistemológicas dependen de las cuestiones metafísicas; a menos que pueda *haber* respuestas a las cuestiones metafísicas; a menos que pueda *haber* una lógica correcta, la pregunta de cómo sabemos si una lógica *es* correcta no se plantea.)

> *¿Debe aspirar un sistema lógico a una aplicación global, es decir, a representar el razonamiento sin tener en cuenta el contenido, o puede una lógica ser localmente correcta, es decir, correcta en un área de discurso limitada?*

La postura pluralista local se distingue por la elección de la segunda de estas opciones.

> *¿Rivalizan las lógicas divergentes con la lógica clásica?*

Los monistas responden afirmativamente a esta pregunta, los pluralistas globales lo hacen negativamente. Todos los temas se refieren a la relación entre argumentos formales e informales, validez relativa al sistema y validez extrasistemática. Por eso la imagen monista puede representarse como en la fig. 8 (pág. 252). (i) aspira a representar (iii) de modo tal que (ii) y (iv) se correspondan en la "lógica correcta". El instrumentalista rechaza completamente (iv); el pluralista local relativiza (iv) para especificar áreas de discurso; el pluralista global o niega que los argumentos formales de la ló-

(i) Argumento formal (fbf) — representa → (iii) argumento informal (enunciado) que es (iv) extrasistemáticamente válido (lógicamente verdadero)

(ii) válido (lógicamente verdadero en L) — que corresponde a lo que sea (iii) → (iv) extrasistemáticamente válido (lógicamente verdadero)

Fig. 8

gica divergente representen los mismos argumentos informales que los de la lógica clásica, es decir, fragmenta la relación entre (i) y (iii), o si no, niega que se intente hacer corresponder la validez en la lógica divergente con la validez extrasistemática en el mismo sentido en el que se la intenta hacer corresponder con la validez en la lógica clásica, es decir, fragmenta la relación entre (ii) y (iv).

Comentarios

En filosofía con bastante frecuencia se da el caso de que plantearse las preguntas oportunas es tener la mitad del camino andado. Sin embargo, la otra mitad no hay que eludirla; y ahora propondré unos comentarios a lo que considero que son los temas más importantes. Pero las preguntas que aquí se han formulado creo que son enormemente difíciles, y hay un serio problema para encontrar un punto de partida respecto al argumento que no demande ninguna pregunta pertinente; por tanto, querría subrayar que los párrafos siguientes son tentativos y, sin duda, inconclusos.

Antes indiqué (cap. 2, § 2) que asumo que hay una idea extrasistemática de validez a la cual aspiran a dar expresión precisa los sistemas lógicos formales. Está bastante claro en la historia de la lógica formal (considérese, por ejemplo, Aristóteles o Frege) que la motivación para la construcción de los sistemas formales ha sido, sobre la base de una concepción inicial de unos argumentos como buenos y otros como malos, separar las características lógicas de los buenos argumentos de las características, por ejemplo, retóticas, y dar reglas que permitan sólo los argumentos lógicamente buenos y excluyan los malos. Por tanto, esto me induce a contestar afirmativamente a las primeras preguntas, y de este modo rechazar la postura instrumentalista. Esta inclinación está reforzada, además, por unas persistentes dudas acerca de si un instrumentalista puede tener algo sensato que decir acerca de cómo elegir entre los sistemas lógicos. El instrumentalista normalmente admite que, al menos para ciertos propósitos, un sistema lógico puede considerarse mejor que otro, quizás como más conveniente, más productivo, más

adecuado, que proporciona las inferencias deseadas... Pero no importa cuán conveniente o productivo pueda ser, si se pudiera inferir "*A* y *B*" de "*A*", éstas no serían razones para preferir un sistema que representa esa inferencia como válida. Soy consciente, por supuesto, de que al hacer comentarios tales como éste corro el riesgo de suponer una concepción extrasistemática de validez, y criticar al instrumentalista por dejar de tenerla en cuenta, cuando, por supuesto, él afirma que no hay una tal concepción (al modo como Russell y Moore supusieron la corrección de una teoría de la correspondencia de la verdad y criticaron la teoría pragmática sobre esa base). Sin embargo, creo que el hecho de que Rescher, al presentar una postura instrumentalista, finalmente admita que es primordial la exigencia de que los argumentos sean preservadores de verdad, puede confirmar justificablemente mis sospechas.

También indiqué (cap. 1, § 2) que entiendo que es característico de la lógica aspirar a presentar los principios que se aplican al razonamiento sobre no importa qué materia; ser *global* en ámbito. Admití que la noción de aplicación al razonamiento de un principio sin tener en cuenta el contenido no estaba perfectamente clara o precisa —comparte algo de la imprecisión de la concepción extrasistemática de validez que se mantiene más bien en virtud de la forma que del contenido. No obstante, aunque pienso que hay lugar para la duda sobre si "cree" o "prefiere" pueden considerarse legítimamente como formas más que como contenidos, sin embargo, me siento bastante segura de que principios que se mantuvieron para el razonamiento sobre contenidos biológicos, pero no para razonamientos sobre física, por ejemplo, no serían principios lógicos (sino biológicos, supongo). En consecuencia, contestaría afirmativamente a la segunda pregunta, y estaría poco dispuesta a aceptar el pluralismo local. Por ejemplo, si resulta, como Birkhoff y von Neumann afirmaron (1936), que cuando se trata de fenómenos cuánticos, "*A* y (*B* o *C*) sii (*A* y *B*) o (*A* y *C*)" no es invariablemente verdadero, entonces la lógica clásica, en la que las leyes distributivas son teoremas, no es correcta. (Estoy completamente dispuesta a conceder que podría ser que aunque los principios clásicos sean estrictamente incorrectos, son válidos para todo razonamiento corriente sobre los fenómenos macroscópicos, por tanto, que sería tan razonable utilizar la lógica clásica a efectos del razonamiento corriente como utilizar la geometría euclidiana con propósitos prácticos, a pesar del hecho de que la geometría euclidiana no es, estrictamente, verdadera para nuestro espacio. Sin embargo, dudo ahora de si esta concesión aplacará al pluralista local.)

Esto nos deja, por un lado, la opción del monismo y, por otro, la de una forma de pluralismo global. Pero creo que en este entorno cambia el carácter del argumento; quiero decir que mientras las dos primeras preguntas tratan de la naturaleza y aspiraciones de la lógica, y se pueden contestar en general, las últimas tratan de las relaciones entre las lógicas clásicas y las divergentes, y, por tanto, no se les puede dar una respuesta general, sino quizás respuestas diferentes para diferentes lógicas divergentes. Es decir, que puede ser que unos lógicos divergentes estén usando metaconceptos diferentes de los que usa el lógico clásico, y otros los mismos; o que la tesis de varianza del significado sea verdadera para unas lógicas divergentes, pero no para otras; o las dos cosas, en efecto. De aquí en adelante es adecuada, más bien, una aproximación por partes. Por supuesto, aunque el monismo y el pluralismo son *asimétricos* de un modo que es relevante, tan sólo *un* ejemplo de una lógica divergente que pudiera ser correcta *lo mismo que* en la lógica clásica, inclinaría la balanza hacia el pluralismo.

Ahora bien, aunque insistí en que hay una idea extrasistemática de validez que los sistemas formales de la lógica aspiran a representar, también observé (págs. 34-35) que esa idea no es de ningún modo completamente precisa, y que puede depurarse y quizás modificarse de acuerdo con el desarrollo de la lógica. El lógico de la relevancia (cap. 10, § 6) rechaza el principio según el cual de "A" y "$A \rightarrow B$" se infiere "B"; propugna que el *modus ponens* es inválido. Además, pone de manifiesto que está hablando de *modus ponens* para la implicación material, clásica y corriente. Sin embargo, no niega que si "A" y "$A \rightarrow B$" son verdaderos, entonces necesariamente "B" es verdadero; lo que quiere decir cuando afirma que MPP *no es* válido, no es lo que quiere decir el lógico clásico cuando afirma que MPP *es* válido, ya que el lógico de la relevancia estaría de acuerdo en que el MPP es válido *en el sentido clásico de "válido"*. Creo que este caso proporciona algunas razones para un pluralismo global (y puede ser que también haya algo que decir sobre la idea de que, en la lógica intuicionista, se esté utilizando una concepción no clásica de verdad lógica).

Sin embargo, optar incondicionalmente por el pluralismo global al llegar a este punto sería, creo, tomar demasiado a la ligera la insistencia de los lógicos de la relevancia en que la concepción de validez de los lógicos clásicos no es solamente *diferente* de la suya, sino también *inadecuada*. Aquí hay auténtica competición, rivalidad genuina, no acerca de qué argumentos son válidos en un sentido de "válido" en el que se está de acuerdo, sino acerca de qué concepción de validez es la más propia y adecuada. (Recuérdese la sugerencia hecha anteriormente, pág. 227, de que a los lógicos de la relevancia se les podría ver como los que insisten en que la relevancia de las premisas para las conclusiones, que los lógicos clásicos consideran como una característica *retórica* de los buenos argumentos, en contraste con los malos, es realmente una característica *lógica* de los buenos argumentos, un asunto de validez.) La consideración que merece el significado de este desacuerdo

parece exigir que se combine un tipo de pluralismo global sobre los sistemas lógicos con el reconocimiento de que puede haber competencia real en el nivel de los metaconceptos.

Ahora bien, ¿qué hay del argumento de varianza del significado para el pluralismo? No es plausible, a mi entender, decir que cuando Lukasiewicz, por ejemplo, niega que "$p \vee -p$" representa una verdad lógica, su aparente desacuerdo con el lógico clásico puede justificarse por completo con el resultado simplemente de dar un nuevo significado a "$\vee$" o a "$-$" o a ambos. Deliberadamente centro la cuestión en este transitado sendero; lo que niego no es que toda lógica divergente siempre supone un cambio del significado de las constantes lógicas —es razonable sospechar, por ejemplo, una idiosincrasia en el significado de la negación y la cuantificación intuicionistas—, sino que toda desviación de la lógica clásica supone siempre inevitablemente varianza de significado a gran escala tal como la que se necesita para evitar la rivalidad real. (He argumentado esto en detalle, con referencia específica a los argumentos de la traducción de Quine, en 1974, págs. 14-21, y 1977c.)

