



PROBLEMAS DE EPISTEMOLOGÍA MODAL

JOSÉ TOMÁS ALVARADO MARAMBIO
PONTIFICIA UNIVERSIDAD
CATÓLICA DE VALPARAISO

RESUMEN ABSTRACT

Parece obvio, o casi obvio, que poseemos una gran cantidad de conocimiento modal. Sabemos mucho sobre hechos necesarios que se dan en matemáticas o, por ejemplo, sobre el carácter contingente de objetos físicos con los que tenemos contacto corriente. La cuestión es cómo explicar este conocimiento. Este trabajo, en particular, va a discutir la relevancia que podría tener el dilema propuesto por Paul Benacerraf para el caso matemático sobre enunciados modales, considerando las concepciones sobre hechos modales y sobre mundos posibles de David Lewis y de Saul Kripke.

It seems obvious –or almost- that we do possess a great amount of modal knowledge. We know a lot about necessary facts obtaining in mathematics, and we also know a lot about the contingent character of “middle size dry goods”, for example. The question is how to explain that knowledge. In particular, this work will discuss the impact that the dilemma posed by Paul Benacerraf in the mathematical case would have over modal statements, considering the conceptions about modal facts and possible worlds advanced by David Lewis and Saul Kripke.

PALABRAS CLAVE

Epistemología modal, necesidad, contingencia, Benacerraf, Lewis, Kripke.

KEYWORDS

Modal Epistemology, Necessity, Contingency, Benacerraf, Lewis, Kripke.

Una proposición modal es una proposición en la que se enuncia que algo es necesario o posible, imposible o contingente. Desde un punto de vista formal estas proposiciones se caracterizan porque aparecen antecedidas por un operador específico (un operador modal) de la forma “es necesario que p” o “es posible que p”. Cualquiera de estos operadores puede ser definido a partir del otro. Una proposición modal “refleja” o “representa” un hecho modal que es aquello que hace tal proposición verdadera (o falsa, si es el caso). Así, si es que hay proposiciones modales verdaderas hay también hechos modales como habitantes del mundo. Si llegamos a conocer que una proposición modal es verdadera llegamos también a conocer que hay un hecho modal en el mundo. La cuestión que va a ser considerada aquí es cómo es posible el conocimiento modal dadas las concepciones metafísicas hoy prevalecientes sobre la modalidad. Estas concepciones modales prevalecientes consideran que la mejor forma de especificar las condiciones de verdad de un enunciado modal es haciendo apelación a mundos posibles en los que el contenido de tal proposición se da (o no se da). Hay que tomar aquí como un punto de partida que no será sometido a discusión que existe conocimiento modal, esto es, hay proposiciones modales que nos constan como verdaderas. La cuestión no es justificar que tales proposiciones sean verdaderas sino sencillamente explicar cómo es que hemos llegado a tal conocimiento que, desde un punto de vista intuitivo, parece obvio. Así, resulta obvio que es necesario que $2 + 2 = 4$, como también parece obvio que es posible que esta silla en la que me encuentro sentado no se hubiese construido.

Este trabajo tratará, en primer lugar, sobre la llamada “aporía de Benacerraf” que es la expresión más clara del tipo de aporía epistemológica que surge respecto de los enunciados modales, tal como éstos son ordinariamente interpretados y tal como se suele entender por muchos de nuestros contemporáneos en la epistemología. La aporía de Benacerraf fue formulada para el caso matemático pero es aplicable perfectamente al caso modal de que se trata aquí. Después se considerará cómo se comportan respecto de los problemas que parecen seguirse de la aporía de Benacerraf dos de las principales concepciones sobre la naturaleza de los mundos posibles: el realismo modal extremo de David Lewis y la concepción moderada que se desprende de los trabajos de Saul Kripke. La conclusión de esta consideración será negativa, tal como se verá, pero es un resultado negativo de carácter

provisional. Esto es, hay problemas que todavía siguen abiertos, pero que, en principio, *podrían* ser resueltos con más indagación. No se pretende argumentar aquí que existan problemas que *no pueden* ser resueltos afectando las nociones modales fundamentales.

1. LA APORÍA DE BENACERRAF

Paul Benacerraf en un famoso artículo sobre filosofía de las matemáticas¹ ha hecho notar la existencia de una tensión profunda entre dos grandes estrategias para tratar de dar cuenta de la objetividad de las matemáticas. Señala Benacerraf:

Mi alegato es que dos tipos de preocupaciones muy distintas entre sí han motivado separadamente explicaciones de la naturaleza de la verdad matemática: (1) la preocupación por poseer una teoría semántica homogénea en la que la semántica para las proposiciones matemáticas sea paralela a la semántica para el resto del lenguaje, y (2) la preocupación de que la explicación de la verdad matemática sea coherente con una epistemología razonable. Será mi tesis general que casi todas las explicaciones del concepto de verdad matemática pueden ser identificadas con una u otra de estas líneas *a la expensa de la otra*.²

Por un lado existe la intuición de que la “forma” de enunciados sobre relaciones entre objetos matemáticos y la “forma” de enunciados sobre relaciones entre objetos físicos cualesquiera es la misma. Considérese, por ejemplo:

(1) La silla está a la izquierda de la mesa

Esta proposición indica que hay un objeto que se encuentra en una cierta relación R con otro objeto. Por otro lado, considérese:

(2) $2 \leq 4$

Esta otra proposición (2) también indica, o parece indicar, que hay un objeto –el número 2– que se encuentra en cierta relación S con otro objeto –el número 4–. Desde un punto de vista semántico, no hay diferencias en estructura entre la afirmación de que la silla está a

¹ P. Benacerraf, “Mathematical Truth” en W. D. Hart (ed.), *The Philosophy of Mathematics*, Oxford: Oxford U.P., 1996, 14-30.

² P. Benacerraf, “Mathematical Truth”, 14.

la izquierda de la mesa y decir que 2 es menor o igual que 4. Ambas proposiciones están simplemente enunciando que hay dos objetos que satisfacen cierta relación y parece ser que en ambos casos la condición de verdad del enunciado es el hecho de que esos dos objetos se encuentren en la relación enunciada. Ambas proposiciones, también, admiten una generalización existencial que tiene esta forma:

$$(3) \quad \exists x \exists y Rxy$$

Un anti-platonista estará aquí inclinado a protestar que un mueble y un número son entidades de muy distinta naturaleza, por supuesto, pero esto no impide de ninguna manera que puedan ser tratadas ambas proposiciones como poseedoras de una estructura semántica semejante y de una especificación también semejante de condiciones de verdad. Nada parece importar para que tanto la proposición (1) como la proposición (2) sean verdaderas sino el hecho de que los objetos a los que se hace referencia en tales proposiciones se encuentren en la relación que tales proposiciones enuncian como obteniendo entre ellos. Resulta *prima facie*, por lo tanto, que este modelo semántico que parece imponerse por sí mismo para el tratamiento de los enunciados matemáticos, está operando bajo la presuposición de que hay entidades matemáticas realmente y de que tales entidades matemáticas poseen propiedades y relaciones tal como las mesas y sillas poseen propiedades (como estar hechas de madera) y relaciones (como estar ubicadas una junta a la otra). La dificultad fundamental que provoca esta concepción semántica sobre en qué consiste la verdad de los enunciados matemáticos es que –indica Benacerraf– no parece ser muy coherente con una explicación epistemológica plausible sobre cómo es que llegamos a poseer conocimiento de tales verdades matemáticas. Ciertamente, conocemos que $2 < 4$ y muchas otras verdades matemáticas incluso con mayor certeza con la que conocemos la estructura física del mundo, pero ciertas explicaciones usuales sobre qué exigencias posee el conocimiento de algo hacen difícil pensar que pueda existir “conocimiento” de un dominio de objetos abstractos en los que obtienen propiedades y relaciones de manera objetiva. Este requerimiento epistemológico es enunciado por Benacerraf de esta manera:

Favorezco una explicación causal del conocimiento según la cual para que X conozca que S es verdadera se requiere que cierta relación causal se dé entre X y los referentes de los nombres, predicados y cuantificadores de S . Creo, además, en una teoría

causal de la *referencia*, haciendo de tal manera que la conexión de mi enunciación el conocimiento de *S* sea *doblemente* causal.³

La aporía a la que hace mención Benacerraf consiste en el hecho de que una entidad matemática de aquellas que postula el platonista, esto es, una entidad de carácter abstracto, sin locación espacio-temporal, no es algo con lo que sea posible una relación causal. El requerimiento epistemológico impuesto por Benacerraf es que si es que *X* conoce que *p*, entonces debe existir algún tipo de conexión causal entre *X* y el estado de cosas *p* que es objeto de conocimiento. Esto implica que no podría haber conocimiento de la verdad de los enunciados matemáticos. Se agrega un requerimiento causal para la referencia que viene a completar el requerimiento propiamente epistemológico. Normalmente, nuestro conocimiento de que *p* es el caso supone nuestra capacidad para formular el contenido de la proposición *p* mediante el despliegue de los recursos conceptuales apropiados para tal formulación. Estos recursos conceptuales tienen su correlato en la comprensión del significado de los términos con los que se puede formular una oración apropiada para expresar la proposición que es el contenido del pensamiento de que *p* es el caso o de que *p* podría ser el caso en un cierto lenguaje. Tales términos poseen una referencia que es aquello en que fundamentalmente consiste su significado –dejando a un lado los pruritos que pudiesen tener algunos filósofos del lenguaje–. El requerimiento adicional de Benacerraf de que exista una conexión causal para la referencia de los términos mediante los que se formula la proposición de que *p* puede tomarse, por lo tanto, como el requerimiento de que, aparte de que se requiere cierta conexión causal entre *X* y *p* para llegar *X* a conocer que *p*, se requiere adicionalmente una conexión causal entre *X* y *p*, o algunos elementos integrantes del estado de cosas de que *p* habla, para que *X* pueda siquiera considerar si es que *p* es o no el caso. Pensar en el contenido de *p* requiere también ciertos vínculos causales entre *X* y *p*.