La pregunta es delicada porque hay razones tanto a favor como en contra de la varianza del significado. Argumenté en el cap. 3, § 2, que el significado de las conectivas puede considerarse derivado en parte de los axiomas/reglas del sistema en el que ocurren y de su semántica formal, y en parte, también, de las lecturas informales dadas para las conectivas y la explicación informal de la semántica formal. Los axiomas/reglas y la semántica formal de los sistemas desviados son, por supuesto, diferentes de los clásicos, y la semántica informal puede también diferir (véase la discusión de si los valores intermedios de los sistemas plurivalentes deben considerarse como valores de verdad, cap. 11, § 3); esto argumenta a favor de una varianza del significado. Por otra parte, los lógicos divergentes normalmente emplean las mismas lecturas informales de sus conectivas ("no", "y", "o", "si") que los lógicos clásicos, lo que, por otra parte, parece ser una indicación *prima facie* de que tratan de presentar representaciones rivales de los mismos argumentos informales.

Pero esto sugiere un pensamiento que se ha tenido tendencia a pasar por alto en el debate sobre la varianza del significado (cfr. Quine, 1973, págs. 77 y ss.). La formalización conlleva una cierta abstracción de lo que se considera que son características irrelevantes o poco importantes del discurso informal; el lógico se siente libre para ignorar las connotaciones temporales de "y" o la pluralidad implicada por "alguno". Y esto deja, por así decirlo, espacio para proyecciones formales alternativas del mismo discurso informal; es decir, espacio para la idea de que, por ejemplo, la implicación material, la implicación estricta, la implicación relevante y otros condicionales formales podrían todos tener derecho a representar algún aspecto del "si", o que las disyunciones bivalentes, trivalentes y no extensionales podrían todas ser posibles proyecciones de (algunos) usos de "o". Y esto proporciona más apoyo a un planteamiento pluralista, según el cual, sin embargo, más que diferentes formalismos que aspiran a representar argumentos informales diferentes, puede ser que den diferentes representaciones de los mismos argumentos.

Otra vez, es probable que haya desacuerdo entre los lógicos clásicos y los divergentes —aun cuando su rivalidad en el nivel de los sistemas lógicos pueda mitigarse como he sugerido— sobre cuál es el modo mejor, o quizás, el más apropiado, de representar los argumentos informales. Sin embargo, soy escéptica ante la idea de que se puede suponer que hay una notación formal única e idealmente perspicua en la cual la única forma lógica de todo argumento informal está correctamente representada (de aquí mi preferencia por "una forma lógica" sobre "la forma lógica" de un argumento, cap. 2, § 4). Algunas representaciones formales pueden ser mejor que otras, ya sea absolutamente o para ciertos propósitos, pero no estoy segura de que haya una única mejor. (Es posible, también, que una representación formal sea preferible en un área de discurso y otra en otra área; y de ser así quizás algo del pluralismo local podría salvarse.)

Por tanto, me inclino a favor de la postura pluralista global: puede haber varios sistemas lógicos que sean correctos en el sentido que he explicado. La imagen monista (fig. 8, arriba) podría reemplazarse por algo más de acuerdo con la fig. 9:

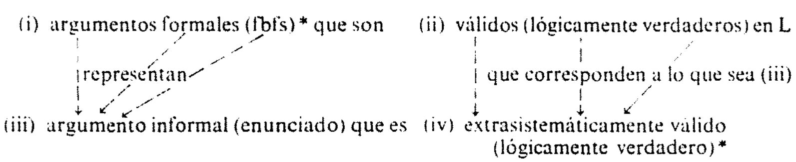


Fig. 9

donde los argumentos informales pueden representarse formalmente en más de una forma, y donde la validez/verdad lógica en L puede corresponder a diferentes concepciones extrasistemáticas de validez/verdad lógica. Sin embargo, subrayo primero que esto no significa que *nunca* se ha de elegir entre una lógica clásica y una divergente, sólo que *a veces* no es necesario (por tanto, mi pluralismo es, por así decirlo, poco sistemático aunque global); y segundo que, incluso en los casos en que una lógica clásica y una divergente pueden ser ambas correctas, hay, sin embargo, competencia entre ellas en el nivel metalógico, por ejemplo, sobre cómo debe entenderse correctamente la idea de validez, o cómo ciertos argumentos informales pueden representarse formalmente mejor. (Los asteriscos de la fig. 9 indican dónde se localiza la rivalidad metalógica.)

Puede ser útil señalar que esta postura es capaz de reconciliar al menos algunas de las consideraciones que se han tomado en cuenta para indicar al monismo, el pluralismo local o el instrumentalismo. He admitido —lo que el monismo subraya principalmente— que un sistema lógico puede competir realmente con otro, en el sentido fuerte en que no puedan ser los dos correctos; sólo he negado que los sistemas lógicos siempre deban competir de este modo. También he propugnado el reconocimiento de la competencia metalógica en la que creo que se puede fundir la rivalidad lógica. Y la sugerencia de que diferentes representaciones formales pueden ser mejores para propósitos diferentes quizás ofrece alguna comodidad a los pluralistas locales. En este campo tengo menos que ofrecer a modo de concesiones a los instrumentalistas. En la sección siguiente, sin embargo, argumentaré a favor de una aproximación bastante radical a la epistemología de la lógica, aproximación que sería bastante compatible con la de un instrumentalista.

2 Cuestiones epistemológicas

> ... ningún enunciado es inmune a revisión. Incluso se ha propuesto la revisión de la ley del tercero excluido como un medio de simplificar la mecánica cuántica; y ¿qué diferencia hay, en principio, entre un cambio tal y el cambio por el que Kepler desbancó a Ptolomeo, o Einstein a Newton, o Darwin a Aristóteles? (Quine, 1951, pág. 43.)

Quine está afirmando que la *lógica es revisable*. Yo creo que tiene razón; pero las cuestiones epistemológicas que plantea esta afirmación son mucho más complejas de lo que se podría sospechar del elegante pero bastante descuidado tratamiento que reciben en "Dos dogmas".

Primero, necesitamos tener claro exactamente qué es lo que se quiere decir con la afirmación de que la lógica es revisable —y con la misma importancia, qué es lo que *no* se quiere decir con ella. Lo que yo quiero decir, en todo caso, no es que las verdades de la lógica podrían haber sido diferentes de lo que son, sino que las verdades de la lógica podrían ser otras que las que nosotros consideramos que son, es decir, *podríamos estar equivocados acerca de lo que son las verdades de la lógica*, por ejemplo, al suponer que la ley del tercio excluso es una de ellas.

Por tanto, un modo mejor de plantear la pregunta, para que se haga más patente su carácter epistemológico, es ésta: ¿alcanza a la lógica el falibilismo? Sin embargo, incluso esta formulación necesita de depuración posterior, pues la naturaleza del falibilismo con frecuencia es mal entendida.

¿Qué es el falibilismo?

Decir que una persona (o un grupo de personas, "la comunidad científica", por ejemplo) es falible —usaré "falible" en el sentido de "falible cognitivamente", que es como decir falible *con respecto a las creencias*, y no, por ejemplo, con respecto a las promesas, resoluciones, etc.— es decir que es susceptible de mantener falsas creencias; decir que un método es falible es decir que es susceptible de producir falsos resultados; por supuesto, una persona puede ser falible porque utiliza métodos falibles para adquirir creencias, como consultar las entrañas o el horóscopo. A mí me parece innegable que las personas son falibles —todos somos susceptibles de mantener falsas creencias; sabemos que son falsas ciertas creencias que la gente solía mantener—; en un tiempo, por ejemplo, la gente creyó que el sol se movía alrededor de la tierra, que la tierra era plana..., etc. —y es razonable, así como modesto, suponer que también nosotros creemos cosas que son falsas, aunque, por supuesto, no sabemos cuáles de las cosas que creemos son las falsas, y naturalmente dejaríamos de creer en ellas si lo supiéramos.

Sin embargo, con frecuencia, los epistemólogos han pensado que, con respecto a ciertos *tipos* de creencias —creencias sobre la propia inmediata experiencia sensible son un ejemplo predilecto—, las personas pueden ser *in*falibles: son susceptibles de tener falsas creencias sobre astronomía, geografía..., etc., pero no son susceptibles de estar equivocadas sobre si tienen dolor, están viendo una mancha roja..., etc. Y tampoco, han argumentado algunos autores, somos susceptibles de estar equivocados sobre las verdades de la lógica; la lógica, piensan, tiene una seguridad epistemológica especial. Popper, por ejemplo, aunque subraya nuestro falibilismo cuando se trata de conjeturas científicas, sin embargo, parece seguro de que la lógica es de fiar; véase 1960 para el falibilismo, y cfr. 1970 para su rechazo a extender el falibilismo a la lógica.

¿Alcanza a la lógica el falibilismo?

(i) *necesidad.* Ahora bien, ¿por qué se encontraría buena voluntad para aceptar que podríamos estar equivocados en lo que consideramos que son las leyes de la física, pero mala voluntad para aceptar que podríamos estar equivocados en lo que consideramos

que son las leyes de la lógica? Una razón importante —importante al menos porque se basa en una confusión significativa— deriva de la presunta necesidad de las leyes lógicas. El argumento sería algo parecido a esto: las leyes de la lógica son necesarias, que es como decir que no podrían ser otra cosa que verdaderas; por tanto, ya que una ley lógica no puede ser falsa, la creencia en una ley lógica no puede ser equivocada y, por tanto, es infalible. Tengo la pequeña duda de que este argumento es poco seguro. (Las verdades de las matemáticas también se supone que son necesarias. Pero, sin embargo, estamos dispuestos a mantener creencias matemáticas falsas, el resultado de equivocaciones en el cálculo, por ejemplo. Y si las leyes de la física son, como algunos suponen, físicamente necesarias, normalmente no se piensa que entraña que somos infaliblemente capaces de decir lo que *son* las leyes de la física.) Pero ¿qué tiene de malo el argumento de que, ya que las leyes de la lógica son necesarias, el falibilismo no alcanza a la lógica?