Nótese que el requerimiento epistemológico de Benacerraf no está suponiendo que el conocimiento *es* un vínculo causal determinado entre cognoscente y objeto conocido y nada más. Todo lo que se requiere es que exista algún tipo de vínculo causal, aunque éste por sí sólo sea insuficiente para generar el fenómeno del conocimiento. El requerimiento epistemológico se podría formular así:

³ P. Benacerraf, "Mathematical Truth", 22-23.

(4) $\forall S \forall p (K_{Sp} \rightarrow \exists C C[S, p])$

Lo que expresa esta tesis (4) es que para todo sujeto S y todo estado de cosas p, vale que si S conoce que p, entonces existe algún vínculo causal C tal que C conecta a S y a p. El vínculo causal es necesario para que exista conocimiento, aunque no se exige que este vínculo causal sea suficiente. Por supuesto, es un presupuesto de todo este planteamiento que la relación de causalidad es una relación de carácter físico que sólo puede darse entre entidades físicas. No hay espacio aquí para decir, por ejemplo, que un ángel –tómese como un espíritu puro, sin ningún tipo de composición material, como lo conciben muchos teólogos⁴– ejerza cierta causalidad sobre otro ángel. Tampoco hay espacio para pensar que Dios ejerce cierta causalidad al crear a un ángel, o –por ejemplo– al aniquilarlo, o enviarlo al abismo. De todos modos, aun cuando la idea de causalidad con la que parece operar la aporía de Benacerraf no deja espacio para las interacciones causales angélicas, un teólogo medieval también se encontraría en el mismo tipo de aporía aunque su concepción de la causalidad sea más amplia, pues un número como entidad abstracta no es el tipo de entidad con la que se pueden tener conexiones causales. Ni siquiera un ángel puede afectar el dominio ontológico de los números naturales, si es que existe tal cosa. Tampoco un ángel puede “sufrir” efectos causales de parte de números. Jamás se ha oído que un número haya golpeado o injuriado a un ángel. La aporía, entonces consiste en lo siguiente: el modelo semántico inclina a pensar que existe un dominio objetivo de entidades de las que tratan las proposiciones matemáticas. El problema es que estas entidades (y las propiedades que le atribuyen los enunciados matemáticos) son tales que no es posible que exista conexión causal entre ellas y el sujeto que llega a conocer que ciertas proposiciones, enunciando propiedades y relaciones de tales entidades, sean verdaderas. Dado el principio indicado en la tesis (4) se sigue que no puede haber conocimiento de las matemáticas, cuando es obvio que lo existe. Por otro lado, si se trata de proponer un modelo semántico en el que resulte más explicable la cuestión epistemológica, entonces comienzan a perderse partes importantes de las matemáticas clásicas, tal como son entendidas por los matemáticos. El intuicionismo matemático, por ejemplo, ha debido

⁴ Cfr. por ejemplo Tomás de Aquino, *Summa Theologiae*, I, q. 50, a. 2, c. *Omnis substantia intellectualis est omnino immaterialis*.

reducir de manera drástica el dominio de lo que puede ser objeto de estudio del matemático y, lo que resulta más curioso, ha sido por esto mismo ignorado por el grueso de los matemáticos que siguen ocupados en describir el paraíso de Cantor y otros afines, sin que les importe lo más mínimo un principio como (4).

Hay que hacer notar, por otra parte, en relación con el principio (4), que existe una interpretación de la conexión “causal” entre el sujeto cognoscente y el estado de cosas conocido que la haría trivial. En efecto, una “causa” en sentido muy lato puede tomarse como algo que, en términos generales, “hace” o “provoca” un cambio en otra entidad. Sea el sujeto A en el tiempo t_1 en el que no conoce que p es el caso. En el instante de tiempo t_2 ese sujeto A llega a conocer que p es el caso, por lo que hay un cambio en el sujeto S de no conocer que p a conocer que p. Lo que “provoca” ese cambio o modificación intrínseca en S (pues el estado mental de conocimiento de que p es el caso es una propiedad intrínseca de S; la propiedad del estado de cosas p de ser conocido por S es, en cambio, una propiedad extrínseca que no genera una verdadera modificación en p) es el estado de S de conocimiento de que p es el caso. Un componente del conocimiento de que p es el caso es el mismo estado de cosas p. Por lo tanto, p es “causa” de que S conozca que p es el caso por la sencilla razón de que S ha llegado a conocer que p, sea como sea que haya llegado a ese conocimiento. Hay que hacer notar aquí, en efecto, que el estado mental de conocimiento por parte de un sujeto es un estado de cosas consistente en la conexión entre sujeto cognoscente y objeto conocido por un estado mental en el sujeto en cuestión. Es este estado de cosas el que, a su vez, tiene como elemento integrante un estado de cosas que es el objeto conocido. Los estados de cosas pueden entrar en relaciones causales entre sí⁵. Sucede, por lo tanto, en un sentido trivial, que el estado de cosas de conocer S que p debe ser causado por el estado de cosas p conocido por la sencilla razón de que se trata del conocimiento de que p es el caso. En este sentido trivial si S conoce que $2 + 2 = 4$,