El argumento se equivoca en dos formas. Primero, depende del uso de "falible" como predicado, no de personas, sino de proposiciones: un predicado que presumiblemente significa "posiblemente falso". Ahora bien, es bastante cierto que, si las leyes de la lógica son necesarias, no son posiblemente falsas y por tanto en este sentido, son "infalibles". Pero la tesis de que algunas proposiciones son posiblemente falsas (a lo que llamaré "falibilidad de la proposición") es una tesis lógica poco interesante, que no debería confundirse con la interesante tesis epistemológica de que somos susceptibles de mantener falsas creencias (a lo que llamaré "falibilismo del agente"). Y el infalibilismo de la proposición no entraña el infalibilismo del agente: aun cuando las leyes de la lógica no son posiblemente falsas, esto no garantiza de ningún modo que no seamos susceptibles de mantener creencias lógicas falsas. Al postular que somos falibles en nuestras creencias lógicas (que el falibilismo del agente *alcanza* a la lógica) no estoy, por supuesto, afirmando la tesis contradictoria de que aunque "$p \vee -p$" es necesario, podríamos creer falsamente que $p \vee -p$; más bien, estoy postulando que, aunque "$p \vee -p$" es necesario, podríamos creer falsamente que $-(p \vee -p)$, o si no, quizás, aunque "$p \vee -p$" no es necesario, falsamente creemos que lo es. (Elijo deliberadamente el tercio excluso como ejemplo de una pretendida ley lógica, ya que hay, por supuesto, discusión acerca de su status.) En segundo lugar, el argumento proporciona una engañosa plausibilidad por la facilidad con que se confunde la tesis de que algunas proposiciones son posiblemente falsas con la tesis de que algunas proposiciones son contingentes. Si las leyes de la lógica son necesarias, nuestras creencias lógicas, en efecto, no serán contingentes, sino necesariamente verdaderas o necesariamente falsas. Pero "posiblemente falso" no debería equipararse con "contingente", pues creencias *necesariamente falsas* son posiblemente falsas[4].

La confianza en que la lógica es inalterable ha sido con bastante frecuencia la base para negar que la lógica es revisable. Una vez esté claro —como espero que lo está ya— que la necesidad de los principios lógicos no nos muestra que sean infalibles lógicamente, estará claro, también, que si la lógica es irrevisable, *no* es porque sea inalterable.

Ahora bien, una razón para creer que somos falibles cuando se trata de nuestras creencias sobre el mundo es que sabemos que una vez unas personas creyeron con seguridad lo que ahora (pensamos) sabemos que es falso; y, aunque estamos seguros de que estaban equivocados al pensar, por ejemplo, que la tierra es plana, el hecho de que sus creencias resultaran ser falsas es para nosotros una razón para aceptar que algunas de nuestras creencias también puede resultar que estén equivocadas. Y yo pensaría que parecidas razones operan, para una modestia comparable, sobre nuestras creencias lógicas. Por ejemplo, Kant escribió que "En nuestros tiempos no ha habido ningún lógico famoso, y, en efecto, no precisamos de nuevos descubrimientos en Lógica..." (1800, pág. 11); su seguridad de que la lógica era una ciencia acabada nos parece —con la ventaja de la retrospección, después de los enormes avances realizados en lógica desde el último cuarto del siglo XIX— dar muestras de un notable y curioso exceso de confianza. (La confianza de Kant en la lógica aristotélica estaba basada en la creencia de que la lógica incorpora las "formas del pensamiento", que no podemos sino pensar de acuerdo con estos principios. La discusión de estas ideas tendrá lugar más adelante.) O, por otra parte: Frege pensó que la reducción de la aritmética a la lógica garantizaría epistemológicamente la aritmética, porque consideró que las verdades de la lógica eran autoevidentes; nosotros, sin embargo, al saber que los axiomas "autoevidentes" de Frege eran inconsistentes, estamos capacitados para encontrar su confianza fuera de lugar. (Lakatos, 1963-4, en un magnífico ensayo filosófico sobre la historia de las matemáticas, derriba de modo similar la tendencia a situar a las matemáticas sobre un pedestal epistemológico.) Otra razón en contra del exceso de confianza epistemológico es el conocimiento de

[4] Si estoy acertada en que un interesante falibilismo auténticamente epistemológico hará "falible" a un predicado de personas más que de proposiciones, esto tiene la consecuencia de que el intento de Popper de hacer sitio para el falibilismo en una "epistemología sin sujeto de conocimiento" (véase el artículo de ese título de 1972) está mal orientado. Y si estoy acertada en que el falibilismo del agente puede alcanzar bastante consecuentemente a contenidos, la verdad de los cuales es necesaria, no hay necesidad de turbarse (tal como incluso un "falibilista contrito" como C. S. Peirce manifiesta) para extender el falibilismo a las matemáticas.

que otras personas mantienen, con tanta confianza, creencias incompatibles con las propias. Y este motivo opera en la esfera de la lógica también; la misma pluralidad de los sistemas lógicos habla contra nuestra posesión de una capacidad infalible para averiguar las verdades de la lógica.

(ii) *autoevidencia.* Con todo, la idea de que las verdades de la lógica son autoevidentes necesita un examen más minucioso. ¿Qué significa afirmar que una proposición es autoevidente? Probablemente algo que especifica que es evidentemente verdadera. Pero una vez que se ha planteado así, la dificultad con el concepto de autoevidencia no puede disimularse. El hecho de que una proposición sea obvia, lamentablemente, no es ninguna garantía de que sea verdadera. (Es oportuno observar que personas diferentes y épocas diferentes, encuentran "obvias" proposiciones diferentes e incluso incompatibles —que unos son esclavos por naturaleza, que todos los hombres son iguales...) Si se dice que los axiomas inconsistentes de Frege sólo *parecían* autoevidentes, pero que, en realidad, no podían serlo, o que eran autoevidentes pero desafortunadamente no eran verdaderos, la autoevidencia debe fracasar al proporcionar una garantía epistemológica; porque o (según el último supuesto) una proposición puede ser autoevidente pero falsa, o si no (según el primer supuesto) aunque si una proposición es autoevidente entonces es, en efecto, verdadera, no tenemos ningún modo seguro para decir cuándo una proposición es autoevidente en realidad[5].

(iii) *analiticidad.* Otra razón para dudar de la revisabilidad de la lógica parece derivar de la idea, primeramente, de que las verdades lógicas son analíticas y además que las verdades analíticas son, por así decirlo, evidentes. Si A es verdadero en virtud de su significado, lo que se piensa es entonces que nadie que lo entienda puede dejar de decir que *es* verdadero. Creo que hay lugar para la duda acerca de si un argumento realmente convincente podría desarrollarse de esta manera; pues la idea de "verdad en virtud del significado" está lejos de ser transparente, no sólo a causa del "significado" (como Quine mucho tiempo ha propugnado), sino también a causa del "en virtud del". Y aun suponiendo que se pudiera, hay lugar para una duda posterior acerca de si su conclusión perjudicaría seriamente al falibilismo, pues aun cuando, si se entendiese una verdad lógica correctamente, no cabría sino reconocer su verdad, esto garantizaría la corrección de las creencias lógicas

[5] Mis comentarios tienen mucho en común con la muy hábil crítica de Peirce (1868) de la infalible facultad de "intuición" de la cual Descartes nos suponía poseedores.

sólo si también se tuviera un medio a prueba de tontos para estar seguros de haber entendido correctamente un candidato a verdad lógica. (La similitud estructural entre este comentario y la critica previa del argumento de la "autoevidencia" es digna de observarse.)

Una digresión: "Dos dogmas" de nuevo

Este parece el lugar adecuado para algunas observaciones sobre la estructura del argumento de Quine en "Dos dogmas". El articulo comienza (simplifico, pero espero que no equivocadamente) con un ataque a la distinción analítico/sintético, y se cierra con un alegato a favor de la revisabilidad de la lógica. ¿Cuál es la conexión entre los dos?

Se puede interpretar que Quine propugna la revisabilidad de la lógica como un argumento en contra de la concepción de analiticidad de los positivistas lógicos. El positivista considera que el significado de una oración está dado por sus condiciones de verificación; y, por tanto, considera que un enunciado es analítico, o verdadero en virtud de su significado, justamente en el caso de que se verifique pase lo que pase. La idea metafísica de analiticidad va junto con la idea epistemológica de una *a prioridad*; por lo que sería adecuado para Quine atacar la afirmación de que la lógica es analítica, en este sentido, argumentando que la lógica es revisable. Según esta interpretación, la revisabilidad de la lógica no es una conclusión, sino una premisa, del argumento de "Dos dogmas".

Otra posibilidad es ver el ataque a la analiticidad como premisa, y el alegato a favor de la revisabilidad de la lógica, como conclusión. Sin embargo, si el argumento es: si las leyes de la lógica fueran analíticas no serían revisables, pero ya que hay verdades no analíticas, las leyes de la lógica no son analíticas y, por tanto, son revisables —sería un usual argumento. Es inválido, teniendo la forma "$A \rightarrow -B$, $-A$, por tanto B". Una premisa es falsa ya que, como acabo de argumentar, el que A sea analítico no impide que estemos equivocados acerca de ello. Y la otra premisa no se ha establecido: Quine ataca el segundo disyunto de la definición "fregeana" de verdad analítica como una verdad lógica o reducible a una verdad lógica mediante la sustitución de sinónimos, pero esto apenas puede mostrar que las verdades lógicas no son analíticas, pues se califican en el primer disyunto.

No obstante, esta interpretación merece alguna atención: porque nos permite entender el creciente conservadurismo más reciente de Quine sobre la lógica. El ataque de "Dos dogmas" sobre la sinonimia, etc., podría amenazar con una consideración de las verdades lógicas como analíticas en tanto que *verdaderas en virtud del significado de las constantes lógicas*. Ahora bien, en *Palabra y Objeto*, Quine renueva su escéptico ataque a las nociones de significado, pero hace una excepción en el caso de las conectivas lógicas, las cuales, afirma, tienen significado determinado (1960a, cap. 2); y esto prepara el terreno para su aceptación (1970, cap. 6) de un un argumento de varianza del significado en el sentido de que los teoremas de las lógicas clásicas y divergentes son, del mismo modo, verdaderos en virtud de las conectivas (clásicas o divergentes); lo cual, a su vez, parece llevarle a comprometer su anterior insistencia en que el falibilismo alcanza incluso a la lógica.

Revisión de la lógica

Si el falibilismo alcanza a la lógica, si, como he afirmado, somos susceptibles de estar equivocados en nuestras creencias sobre la lógica, entonces sería prudente estar dispuestos, si fuera necesario, a revisar nuestras opiniones lógicas. Pero esto no quiere decir que se emprendan a la ligera las revisiones de la lógica, pues la extremada generalidad de los principios lógicos supone que tales revisiones tendrán consecuencias de lo más trascendental; la lógica *es* revisable, pero las razones para su revisión es mejor que sean buenas. Como argumenté en el cap. 11, § 3, los argumentos presentados a favor de las lógicas desviadas han sido, con demasiada frecuencia, bastante débiles.

3 LÓGICA Y PENSAMIENTO

La confianza de Kant en la irrevisabilidad de la lógica aristotélica descansaba en la idea de que los principios lógicos representan "las formas de pensamiento", que no podemos pensar si no es de acuerdo con ellas: una idea que plantea un montón de preguntas intrigantes acerca de qué tiene que ver la lógica con "el modo en que pensamos".

Aunque en un momento era bastante normal suponer que los principios de la lógica son "las leyes del pensamiento" (véase Boole, 1854), la vigorosa crítica de Frege fue tan influyente que, últimamente, ha habido bastante poco apoyo para el "psicologismo" de cualquier forma o condición. Sin embargo, los argumentos de Frege en contra del psicologismo sospecho que son menos concluyentes, y por lo menos alguna forma de psicologismo más plausible, que lo que hoy en día se puede suponer. Una revaloración completa del psicologismo exigiría, sin embargo, una consideración más extensa y sofisticada de la naturaleza del pensamiento que la que soy capaz de presentar; por tanto, lo que sigue puede que sea, en el mejor de los casos, incompleto.

Se puede empezar por distinguir —la distinción es bastante tosca, pero no obstante puede ser útil como punto de partida— tres tipos de posturas:

(i) la lógica es descriptiva de los procesos mentales (describe cómo *pensamos* o quizás cómo *tenemos que* hacerlo)
(ii) la lógica es preceptiva de los procesos mentales (prescribe cómo *deberíamos* pensar)
(iii) la lógica no tiene nada que ver con los procesos mentales.

A estas posturas se las podría llamar *psicologismo fuerte*, *psicologismo débil* y *antipsicologismo*, respectivamente. Ejemplos: Kant mantuvo algo parecido a (i); Peirce una versión de (ii); Frege (iii).

En lo que sigue presentaré algunos argumentos a favor de una forma de psicologismo débil bastante cercana a la adoptada por Peirce (1930-58, 3, 161 y ss.): que la lógica es normativa con respecto al razonamiento. Pasaré entonces a señalar algunas ventajas del psicologismo débil frente al antipsicologismo, por una parte, y al psicologismo fuerte, por otra.

La lógica ante todo trata de los *argumentos*: ¿Cómo, entonces, puede relacionarse con los procesos mentales que constituyen el *razonamiento?* Abordaré esta pregunta en dos etapas, ofreciendo, primero una respuesta platónica, y luego, una versión nominalista de esa respuesta; la razón para esta estrategia es que la conexión entre lógica y pensamiento se pone de relieve con más nitidez en la consideración platónica, pero creo que está mejor explicada, aunque con menos simplicidad, en la versión nominalista.

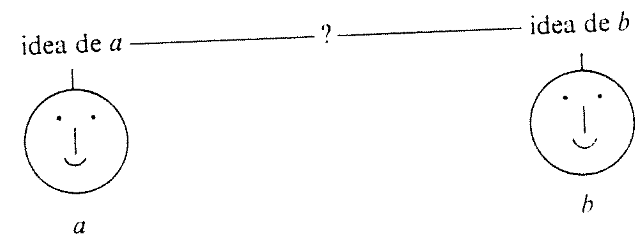
La respuesta platónica: La lógica trata de la (in)validez de los argumentos, de la conexión entre las premisas y la conclusión; las relaciones lógicas son relaciones entre proposiciones tales como el entrañamiento o la incompatibilidad. Razonar es un (una cierta especie de) proceso mental, tal como llegar a creer que q por la fuerza de la propia creencia en que p (inferir q de p), o llegar a reconocer que si p fuera el caso, entonces q sería el caso; y creer que p, o preguntarnos si, o qué si, p, es colocarse en cierta relación a una proposición. Por tanto, la lógica es normativa con respecto al razonamiento en este sentido: que si, por ejemplo, se infiere q de p, entonces, si el argumento de p a q es válido, la inferencia es *segura*, con esto se garantiza que no tengamos como resultado el mantenimiento de una falsa creencia sobre la base de una verdadera.

La versión nominalista: que s cree que p, o se pregunta si, o qué si, p, puede analizarse, en el fondo, en términos de una complicada relación entre s y la oración "p"; y el modo de hablar platónico de la creencia en o la recepción de una proposición ha de considerarse como una oportuna taquigrafía de este complicado análisis. La lógica trata de la validez de los argumentos, los cuales,

sin embargo, se han de concebir (cap. 2, § 1) como trozos del discurso/cadenas de oraciones y el modo de hablar platónico de las relaciones lógicas entre proposiciones se ha de considerar, de nuevo, como una taquigrafía oportuna (especialmente, para requisitos bastante complicados acerca de qué oraciones han de considerarse como intersustituibles, cap. 6, § 4). Una vez más, se sigue que la lógica es normativa con respecto al razonamiento, en el sentido antes explicado.

La versión nominalista del psicologismo débil creo que es preferible a la platónica, por razones que surgirán de la consideración de un argumento de Frege contra el psicologismo.

Las objeciones de Frege al psicologismo son bastante complejas y sólo consideraré el argumento que es más relevante para la postura que he defendido[6]. Este argumento se desarrolla como sigue. La lógica no tiene nada que ver con los procesos mentales; pues la lógica es objetiva y pública, mientras que lo mental, según Frege, es subjetivo y privado. Por esto es por lo que Frege trata, por tanto, de insistir (véase especialmente Frege, 1918; y cfr. pág. 82n) en que el sentido de una oración no es una idea (una entidad mental), sino un pensamiento (*Gedanke*: un objeto abstracto, una proposición). Ya que las ideas son mentales, argumenta Frege, son esencialmente privadas; tú no puedes tener mi idea lo mismo que no puedes tener mi dolor de cabeza. Si el sentido de una oración fuera privado, entidad mental, una idea en el sentido de Frege, habría un misterio acerca de la relación entre la idea de una persona y la de otra:



Las proposiciones, sin embargo, son públicas: tú y yo podemos "captar" los dos la misma proposición, y esto es lo que hace posible, que haya conocimiento público, objetivo[7].

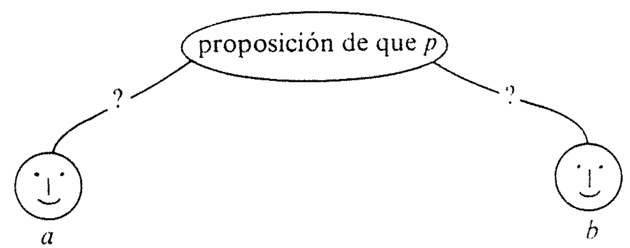
Este argumento se podría cuestionar en más de un punto: por ejemplo, ¿por qué supone Frege que todo lo mental es subjetivo y

[6] Ignoraré por completo los argumentos de Frege contra las explicaciones psicologistas del número; tan sólo observaré que en vista de su logicismo, habría entendido que estos argumentos se referían indirectamente al psicologismo con respecto a la lógica.

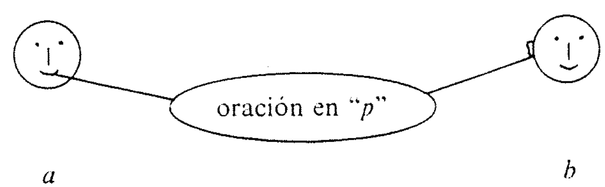
[7] Las razones de Popper para divorciar la epistemología de la psicología son muy parecidas.

privado? ¿Es relevante el que la psicología con la que estaba familiarizado era la introspeccionista? Pero, de todos modos, está bastante claro que el argumento *no* nos obliga a divorciar a la lógica de los procesos mentales de la manera que supone Frege. Pues la postulación de proposiciones sólo garantizaría la publicidad del conocimiento si las proposiciones son no solamente *objetivas*, sino también *accesibles*, si podemos "captarlas"; y esto es exactamente lo que exige la versión platónica del argumento a favor del psicologismo débil.

De hecho, sin embargo, Frege no tiene nada muy consistente que decir para mitigar el carácter misterioso de nuestra supuesta habilidad para "captar" su *Gedanken:*



Pero este misterio se *puede* disipar al concentrarnos no en las ideas (que crean problema sobre la objetividad), ni en las proposiciones (que crean problema sobre la accesibilidad), sino en las *oraciones;* pues el comportamiento verbal de los usuarios de un lenguaje es las dos cosas, *objetivo* y *accesible:*



(Dewey vio esto: véase 1929, pág. 196.) Y esto da razón para preferir, como propugné, la versión nominalista del argumento a favor del psicologismo débil.

La lógica, sugerí, es preceptiva del razonamiento en el limitado sentido de que la inferencia de acuerdo con los principios lógicos sea segura. (Por supuesto, la seguridad no es necesario que tenga una consideración primordial; bastante racionalmente, se podrían preferir procedimientos seguros, pero relativamente poco interesantes; cfr. la defensa de de Bono, por ejemplo, en 1969, del "pensar lateral".) Sin embargo, es importante que desde el punto de vista del psicologista débil, aunque la lógica es aplicable al razonamiento, la

validez de un argumento consiste en su carácter preservador de verdad; lo que en ningún sentido es una propiedad psicológica. En consecuencia, el psicologismo débil evita la dificultad principal del psicologismo fuerte, el problema de justificar el error lógico: pues ya que, sin duda, las personas, de cuando en cuando, argumentan inválidamente, ¿cómo podría consistir la validez de un argumento en su conformidad con el modo en que pensamos? Esto no es decir que el psicologismo fuerte es rotundamente incompatible con los errores lógicos; sino que los dos pueden conciliarse sólo por medio de una explicación de tales errores como el resultado de alguna irregularidad o mal funcionamiento de nuestros poderes de razonamiento. (Según Kant, los errores lógicos son el resultado de una inadvertida influencia de la sensibilidad sobre el juicio.) Sin embargo, su mucho más rápida conciliación con el falibilismo creo que habla a favor del psicologismo débil sobre el fuerte.