⁵ Cfr. para estas nociones de “estado de cosas” y las conexiones causales entre estos estados de cosas D. M. Armstrong, *A World of States of Affairs*, Cambridge: Cambridge U.P., 1997, 113-138, 202-219. Señala Armstrong: “estamos preguntando qué es lo que en el mundo va a asegurar, hacer verdadero, sostener o servir como fundamento ontológico para la verdad de que a es F. El candidato obvio parece ser el estado de cosas de ser a F. En este estado de cosas (hecho, circunstancia) a y F son puestos conjuntamente” (116). También D. M. Armstrong, *Truth and Truthmakers*, Cambridge: Cambridge U.P., 2004, especialmente 4-25.

entonces el conocimiento de que $2 + 2 = 4$ del sujeto S ha de ser causado (entre otras cosas) por el estado de cosas de ser $2 + 2 = 4$, sin que tenga ninguna importancia el hecho de que ese estado de cosas pueda o no entrar en interacciones físicas con el sujeto S. Es obvio, sin embargo, que Benacerraf no está pensando en esta tesis trivial cuando ha formulado el requerimiento epistemológico. Debe verse este requerimiento como exigiendo que existan conexiones *físicas* entre el sujeto y el objeto conocido. Esto hace que el requerimiento de Benacerraf haga imposible hablar, entre otras cosas, que hay conocimiento en Dios o hablar de que hay conocimiento en los ángeles. En efecto, para que exista una relación causal de tipo físico entre los estados de cosas p y q se requiere que tanto p como q sean entidades físicas, esto es, entidades cuya descripción completa pueda darse entregando la indicación de la totalidad de sus propiedades físicas o bien, que se trate de entidades que poseen un componente que sea él mismo una entidad física. Son propiedades “físicas” aquellas propiedades que resultan relevantes de acuerdo a las teorías físicas; en este caso habría que hablar de una teoría física “ideal” que fuese en el futuro a describir completamente las leyes que rigen la naturaleza (una teoría que hoy *no* poseemos, por cierto) para la determinación del comportamiento de un sistema, de acuerdo a lo que puede predecir un cuerpo de leyes naturales. Si alguno de los términos de la relación falla, entonces no puede sostenerse que la conexión entre ellos sea de carácter físico. Si el sujeto cognoscente es una entidad que no posee ninguna “parte” que pueda ser completamente descrita mediante la indicación exhaustiva de todas las propiedades físicas que posee, entonces no puede decirse que exista una relación causal que pueda satisfacer el requerimiento causal de Benacerraf. Un ángel o Dios no satisfacen este requerimiento, un ser humano concebido de acuerdo a una teoría dualista, en la que pueda distinguirse un “cuerpo” y un “alma” sí lo satisface, porque posee una parte que es una entidad que puede ser completamente descrita mediante la indicación exhaustiva de todas sus propiedades físicas. Por otro lado, una entidad abstracta como un número o construcciones conjuntistas en las que no existen átomos diferentes del conjunto vacío son también entidades que no permiten satisfacer este requerimiento. Tampoco lo podrían satisfacer universales trascendentes⁶. La aporía de Benacerraf surge

⁶ Se dice que un universal es “trascendente” cuando no está instanciado en el mundo actual. Un universal trascendente puede ser instanciado en un mundo posible o puede no ser instanciable. Por ejemplo, hay un universal de ser círculo y un universal de ser cuadrado. Dados dos universales A y B existe el universal estructural consistente en la conjunción de A

precisamente cuando se postula un dominio de entidades matemáticas con las que no pueda entablarse una relación causal-física del tipo descrito.