Hay, inevitablemente, muchas preguntas intrigantes que esto deja sin responder: por ejemplo, ¿qué es lo que distingue exactamente al estudio lógico del razonamiento del psicológico? (No puede ser, como se ha supuesto a veces, que la psicología, a diferencia de la lógica, nunca es normativa, ni incluso que nunca es normativa con respecto a la verdad; considérense, por ejemplo, los estudios psicológicos de las condiciones de la percepción fiable/ilusoria.) ¿Qué consecuencias tendría el psicologismo sobre la lógica para preguntas acerca de las relaciones entre epistemología y psicología? ¿Qué consecuencias tendría para el psicologismo (particularmente, en vista de la afirmación de Chomsky de que ciertas estructuras gramaticales son innatas) la conjetura de que la forma lógica se pueda identificar con la forma gramatical?

¡Es bueno saber (tomo prestada una frase de Davidson) que no nos quedaremos sin trabajo!

Glosario

* al lado de un término indica que tiene una entrada independiente. Para terminología que no se cite aquí, el lector puede encontrar útil consultar el *Dictionary of Philosophy* (Runes, 1966) o las entradas de *Logical terms, glossary of*, en Edwards, 1967.

Analítico/sintético
: Un juicio *analíticamente* verdadero es tal que el concepto de su predicado está contenido en su sujeto, o tal que su negación es contradictoria (Kant); una proposición analíticamente verdadera es o una verdad lógica*, o si no, reducible a verdad lógica mediante definiciones en términos puramente lógicos (Frege; y véase logicismo*); un enunciado analíticamente verdadero es verdadero únicamente en virtud del significado de sus términos (positivistas lógicos*). "Analítico" se utiliza generalmente en equivalencia con "analíticamente verdadero"; la negación* de una verdad analítica es analíticamente falsa. "Sintético" se utiliza generalmente en equivalencia con "ni analíticamente verdadero ni analíticamente falso". Véase la discusión de la crítica de la analiticidad de Quine, págs. 196-8 y 262-3.

A priori/a posteriori
: Una proposición es *a priori* si se la puede conocer independientemente de la experiencia, de lo contrario es *a posteriori* (una distinción epistemológica en contraste con la distinción metafísica de *analítico/sintético**). Véase la discusión del falibilismo con respecto a la lógica, cap. 12, § 2.

Atómico
: Una fbf* atómica del cálculo de oraciones es una letra de oración (por ejemplo, "p"), en contraste con una fbf* compuesta o "molecular" (por ejemplo, "$p \vee q$"). Una fbf atómica del cálculo de predicados es una letra de predicado n-posicional* seguida de n variables* o términos singulares. Un enunciado atómico, análogamente, es un enunciado que no contiene otros enunciados como componentes.

Atomismo lógico
: Escuela de filosofía (primer Wittgenstein, Russell) que trata de analizar lógicamente la estructura del mundo en sus componentes fundamentales (los "átomos lógicos"). Véase la discusión de la teoría de la correspondencia de la verdad, cap. 7, § 2, y de las afinidades con el programa de Davidson, pág. 147n.

Axioma
: Una fbf* A es un *axioma de* L si A está establecida y su verdad incuestionada en el sistema L (trivialmente, todos los axiomas de L son teoremas* de L). Una presentación axiomática de la lógica utiliza axiomas así como reglas de inferencia*. Véase cap. 2, § 3.

Bivalencia

Toda fbf* (oración, enunciado, proposición) es o verdadera o falsa; véase también *tercio excluso**. Véase cap. 11, § 3.

Combinatoria, lógica

Una rama de la lógica formal en la que las variables* se eliminan en favor de símbolos de función. Véase la discusión del criterio ontológico de Quine, cap. 4, § 2.

Completo

(i) Un sistema formal es *completo en sentido débil* si toda fbf* que es lógicamente verdadera* en el sistema es un teorema* del sistema; o *completo en sentido fuerte*, si cualquier nuevo axioma* independiente* que se le añada lo hace inconsistente*. Ejemplos: el cálculo de oraciones es completo en sentido fuerte; los sistemas modales corrientes son completos en sentido débil; la teoría de conjuntos y la aritmética son incompletos. Véase la discusión de la completud como criterio para considerar a un sistema como un sistema de lógica, cap. 1, § 2.

(ii) En lo que respecta a *completud funcional*, véase *veritativo-funcional**.

Condicional

Los operadores "$\rightarrow$", "$\dashv$", etc.

Una fbf* de la forma $A \rightarrow B$ (o enunciado de la forma "Si A entonces B") se llama también una fbf condicional o hipotética. "A" se llama el *antecedente* y "B" el *consecuente* del condicional. Un condicional *subjuntivo* es el que tiene el verbo en subjuntivo (como "Si el impuesto sobre la renta se redujera a la mitad, estaríamos todos encantados"); un condicional *contrafáctico* es un condicional subjuntivo que implica que su antecedente es falso (como "Si se hubiera reducido a la mitad el impuesto sobre la renta en el último presupuesto, todos estaríamos encantados").

Véase cap. 3, § 3; cap. 10, § 6.

Condiciones necesarias/suficientes

A es *condición necesaria* de B, si B no puede ser el caso si no es A; A es *condición suficiente* de B, si, si A es el caso, es B.

Conjunción

Una fbf* (enunciado) de la forma "A & B" ("A y B").

Conjunto

"Una colección en un todo... de objetos definidos y distinguibles" (Cantor); sin embargo, la teoría de conjuntos incluye el conjunto vacío, que no tiene miembros. Véase la paradoja* de Russell, págs. 159-160. "$\{a, b, c\}$" significa "el conjunto que consiste de a, b, c"; "$\{x|Fx\}$" significa "el conjunto de las cosas que son F"; "$a \in \{x|Fx\}$" significa "a es un miembro del conjunto de las cosas que son F". (En la teoría de conjuntos de Gödel-von Neumann-Bernays se hace una distinción entre los conjuntos, que pueden tanto tener miembros como ellos mismos ser miembros, y las clases, que tienen miembros, pero ellas mismas no pueden ser miembros.)

Consecuencia

Una fbf* (enunciado) B es una consecuencia lógica de A sii hay un argumento válido* de A a B.

Consistente

Un sistema formal es *consistente* sii ninguna fbf* de la forma "A & $-A$" es un teorema*; o sii no toda fbf del sistema es un teorema; o (en el sentido de Post, aplicable al cálculo de oraciones) sii ninguna letra de oración sola es un teorema.

Constante

Una *constante* es un símbolo empleado siempre para representar la misma cosa (como los términos singulares como "a", "b", ..., etc., o los operadores como "&", "$\vee$", ..., etc.) en contraste con las variables* (como "x", "y", "z", ..., etc.) que recorren un dominio* de objetos.

Contradicción

Fbf* de la forma "A & $-A$"; enunciado de la forma "A y no A". *Principio de no contradicción:* $-(A$ & $-A)$; o: ninguna fbf (oración, enunciado, proposición) es verdadera y falsa a la vez.

Contradictorio

El contradictorio de una fbf* (enunciado) A es una fbf* (enunciado) que debe ser falso si A es verdadero y verdadero si A es falso.

Contrario

Las fbfs* (enunciados) A y B son contrarios si no pueden ser los dos verdaderos, pero pueden ser los dos falsos.

Correcto

(i) Un *argumento* es *correcto* si (i) es válido* y (ii) sus premisas, y, por tanto, su conclusión, son verdaderas.

(ii) Un *sistema lógico* es *correcto* sii todos sus teoremas son lógicamente verdaderos*; corrección es la conversa de completud*.

Correspondencia de uno a uno

Dos conjuntos* x e y están en correspondencia de uno a uno si hay una *relación* uno a uno*, R, por la cual cada miembro de x se relaciona con exactamente un miembro de y, y cada miembro de y con exactamente un miembro de x.

Cuantificador

Expresión ("$(\exists ...)$" —el cuantificador existencial— "$(...)$" —el cuantificador *universal*) que liga *variables**. Véase cap. 4.

Decidible

Un sistema es *decidible* si hay un procedimiento mecánico ("un procedimiento de decisión") para determinar, para cada fbf* del sistema, si esa fbf es un teorema* o no. Ejemplos: el cálculo de oraciones es decidible; todo el cálculo de predicados (incluyendo los predicados poliádicos* así como los monádicos*) no lo es. Las *tablas de verdad* proporcionan el procedimiento de decisión para el cálculo de oraciones; una prueba de tablas de verdad determina si una fbf es una tautología*, y, por los resultados de corrección* y completud*, todas y solamente las tautologías* son teoremas*.

Deducción

Una secuencia de fbfs* (de L) es una deducción (en L) de B a partir de $A_1 \dots A_n$ sii es un argumento válido* (en L) con $A_1 \dots A_n$ como premisas y B como conclusión.

Deducción natural

Una presentación de deducción natural de un sistema lógico descansa en reglas de inferencia* más que en axiomas*. Véase cap. 2, § 3.

Definición

Una definición *explícita* define una expresión (el *definiendum*) por medio de otra (el *definiens*) que puede reemplazar a la primera dondequiera que ésta ocurra. Una definición *contextual* proporciona una sustitución para ciertas expresiones más largas en las que ocurre el *definiendum*, pero no un equivalente para esa misma expresión. (Si los x se pueden definir contextualmente en términos de y, a veces se dice que los x son construcciones lógicas fuera de las y, y que "x" es un *símbolo incompleto**.) Una definición *recursiva* proporciona una regla para eliminar el *definiendum* en un número finito de pasos. Un conjunto de axiomas* se dice a veces que da una definición *implícita* de sus términos primitivos*. Véase cap. 3, § 1, para la interdefinibilidad de las conectivas; cap. 4, § 3, para la definición contextual de Russell de las descripciones definidas; cap. 7, § 5, para la definición recursiva de satisfacción de Tarski; pág. 125, para las condiciones formales en las definiciones.