Lo que interesa aquí, sin embargo, no es el problema en filosofía de las matemáticas sobre cómo dar cuenta al mismo tiempo de la condiciones de verdad de los enunciados matemáticos y de cómo es que llegamos a conocer su verdad cuando la conocemos. Lo que interesa es una generalización de este tipo de aporía para el caso modal. Christopher Peacocke ha visto aquí un desafío de carácter general para cualquier dominio de entidades y de proposiciones que traten sobre las entidades de ese dominio que ha denominado el “desafío de integración” para ese dominio (*integration challenge*) y que consiste en afinar la coherencia entre metafísica –esto es, especificación de las condiciones de verdad– y la epistemología para él⁷. ¿Cómo es que este tipo de dificultad se manifiesta para el caso modal? En el caso modal cuando se afirma que, por ejemplo, es posible que llueva hoy, se está diciendo que hay un mundo posible en el que llueva (en el contexto pragmático en el que esta oración es proferida esto puede tener la implicación de que ese mundo posible podría ser el mundo actual, el hablante está “esperando” la ocurrencia de lluvia actualmente, o tomando precauciones contra ella). Un “mundo posible” es una entidad con la que no pueden haber interacciones causales, ya sea porque se lo concibe como una entidad abstracta, o bien porque se lo concibe como la suma mereológica de todos los objetos que se encuentran en relaciones espacio-temporales entre sí. Se sigue, entonces, que no puede haber conocimiento de esos mundos posibles y, por lo tanto, no puede haber conocimiento de que puede llover hoy. Se verá ahora si dos de las más importantes concepciones metafísicas sobre la modalidad pueden responder de manera satisfactoria esta aporía.

y B. Se sigue, por lo tanto, que ha de existir el universal trascendente de ser un círculo cuadrado que no posee instanciación en ningún mundo posible (cfr. sobre universales trascendentes D. M. Armstrong, *Los universales y el realismo científico*, México: UNAM, 1988 (1978), 100-116; *Universals. An Opinionated Introduction*, Boulder: Westview, 1989, 75-82).

⁷ Cfr. Ch. Peacocke, *Being Known*, Oxford: Clarendon Press, 1999.

2. EL REALISMO MODAL EXTREMO Y LA APORÍA DE BENACERRAF

Se ha venido a denominar “realismo modal extremo” la posición metafísica de David Lewis según la cual las proposiciones modales tratan sobre lo que acaece o lo que no acaece en el dominio constituido por la totalidad de mundos posibles. Estos “mundos posibles” son entidades de la misma naturaleza que el mundo actual. La noción de la “actualidad”, por otra parte, no designa cierto mundo que pueda ser singularizado por algún tipo de privilegio ontológico respecto de los restantes⁸. No se trata, enfáticamente, de que el único mundo real sea el mundo actual y los restantes mundos posibles sean construcciones abstractas, tomando como piezas básicas para la construcción elementos existentes en el mundo actual, generalmente. La noción de “actualidad” es una noción indexical como “yo”, “aquí” o “ahora”. En términos generales, un enunciado de la forma “yo F” es verdadero si y sólo si quien profiere tal enunciado, sea b, es tal que Fb. Un enunciado de la forma “aquí Fb” es verdadero si y sólo si es el caso que el objeto b cae bajo la propiedad F en el lugar en el que es proferido el enunciado. Un enunciado de la forma “ahora Fb” es verdadero si y sólo si el objeto b cae bajo la propiedad F en el tiempo en el que es proferido el enunciado. Del mismo modo, un enunciado de la forma “actualmente Fb” es verdadero si y sólo si el objeto b cae bajo la propiedad F en el mundo posible en el que ha sido proferido el enunciado. Qué sea “actual”, por lo tanto, depende de qué mundo sea aquel en el que se profieren los enunciados de que se trata. Desde nuestra perspectiva *este* mundo es el mundo actual, pero esto no es ningún privilegio del mundo actual, como el hecho de que aquí sea *este* lugar en el que me encuentro, no constituye ningún privilegio para el lugar en el que me encuentro por relación a otros lugares en los que podría encontrarme.

Se ha presentado la metafísica de Lewis como una metafísica que postula solamente individuos y conjuntos compuestos por tales individuos, como elementos primitivos⁹. También en Lewis se admite

⁸ La exposición más desarrollada de la concepción metafísica de Lewis, al menos en lo que respecta a cuestiones modales, se encuentra en *On the Plurality of Worlds*, Oxford: Blackwell, 1986. Una excelente presentación general de la posición de Lewis con referencia a las discusiones que ha generado en J. Divers, *Possible Worlds*, London: Routledge, 2002, 41-165.

⁹ Está exposición sigue estrechamente a J. Divers, *Possible Worlds*, 45-46.

que dados dos individuos cualesquiera existe la suma mereológica que posee exactamente esos dos individuos como partes. En cuanto a los conjuntos, existe el conjunto vacío, hay conjuntos singleton por cada individuo existente, esto es, dado cada individuo 'a' existente, existe además el conjunto cuyo único elemento es 'a', {a}. Luego, dados los axiomas usuales, deben aceptarse todos los conjuntos que puedan ser formados por los procedimientos iterativos. Esto garantiza de inmediato conjuntos para cada cardinalidad transfinita. En este esquema metafísico general los mundos posibles quedan definidos como cierta clase de individuos. Esto es:

- (5) $\forall x [(x \text{ es un mundo posible}) \leftrightarrow (x \text{ es tal que: } \forall y \forall z ((y \text{ es una parte de } x) \& (z \text{ es una parte de } x)) \rightarrow (y \text{ está espacio-temporalmente relacionado con } z))]$

Esto es, un mundo posible debe concebirse como aquellos individuos "máximos" que son la suma mereológica de todos los individuos que se encuentran entre sí relacionados espacio-temporalmente. Dos individuos se encuentran relacionados espacio-temporalmente si y sólo si se encuentran a cierta distancia espacial o se encuentran a cierta distancia temporal. Todo individuo que esté espacio-temporalmente relacionado con una parte de un mundo posible es también parte de ese mundo posible. A esta noción de mundo posible se debe agregar que hay un mundo posible, α , que es el mundo actual. La noción de "actualidad", tal como se ha indicado, es indexical. Un individuo que es parte de un mundo posible es sólo parte de un único mundo posible, esto es, no hay identidades de objetos entre diversos mundos. Otro principio importante para saber cuál es la extensión de la totalidad de mundos posibles es el siguiente:

- (6) $\forall x_1, x_2, \dots, x_n [\exists w (w \text{ está compuesto exactamente por réplicas de } x_1, x_2, \dots, x_n) \& \forall R ((R x_1, x_2, \dots, x_n) \rightarrow \exists w (\text{en } w \text{ es verdadero que réplicas de } x_1, x_2, \dots, x_n \text{ están en la relación } R))]$

Lo que establece este principio (6) es que dado un conjunto cualquiera de individuos existirá un mundo posible tal que va a consistir en la suma mereológica de exactamente réplicas de esos individuos (y no otros). Para estos efectos se dice que dos individuos 'a' y 'b' son "réplicas" si y sólo si 'a' y 'b' poseen exactamente las mismas propiedades intrínsecas. Recuerdese que un individuo no existe más

que en un mundo posible, por lo que no puede ser que dado un conjunto de individuos existentes en un mundo posible exista otro mundo constituido exactamente por *esos mismos* individuos. De la misma manera si dos individuos se encuentran en un mundo posible en cierta relación espacio-temporal, entonces existe un mundo posible compuesto exactamente por réplicas de esos individuos en exactamente esa relación. Recuérdese que el encontrarse dos individuos en una relación espacio-temporal es una propiedad extrínseca de esos individuos.

En esta concepción metafísica modal de Lewis no existe conexión causal posible entre los mundos posibles, aún cuando se trata de entidades “físicas”, esto es, individuos tal como los individuos con los que poseemos contacto perceptivo usualmente y sumas mereológicas de estos individuos. Esto acaece porque no hay ninguna distancia o relación espacio-temporal entre los mundos posibles, y la causalidad física es una relación que debe darse entre entidades que deben encontrarse ubicadas a cierta distancia espacial o temporal. Aún si fuesen definidas las relaciones espacio-temporales en términos causales, esto es, en términos de conexión causal o de posibilidades de conexión causal, el hecho de que no existan relaciones espacio-temporales entre cualesquiera dos mundos posibles diferentes, o entre cualesquiera dos partes de mundos posibles diferentes, hace que no puedan existir relaciones causales entre estos mundos. Sucede, sin embargo, que para Lewis existe un motivo adicional para que no puedan existir conexiones causales entre mundos posibles o partes de mundos posibles. En la concepción lewisiana de la causalidad, ésta debe ser comprendida como cierta relación de dependencia contrafáctica¹⁰. Para Lewis, las condiciones de verdad de un condicional contrafáctico son dadas del siguiente modo: éste es verdadero si y sólo si en la clase de los mundos posibles más cercanos al mundo en el que el condicional es evaluado y en los que el antecedente es verdadero, el consecuente es también verdadero. Un condicional contrafáctico cuyo antecedente es falso en todos los mundos posibles suficientemente cercanos es verdadero de manera vacua. Se dice que dos eventos A y B tienen entre sí una relación causal si es que $\neg A \square \rightarrow \neg B$, esto es, si es que fuese el caso que: si no se hubiese dado A, entonces no se hubiese dado B. En otras palabras, el evento A produce una “diferencia” en el curso de sucesos del mundo,

¹⁰ Cfr. D. Lewis, “Causation” en *Philosophical Papers Volume II*, Oxford: Oxford U.P., 1986, 159-213.

precisamente contribuyendo para que B sea el caso. Nótese bien que en el análisis contrafáctico de la causalidad de Lewis no se trata de que *porque* existe una conexión causal entre los eventos A y B vale un condicional contrafáctico de la forma $\neg A \Box \rightarrow \neg B$, sino que se trata de que decimos que hay conexión causal entre los eventos A y B, esto es, decimos que A causa a B *porque* hay dependencia contrafáctica entre A y B al ser verdadero un condicional de la forma $\neg A \Box \rightarrow \neg B$.