Descripción definida

Expresión de la forma "El tal y tal", que se escribe, formalmente, "$(\iota x)\, Fx$". Véase cap. 5, § 3.

Disposicional
Un predicado disposicional adscribe una tendencia o "hábito"; en castellano muchos de tales predicados terminan en "-ble" (como: "irritable", "soluble"). Los enunciados disposicionales ("este terrón de azúcar es soluble") son equivalentes a los *condicionales subjuntivos** ("si este terrón de azúcar se pusiera en agua, se disolvería"). Véase cap. 10, § 3.

Disyunción
Fbf* (enunciado) de la forma "$A \vee B$". *Dilema disyuntivo* es la forma del argumento: si $A \vdash C$, $B \vdash C$, entonces $A \vee B \vdash C$.

Divergencia
L_1 es una divergencia de L_2 si tiene un conjunto diferente de teoremas/inferencias válidas que esencialmente suponen un vocabulario compartido con L_2. Una divergencia de la lógica clásica es una *lógica divergente* Véase caps. 9, 11, 12.

Doble negación, principio de
$A \equiv - - A$. Véase la discusión de la lógica intuicionista, cap. 11, § 4.

Dominio (Universo de discurso)
Rango de las variables* de una teoría. Véase cap. 4, § 1.

Entimema
Argumento con una premisa suprimida.

Epistemología
Teoría del conocimiento.

Equivalencia
Dos fbfs (enunciados) son *lógicamente equivalentes* si necesariamente tienen el mismo valor de verdad. Son *materialmente equivalentes* si tienen el mismo valor de verdad.

Esquema (T)
La condición de adecuación material de Tarski exige que una definición aceptable de verdad tenga como consecuencia todas las instancias del esquema (T):

$$S \text{ es verdadera sii } p$$

donde "S" nombra a la oración del lado de la derecha. Véase cap. 7, §§ 5 y 6.

Extensión
L_1 es una extensión de L_2 si contiene nuevo vocabulario, además del vocabulario compartido con L_2, y tiene nuevos teoremas*/inferencias* válidas que esencialmente suponen el nuevo vocabulario. Una extensión de la lógica clásica es la *lógica extendida*. Véase caps. 9, 10, 12.

Extensión/intensión
Referencia (extensión) versus *sentido (intensión)* de una expresión. Para un término singular, la extensión es su referente; para un predicado, el conjunto de cosas para las que es verdadero; para una oración, su valor de verdad. Dos expresiones con la misma extensión son *co-extensivas*. Terminología relacionada: *Bedeutung* (= extensión) versus *Sinn* (= intensión) de una expresión (Frege); *denotación* versus *connotación* (Mill); contextos *extensionales* versus contextos *intensionales**. Véase la discusión de Frege del sentido y la referencia, cap. 5, § 2; cfr. distinción de Quine entre la teoría de la referencia y la teoría del significado, pág. 141.

Extensional/intensional
Un contexto es *extensional* si expresiones co-referenciales —términos singulares con la misma denotación, predicados con la misma extensión u oraciones con el mismo valor de verdad— son sustituibles en él sin que cambie el valor de verdad del todo, "*salva veritate*", es decir, si para él se mantiene la *ley de Leibniz*; de lo contrario, es *intensional*. Ejemplos: "No es el caso que..." es extensional, "Necesariamente..." o "*s* cree que..." son intensionales. Terminología relacionada: contexto *oblicuo* (= intensional) (Frege); contexto *referencialmente transparente* (= extensional) versus contexto *referencialmente opaco* (= intensional), *ocurrencia puramente referencial* (es decir, ocurrencia en un contexto extensional) de un término singular (Quine); conectiva *veritativo-funcional** (= en las oraciones, operador extensional formador de oración). Véase discusión del programa de Davidson, cap. 7, § 5; cfr. crítica de Quine de la distinción analítico/sintético, cap. 10, § 1.

Fbf
Fórmula bien formada, es decir, cadena de símbolos de un lenguaje formal construida correctamente con respecto a sus *reglas de formación*. Una *fórmula* es una cadena de símbolos del lenguaje formal.

Finito-infinito
Un conjunto es *enumerablemente* infinito si tiene un subconjunto propio tal que sus miembros se pueden poner en correspondencia* de uno a uno con los miembros de ese subconjunto propio. Un conjunto es finito si no es infinito. Un conjunto es *enumerablemente* finito si se puede poner en correspondencia de uno a uno con los números naturales.

Formalismo
Escuela de filosofía de la matemática (Hilbert, Curry) caracterizada por la opinión de que los números se pueden identificar con marcas sobre el papel. Véase la discusión del planteamiento formalista de la lógica, página 249.

Gödel, teorema de (incompletud)
La aritmética es incompleta; hay una fbf aritmética que es verdadera, pero que no es ni demostrable ni refutable (Gödel, 1931). Véase pág. 163 para comentarios sobre el papel de la autorreferencia en la prueba de Gödel.

Goldbach, conjetura de
Hipótesis de que todo número par mayor que 2 es la suma de dos números primos.

Implicación
(i) "p" *implica materialmente* "q" ($p \rightarrow q$) si no es el caso que p y no q; "p" *implica estrictamente* "q" ($p \multimap q$) si es imposible que p y no q ($p \multimap q \equiv L(p \rightarrow q)$). Véase cap. 3, § 2, sobre "$\rightarrow$" y "si, entonces"; cap. 10, § 6, sobre las relaciones entre condicionales materiales, estrictos y relevantes, y la idea de *entrañamiento*.
(ii) "Implica" se usa también de otra manera, como "*s* implicó que *p*" (donde la relación es entre hablantes y proposiciones, más que, como antes, entre proposiciones). En este uso se quiere decir algo como "*s* insinuó, aunque no dijo que *p*". Compárese con la discusión de la "implicación conversacional" de Grice, pág. 56.

Independiente
Los axiomas* de un sistema formal son independientes unos de otros si ninguno es una consecuencia* lógica de los otros.

Índice
Expresión cuya referencia depende del tiempo, lugar o hablante, por ejemplo, "ahora", "yo", "aquí". Véase cap. 7, § 6(c).

Inducción
(i) Un argumento es *inductivo en sentido fuerte* si la verdad de sus premisas hace probable la verdad de su conclusión. Véase cap. 2, § 2.
(ii) *Inducción matemática:* una forma de argumento (deductivamente válido) usado en matemáticas, para mostrar que todos los números tienen una propiedad mostrando que 0 tiene esa propiedad, y que si un número tiene esa propiedad su sucesor también la tiene.

Inferencia
Una persona *infiere* q de p si llega a aceptar q teniendo como base p, o llega a aceptar que si p fuera el caso, entonces q sería el caso. Véase ca-

pítulo 12, § 3, sobre la relevancia de la lógica para la inferencia; y capítulo 2, § 3, sobre *reglas de inferencia.*

Interpretación (de un sistema formal)
: Un conjunto (el dominio* D) y una función que asigna elementos de D a términos singulares*, n-tuplos de elementos de D a predicados n-posicionales, y funciones con n-tuplos de elementos de D como argumento y elementos de D como valor, a símbolos de función. Véase sistemas interpretados y no interpretados, págs. 23 y ss.; caps. 4 y 5; cap. 10, § 4, sobre semántica "pura" versus semántica "depravada".

Intuicionismo
: Escuela de filosofía de la matemática (Brouwer, Heyting), caracterizada por la opinión de que los números son construcciones mentales; se apoya en una aritmética restringida y en una lógica no estándar. Véase capítulo 11, § 4.

Lenguaje objeto/metalenguaje
: Si se está hablando sobre sistemas, el sistema del que se habla se conoce como el *lenguaje objeto;* el sistema que se usa para hablar sobre él, el *metalenguaje.* (N.B. ésta es más una distinción relativa que absoluta, por ejemplo, se podría usar el francés (el metalenguaje) para hablar sobre el inglés (el lenguaje objeto) o el inglés para hablar sobre el francés.) Así pues, metalógica* es el estudio de los sistemas lógicos. Véase la discusión del uso de Tarski de la distinción en la definición de verdad, capítulo 7, § 6; cfr. su relevancia para las paradojas semánticas, cap. 8, § 2.

Logicismo
: Escuela de filosofía de la matemática, caracterizada por la tesis (Frege, Russell) de que las verdades de la aritmética son reducibles a la lógica (o *analíticas**, en el sentido de Frege); los números son reducibles a conjuntos. Véase la discusión del programa logicista y la cuestión del ámbito de la lógica, cap. 1, § 2; y del efecto de la paradoja de Russell, página 161.

Matriz característica
: Una matriz es un conjunto de tablas de verdad. Una matriz M es *característica* para un sistema S sii todas y sólo las fbfs* designadas uniformemente* *(tautológicas)** en M son teoremas* de S. Un sistema es n-valente si tiene una matriz característica n-valente y ninguna matriz característica con menos de n valores; *plurivalentes* si es n-valente para $n > 2$, infinitamente plurivalente si es n-valente para infinitas n. Véase cap. 3, § 1, y cap. 11.

Mecánica cuántica
: Una teoría física que trata de la estructura atómica, emisión y absorción de radiaciones por la materia. Véase la discusión de la "lógica cuántica", cap. 11, § 2.

Metalógica
: Estudio de las propiedades formales —por ejemplo, consistencia*, completud*, decidibilidad*— de los sistemas lógicos formales. Véase la discusión de las relaciones entre la filosofía de la lógica y la metalógica, cap. 1, § 1; y de la lógica modal concebida como un cálculo metalógico, págs. 205-206.

Metafísica
: Tradicionalmente "la ciencia del ser en cuanto ser". Uso "metafísico" principalmente para subrayar la distinción entre preguntas sobre el modo en que son las cosas (por ejemplo, "¿Hay una lógica correcta?") y preguntas epistemológicas, preguntas sobre nuestro conocimiento de cómo son las cosas (por ejemplo, "¿Podrían ser las leyes de la lógica distintas de las que consideramos que son?"). Véase cap. 12.