Pues bien, en la concepción de Lewis, si es que existiese un vínculo causal entre el evento A del mundo posible w_1 y el evento B del mundo posible w_2 , para $w_1 \neq w_2$, debería suceder que valiera un condicional contrafáctico de la forma $\neg A\text{-en-}w_1 \Box \rightarrow \neg B\text{-en-}w_2$, pero el problema es que no existe forma en que este condicional pudiese siquiera ser formulado como poseyendo condiciones de verdad, pues para ello se deberían considerar todos los mundos posibles más cercanos en los que el antecedente es verdadero para constatar si también ahí el consecuente es verdadero. Pero, ¿más “cercanos” a qué mundo posible? ¿El mundo posible del antecedente o el mundo posible del consecuente? Decididamente para la concepción contrafáctica de la causalidad de Lewis ésta debe ser entendida como una cuestión que acaece sólo al interior de un mundo y no entre diferentes mundos. En efecto, la causalidad “surge” de la consideración de la diferencia que hace un cierto evento en el curso de acontecimientos del mundo posible en el que ese evento existe, comparando lo que acaece en ese mundo posible con lo que acaece en otros mundos posibles semejantes en los que – suponiendo un “horizonte” (o *background*) de condiciones que se mantienen constantes– se constata cómo es que la ausencia del evento causa va acompañada de manera constante con la ausencia del evento efecto. Es este análisis comparativo el que aquí no puede ser realizado, pues el evento causa no provoca ninguna diferencia en su propio mundo sino que, *per impossibile*, provocaría una diferencia en otro mundo.

Esto pareciera sugerir que aunque en la concepción metafísica de la modalidad de Lewis los mundos posibles vienen dados como entidades de carácter concreto¹¹, al ser simples sumas mereológicas máximas de individuos que son cada uno de ellos entidades físicas que se encuentran entre sí en relaciones espacio-temporales, existe un problema

¹¹ Hay que hacer notar aquí que Lewis no se siente muy cómodo con la distinción entre abstracto y concreto. Sólo bajo algunas formas de efectuar la distinción entre abstracto y concreto puede decirse que los mundos posibles son entidades “concretas” en su concepción. Aquí estas dificultades se han obviado para facilitar la exposición (cfr. D. Lewis, *On the Plurality of Worlds*, 81-86).

del tipo apuntado por la aporía de Benacerraf, pues no es dable un vínculo causal-físico con tales entidades. Como la existencia de algún tipo de conexión causal-física es tomada como una condición necesaria para que exista conocimiento, no puede haber conocimiento modal.

La posición de Lewis, sin embargo, cuando se trata de justificar la existencia de la totalidad de mundos posibles, no es hacer intervenir algún tipo de conexión causal con los mundos posibles diferentes del mundo actual, como podrían ser las conexiones que están implicadas en las percepciones sensoriales ordinarias. Lewis postula la pluralidad de mundos posibles como fruto de una inferencia a la mejor explicación.

¿Por qué creer en una pluralidad de mundos? Porque la hipótesis es útil, y ésta es una razón para pensar que es verdadera. El análisis familiar de la necesidad como verdad en todos los mundos posibles fue sólo el comienzo. En las últimas dos décadas los filósofos han ofrecido muchos otros análisis que hacen referencia a mundos posibles o a individuos posibles que habitan mundos posibles. Considero este logro algo de lo más notable. Pienso que es claro que hablar de *possibilia* ha clarificado cuestiones en muchas áreas de la filosofía de la lógica, la filosofía de la mente, la filosofía del lenguaje y la filosofía de la ciencia –sin mencionar la metafísica por sí misma. Incluso aquellos que oficialmente se burlan no pueden resistir la tentación de ayudarse avergonzados mediante esta forma útil de hablar.

Hilbert llamaba al universo de la teoría de conjuntos un paraíso para los matemáticos. Y él estaba en lo cierto (aunque tal vez no debió haber sido él quien lo debió decir). Sólo debemos creer en la vasta jerarquía de conjuntos y entonces encontramos entidades apropiadas para las necesidades de todas las ramas de las matemáticas. Encontramos también que el escaso vocabulario primitivo de la teoría de conjuntos, extendido mediante definiciones, es suficiente para cubrir nuestras necesidades para los predicados matemáticos. Encontramos también que los escasos axiomas de la teoría de conjuntos son primeros principios suficientes para entregar los teoremas que son el contenido del objeto de estudio (...) Los filósofos pueden querer ver el objeto de estudio reconstruido o re-interpretado, pero los matemáticos insisten en perseguir su objeto en el paraíso y no van a ser sacados de ahí. Su tesis de la pluralidad de conjuntos es fructífera y esto les da una buena razón para creer que es verdadera.¹²

¹² D. Lewis, *On the Plurality of Worlds*, 3-4.