Modus ponens (MPP)
: La regla de inferencia*, para inferir "B" de "A" y "$A \rightarrow B$". Véase la discusión del fracaso del MPP en la lógica de la relevancia, cap. 10, § 7.

Monádico/diádico/poliádico
: Una oración/conectiva abierta es monádica (1-posicional) si tiene un argumento, diádica (2-posicional) si tiene dos, poliádica (pluri-posicional) si tiene más de dos argumentos; por ejemplo, "... es rojo" es una oración abierta monádica, "... es más largo que..." es diádica. Véase la discusión de la función de las secuencias de objetos en las definiciones de satisfacción/verdad de Tarski, págs. 127 y ss.

Monismo/pluralismo/instrumentalismo
: (i) En metafísica*, monismo es la tesis de que hay sólo un último tipo de cosas, dualismo la tesis de que hay dos, pluralismo la tesis de que hay más de dos.
: (ii) Monismo en lógica es la tesis de que sólo hay un sistema lógico correcto, pluralismo la tesis de que hay más de un sistema lógico correcto, instrumentalismo la tesis de que la noción de "corrección" no se aplica a los sistemas lógicos. Véase cap. 12, § 1.

Negación
: La negación de "A" es "$-A$".

Nominalismo/platonismo/conceptualismo
: El *nominalista* niega y el *platónico* afirma que hay *universales reales* (por ejemplo, rojez, cuadratura, etc.); el *conceptualista* declara que los universales son entidades mentales. Terminología relacionada: reísmo, materialismo, pansomatismo (formas del nominalismo) versus realismo (forma del platonismo). Véase la discusión de los cuantificadores* de segundo orden, cap. 4, § 3; la del extensionalismo* de Davidson y el nominalismo de Kotarbinski, pág. 124n; la del status de los mundos posibles, cap. 10, § 4.

Ontología
: Parte de la metafísica que trata de la pregunta sobre qué tipos de cosas existen. Véase cap. 4, § 2, para la discusión de las relaciones entre lógica y ontología.

Oración oblicua
: Habla indirecta como: "s dijo que p". Véase cap. 7, § 6(c).

Paradojas
: (i) (También conocidas como "antinomias".) Contradicciones derivables en teoría de conjuntos* y en semántica*; incluyen la del *Mentiroso* ("Esta oración es falsa") y la *paradoja de Russell* ("El conjunto de todos los conjuntos que no son miembros de sí mismos es un miembro de sí mismo, sii no es miembro de sí mismo"). Véase cap. 8.
: (ii) Las "paradojas" de la implicación* estricta y material son teoremas* de lógica clásica, bivalente y modal ("$p \rightarrow (-p \rightarrow q)$", "$L - p \rightarrow (p \rightarrow q)$)" que parecen bastante contraintuitivos con "$\rightarrow$" o "$\rightarrow$" cuando se lee "si...". Utilizo comillas de advertencia porque estas "paradojas" no encierran contradicción. Véase cap. 10.

Positivismo lógico
: Escuela de filosofía centrada en el *Círculo de Viena* (Schlick, Carnap); caracterizada por el *principio de verificación*, según el cual el significado de un enunciado viene dado por sus condiciones de verificación, y los enunciados no verificables carecen de sentido. Véase la discusión del ataque de Quine a la distinción *analítico/sintético*, cap. 10, § 1.

Postulados de Peano
: Conjunto de axiomas para la teoría de los números naturales:

1. 0 es un número.
2. El sucesor de un número es otro número.
3. Dos números no tienen el mismo sucesor.
4. 0 no es el sucesor de ningún número.
5. Si 0 tiene una propiedad, y, si un número tiene esa propiedad, entonces el sucesor de ese número tiene esa propiedad, entonces todos los números tienen esa propiedad. (Axioma de inducción*.)

Pragmatismo

Escuela americana de filosofía iniciada por Peirce y James (otros pragmatistas incluyen a Dewey y F. C. S. Schiller); caracterizada por la "máxima pragmática", según la cual el significado de un concepto ha de buscarse en las consecuencias prácticas o empíricas (Kant —*pragmatische*— empíricamente condicionadas; griego —*praxis*— acción) de su aplicación. Véase la discusión de la teoría pragmática de la verdad, cap. 7, § 4.

Presuposición

"A" presupone "B" si "A" no es ni verdadero ni falso a no ser que "B" sea verdadero. Véase cap. 5, § 3.

Primitivo

Término no definido (véase definición*).

Principio del círculo vicioso

Poincaré y Russell diagnosticaron a las paradojas* como el resultado de violaciones del principio del círculo vicioso (P.C.V.): "cualquier cosa que encierre el todo de una colección no puede ser un miembro de esa colección". Véase cap. 8, § 2.

Prueba

Una prueba de A (en L) es una deducción* (en L) de A a partir de ninguna otra premisa excepto de los axiomas* (de L), si los hay. Una fbf* A es demostrable (en L) si hay una prueba de A (en L); es refutable si su negación* es demostrable.

Proposicional, actitud

Verbos tales como "sabe", "cree", "espera", etc., que toman la construcción "s Φ que p", se conocen como verbos de actitud proposicional (Russell).

Refutar

Mostrar que una tesis (o teoría, etc.) es falsa. N.B. negar que p *no* es refutar "p".

Relación

Se llama símbolo de *relación* a un predicado de 2 o más posiciones; su extensión* —el conjunto* de los pares ordenados (triplos ... n-tuplos) para los que se establece— se conoce como una *relación de extensión*. Una relación R es *transitiva* si, si (x) (y) (z) si Rxy y Ryz, entonces Rxz; es *simétrica* si (x) (y), si Rxy entonces Ryx; es *reflexiva* si (x) Rxx.

Salva veritate

Sin cambio del valor de verdad.

Satisfacción

(i) En la definición de verdad de Tarski (cap. 7, § 5): relación entre oraciones abiertas y secuencias* de objetos (como, por ejemplo ⟨Edimburgo, Londres, ...⟩ satisfacen "x está al norte de y").

(ii) En lógica imperativa (pág. 106): análogo del valor de verdad, asignado a las oraciones imperativas (como, por ejemplo, "¡Cierra la puerta!" se satisface sii la puerta se cierra).

Secuencia

Par, triplo ... n-tuplo ordenado de objetos (es decir, como un conjunto* excepto por la cuestión del orden; mientras $\{a, b\} = \{b, a\}$, $\langle a, b\rangle \neq \langle b, a\rangle$). Véase la función de las secuencias de objetos en la definición de satisfacción* de Tarski, cap. 7, § 5.

Símbolo incompleto

Expresión definida* contextualmente. Véase cap. 5, § 3; cap. 7, § 7.

Sintaxis/semántica/pragmática

Sintaxis es el estudio de las relaciones formales entre las expresiones: así pues, al vocabulario, a las reglas de formación y a los axiomas*/reglas de inferencia* de un sistema se les llama la sintaxis del sistema. *Semántica* es el estudio de las relaciones entre las expresiones lingüísticas y los objetos no lingüísticos a los que se aplican: así pues, a la interpretación* de un sistema se le llama la semántica del sistema. (Aproximadamente, la distinción entre sintaxis y semántica podría compararse a la que se hace entre gramática y significado.) *Pragmática* es el estudio de las relaciones entre expresiones y el uso o los usuarios de esas expresiones. Véase la interpretación sintáctica y semántica de validez, cap. 2, § 2; aproximaciones sintáctica, semántica y pragmática a las proposiciones, etcétera, cap. 6, § 1; relaciones pragmáticas de la implicación conversacional, pág. 57; y presuposición*, págs. 81 y ss.

Tautología

Sentido técnico: una fbf* es una *tautología* si toma el valor "verdadero" para todas las asignaciones de valores de verdad a sus componentes atómicos* (extendido, en el caso de las lógicas plurivalentes a: si toma un valor designado* para todas las asignaciones a sus componentes atómicos). La prueba de corrección* para el cálculo de oraciones muestra que sólo las tautologías* son teoremas*; la prueba de completud* muestra que todas las tautologías son teoremas.
Sentido no técnico: un enunciado es *tautológico* si dice la misma cosa dos veces, y, por tanto, es trivialmente verdadero. Véase la discusión de la idea presistemática que se corresponde con la noción técnica de verdad lógica*, págs. 34-35.

Teorema

Una fbf* A es un teorema de L sii A se sigue de los axiomas* de L, si los hay, mediante las reglas de inferencia* de L. Véase cap. 2, § 2; cap. 12, § 1.

Teorema de deducción

Si, en un sistema formal L, si

$$A_1 \ldots A_n \vdash_L B, \quad \text{entonces} \quad \vdash_L A_1 \rightarrow (A_2 \rightarrow (\ldots(A_n \rightarrow B)))$$

entonces se establece el teorema de deducción para L.

Teorema de Skolem-Löwenheim

Toda teoría que tiene un modelo (es consistente*) tiene un modelo enumerable (véase la entrada de finito/infinito*). Véase pág. 71 para lo que se refiera a cuantificación sustitucional.

Teoría de tipos

Solución formal de Russell a las paradojas*: la *teoría simple de tipos* evita las paradojas de la teoría de conjuntos*, la *teoría ramificada de tipos* evita las paradojas semánticas. Véase cap. 8, § 2.

Tercio excluso

$p \vee -p$ (cfr. *bivalencia**). Véase cap. 11, § 3.

Término masa

Expresión que denota un tipo de materia o de material (como "agua", "nieve", "hierba") más que, al igual que el "término singular", un objeto individual (como "vaso de agua"). Véase cap. 7, § 6(c).

Válido

Un argumento formal es:
sintácticamente válido en L sii su conclusión se sigue de sus premisas y los axiomas* de L, si los hay, mediante las reglas de inferencia* de L;
semánticamente válido en L sii su conclusión es verdadera en todas las interpretaciones de L en las que todas sus premisas son verdaderas.
Un argumento informal es:
válido sii sus premisas no pueden ser verdaderas y su conclusión falsa. Véase cap. 2, § 2; cap. 10, § 6.

Valor designado

Valor similar al de verdad, tal que todas las fbfs* compuestas que toman un valor designado para todas las asignaciones a sus componentes son *tautologías**.

Variable

Expresión como: "x", "y" ... (en el cálculo de predicados de primer orden) que *recorre* un dominio* de objetos: en contraste con las constantes como: "a", "b" ... cada una de las cuales denota un elemento específico del dominio. Una expresión que puede sustituir a una variable se llama un *substituyente* para la variable; los elementos que recorre, sus *valores*. Una variable *ligada* es la que está dentro del alcance de un cuantificador*, una variable *libre* es la que está fuera del alcance de un cuantificador. Véase cap. 4.

Verdad lógica

Una fbf* es lógicamente verdadera en L sii es verdadera en todas las interpretaciones de L. Véase cap. 2, § 2.

Veritativo-funcional

Una *conectiva* (en las oraciones, operador, formador de oración) es *veritativo-funcional* si el valor de verdad de un compuesto del cual es la principal conectiva depende únicamente de los valores de verdad de sus componentes; en cuyo caso se puede construir una *tabla de verdad* para esa conectiva. Un *sistema lógico* es *veritativo-funcional* si todas sus constantes son veritativo-funcionales. Un sistema n-valente es *funcionalmente completo* —tiene un *conjunto adecuado de conectivas*— si tiene bastantes conectivas para expresar todas las funciones n-valentes de verdad. Ejemplos: las conectivas del cálculo de oraciones clásico y del finito plurivalente son veritativo-funcionales; los operadores modales y los operadores epistémicos no lo son. Véase la discusión de la preferencia de los lógicos por las conectivas veritativo-funcionales, cap. 3, § 2, y la del cálculo plurivalente y no veritativo-funcional, cap. 11.

Verosimilitud

Proximidad a la verdad (Popper); véase cap. 7, § 6(b).

Consejos sobre lecturas

He dado referencias completas en el texto, para capacitar al lector para localizar la literatura relevante sobre temas específicos. El objetivo de la presente sección es dar, en aquello que es nuevo en la materia, unas sugerencias sobre dónde empezar a leer.

He dado por supuesto un conocimiento de la *lógica formal elemental*, como la presentaron Lemmon en 1965, o Quine en 1950, ésta algo más ardua, pero mucho más rica. Una presentación compacta de los resultados de la *metalógica* se puede encontrar en Hunter, 1971, o en Boolos y Jeffrey, 1974. Para la *historia de la lógica*, consúltese Kneale y Kneale, 1962.

Aunque hay varias "*introducciones*" *a la filosofía de la lógica*, por lo general son más duras, y exigen más sofisticación en el lector, de lo que sus títulos sugieren: *Introduction to Logical Theory* de Strawson (1952) presenta una fundamentada crítica de la lógica formal desde el punto de vista de la filosofía del lenguaje ordinario, y debería leerse conjuntamente con la recensión de Quine (1953c); *Philosophy of Logic* de Quine (1970) es, aunque corta, rica y de gran amplitud, pero da por supuesta mucha filosofía característicamente "quineana", y es más adecuada para los estudiantes avanzados que para los principiantes; *Philosophy of Logic* de Putnam (1971) es fiel a un único tema, la necesidad de entidades abstractas en la lógica.

Hay varias *colecciones* valiosas de artículos. Van Heijenoort, 1967a, contiene los artículos clásicos desde la iniciación por Frege de la lógica moderna con la *Begriffsschrift* (1879) al teorema de incompletud de Gödel (1931). Otras colecciones útiles de artículos filosóficos más recientes incluyen la de Copi y Gould, 1967; Strawson, 1967; Iseminger, 1968.

Si desea encontrar lectura sobre un tema específico, pero no sabe por dónde comenzar a buscar, puede encontrar útiles los artículos sobre lógica y filosofía de la *Encyclopaedia of Philosophy* (Edwards, 1967): son informativos en general, y tienen bibliografías útiles. Las recensiones del *Journal of Symbolic Logic* de artículos (filosóficos así como formales) de otras revistas también pueden resultar valiosas. En general, recomiendo que se empiece con materiales de primera mano más que de segunda —que se lean los propios artículos de Frege antes que los de los comentaristas de Frege, por ejemplo; se encontrará que el material de segunda mano es generalmente mucho más útil si ya se ha tenido un conocimiento del trabajo sobre el que se basa.

Algunas sugerencias sobre dónde empezar las lecturas sobre los tópicos discutidos en este libro:

Capítulo

1 Sobre las aspiraciones del formalismo: Frege, 1882a, b.
Sobre el ámbito de la lógica: Kneale, 1956; Quine, 1970, cap. 5.

2 Sobre la inducción y la deducción: Skyrms, 1966, cap. 1.
Sobre *lógica utens* y *lógica docens*: Peirce, "Why study logic?", en 1930-58, vol. 2, especialmente 2, 185 y ss.
Sobre validez y forma lógica: Cargile, 1970; Davidson, 1970; Harman, 1970.

3 Sobre "*tonk*": Prior, 1960, 1964; Belnap, 1961; Stevenson, 1961.
Sobre "si" y "→": Faris, 1962.

4 Sobre el desarrollo de los cuantificadores: Frege, 1891.
Sobre interpretaciones sustitucionales versus interpretaciones objetuales: Belnap y Dunn, 1968.
Sobre los tratamientos no estándars de los cuantificadores: Montague, 1973.

5 Frege, 1892a; Russell, 1905; Strawson, 1950; Quine, "On what there is", en 1953a; Kripke, 1972.

6 Frege, 1918 (y cfr. Popper, "Epistemology without a Knowing subject", en 1972); Quine, 1970, cap. 1; Putnam, 1971; Lemmon, 1966.

7 Definiciones versus criterios: Rescher, 1973, caps. 1 y 2.
Teorías de la correspondencia: Russell, 1918; Austin, 1950; Prior, en Edwards, 1967.
Teorías de la coherencia: Bradley, 1914; Hempel, 1935; Rescher, 1973.
Teorías pragmáticas: Peirce, 1877; James, 1907; Dewey, 1901; Rescher, 1977, cap. 4.
La teoría semántica: Tarski, 1944 (y cfr. Quine, 1970, cap. 3; Rogers, 1963); Popper, "Truth, rationality and the growth of scientific knowledge" (en 1963); "Philosophical comments on Tarski's theory of truth" (en 1972); Davidson, 1967.
La teoría de la redundancia: Ramsey, 1927; Prior, 1971; Grover y otros, 1975.

8 Russell, 1908a; Mackie, 1973, cap. 7; Kripke, 1975.

9 Sobre la lógica temporal: Quine, 1960a, § 36; Prior, 1957, 1967; Lacey, 1971; Geach, 1965.
Sobre lógica vaga: Zadeh, 1975; Gaines, 1976.

10 Sobre verdad necesaria: Quine, 1951.
Presentación formal de las lógicas modales: Hughes y Cresswell, 1968.
Temas filosóficos: Quine, 1953b; Linsky, 1971; Plantinga, 1974.

11 Rescher, 1969; Haack, 1974.

12 Sobre cuestiones metafísicas: van Heijenoort, 1967b; Rescher, 1977, caps. 13, 14.
Sobre cuestiones epistemológicas: Quine, 1951; Putnam, 1969; Popper, 1970.
Sobre psicologismo: Mill, 1843, libro II; Frege (sobre matemática), 1884, especialmente § 26; (sobre lógica) 1918; Peirce, 1930-58, 3, 161 y ss.; Russell, 1938.

Bibliografía

* En una entrada indica a qué edición, y números de página se hace referencia en el texto.

Ackerman, R. (1967), *Introduction to Many-Valued Logics*, Routledge and Kegan Paul.

Alston, W. P. (1958), "Ontological commitments", *Philosophical Studies*, 9; y en Iseminger, 1968.

Altham, J. E. J. (1971), *The Logic of Plurality*, Methuen.

Anderson, A. R. (1970), "St Paul's epistle to Titus", en *The Paradox of the Liar*, ed. Martin, Yale University Press.

Anderson, A. R. y Belnap, N. D. Jr. (1962a), "Tautological entailments", *Philosophical Studies*, 13.

— (1962b), "The pure calculus of entailment", *Journal of Symbolic Logic*, 27 (1962a y b están parcialmente reimpresos en Iseminger, 1968).

— (1975), *Entailment*, vol. 1, Princeton University Press.

Anscombe, G. E. M. (1957), *Analysis* puzzle, 10, *Analysis*, 17.

— (1959), *An Introduction to Wittgenstein's Tractatus* (Hutchinson) [*Introducción al Tractatus de Wittgenstein*, Buenos Aires, Ateneo, 1977].

Austin, J. L. (1950), "Truth", *Proceedings of the Aristotelian Society, Supplement*, 24; y en Urmson y Warnock (eds.), 1961; y Pitcher, 1964.

Ayer, A. J. (1958), *The Problem of Knowledge* (Macmillan) [*El problema del conocimiento*, Buenos Aires, Eudeba].

Baker, A. J. (1967), "'If' and '⊃'", *Mind*, 76.

Bar-Hillel, Y. (1957), "New light on the Liar", *Analysis*, 18.

Barker, S. F. (1965), "Must every inference be either inductive on deductive?", en *Philosophy in America*, comp. Black (Allen and Unwin).

Belnap, N. D. Jr. (1961), "Tonk, plonk and plink", *Analysis*, 22; y en Strawson, 1967.

— (1974), "Grammatical Propaedeutic", en *The Logical Enterprise*, comp. Anderson, Marcus y Martin, Yale University Press, y en Anderson y Belnap, 1975.

Belnap, N. D. Jr. y Dunn, M. (1968), "The substitution interpretation of the quantifiers", *Noûs*, 2.

Belnap, N. D. Jr. y Grover, D. L. (1973), "Quantifying in and out of quotes", en *Truth, Syntax and Modality*, comp. Leblanc, North Holland.

Bennett, J. (1954), "Meaning and implication, *Mind*, 63.

Bergmann, G. (1960), "The philosophical significance of modal logic", *Mind*, 69.