La idea de Lewis es postular el espacio lógico de manera análoga a la jerarquía conjuntista. Así como la teoría de conjuntos se justifica a los ojos de los matemáticos por su fertilidad para domeñar mediante un cuerpo de teoría (y una postulación de entidades correlativa) todas las áreas de la matemática, el filósofo tiene una buena razón para pensar que existe una pluralidad completa de todos los mundos posibles cubriendo el completo espacio lógico por su fertilidad para la clarificación y domeñación de áreas filosóficas de lo más diversas. Si la postulación de la pluralidad de mundos posibles permite explicar todos los fenómenos que acaecen en estos campos, entonces es razonable postular que esta pluralidad existe, en especial, si es que las explicaciones alternativas no son tan fructíferas y son mucho más costosas para conseguir el mismo resultado explicativo. Lewis luego hace notar cómo es que estas finalidades teóricas no parecen poder ser servidas por una pluralidad de mundos posibles entendidos como entidades “abstractas”, tal como lo hace la gran parte de los filósofos actualistas. La pluralidad de mundos posibles como entidades de la misma naturaleza que el mundo actual es la única hipótesis que logra los frutos explicativos que esperamos de tal pluralidad constitutiva del espacio lógico¹³. Se trataría de la mejor explicación de los fenómenos modales y de toda otra vasta gama de cuestiones filosóficas en casi todas las áreas.

¿Es ésta una respuesta aceptable de la aporía de Benacerraf? Lewis explícitamente sostiene que lo es. Es más, sostiene de manera general, contestando a la aporía, que sería una mala cosa reformar la matemática por los pruritos que puedan tener los epistemólogos. Si los epistemólogos no son capaces de explicar cómo es que tenemos conocimiento matemático, entonces más vale revisar nuestra epistemología¹⁴. La cuestión que surge aquí, sin embargo, tiene que ver con que pareciera que existe una importante diferencia entre el caso matemático y el caso de mundos posibles concebidos como individuos constituidos cada uno de ellos por la suma mereológica máxima de todos los individuos que poseen entre sí relaciones espacio-temporales. Quizás el procedimiento más fructífero de la ciencia para avanzar en la comprensión de los fenómenos naturales es la abducción de explicaciones que, postulando muchas veces entidades teóricas, son

¹³ Cfr. D. Lewis, *On the Plurality of Worlds*, 136-191.

¹⁴ Cfr. D. Lewis, *On the Plurality of Worlds*, 108-115. Dice Lewis, “si usted piensa que todo conocimiento requiere contacto (*acquaintance*) causal con la materia de que se trate, pienso que tal cosa es simplemente una generalización apresurada” (112).

capaces de dar cuenta de una pluralidad de fenómenos. Así, ha sido mediante una inferencia a la mejor explicación que se ha postulado la existencia de una entidad como el electrón aún cuando no tenemos percepción directa de estas entidades ni podemos tenerlas, dado como está dotado el ser humano en sus facultades sensoriales. ¿Por qué se postula, entonces, el electrón? Porque *debe* existir algo que sea causalmente responsable de cierta clase de fenómenos y porque la suposición de que hay una cierta entidad que es la que resulta causalmente responsable de esos fenómenos ha resultado luego útil para explicar otros fenómenos de los que no se poseía noticia originalmente. El electrón es también, o parece también, aquella entidad causalmente responsable de esas otras áreas de fenómenos. La inferencia a la mejor explicación por la que se postula la existencia del electrón es una inferencia que supone que la entidad postulada ha de poder explicar *causalmente* un cuerpo de apariencias observacionales dado. El electrón se postula como una entidad concreta, localizada espacio-temporalmente (con todas las peculiaridades cuánticas conocidas, pero esto no impide que se trate de una entidad física con localización espacio-temporal) y capaz de entrar en interacciones físicas. Precisamente esta eficacia físico-causal es lo que constituye el cometido explicativo del electrón como entidad teórica. En el caso de la teoría de conjuntos, en cambio, nadie está pensando en nada parecido. De entrada se trata de entidades abstractas que ni son capaces de localización espacio-temporal ni guardan entre sí relaciones causales de dependencia. Todas las relaciones de dependencia que se esperan entre estas entidades abstractas es la posibilidad de probar ciertas tesis a partir de otras, la posibilidad de “definir” cierto tipo de entidades como propiedades y relaciones que se dan entre entidades de otro tipo, o bien la posibilidad de que cierto dominio de entidades sea isomórfico a otro. La teoría de conjuntos es “fructífera” en matemáticas porque es extremadamente simple y en su seno pueden ser definidas estructuras isomórficas a todas las ramas de la matemática conocidas. Esto es, todas las ramas de la matemática pueden ser vistas como “hablando” de ciertas estructuras conjuntistas.

Pues bien, cuando se trata de la postulación de una pluralidad completa de mundos posibles, tal como lo hace Lewis, ¿de qué tipo de postulación se está hablando? Lewis no puede estar hablando de una postulación que vaya a efectuar una tarea explicativa de tipo causal, tal como sucede cuando se postula –por ejemplo– el electrón. Su modelo debe ser el tipo de explicación por el que se postulan entidades en la matemática, por

el que –por ejemplo y siguiendo la analogía del propio Lewis– se postula a la teoría de conjuntos como la teoría matemática fundamental. El problema que tiene esta identificación, sin embargo, es que el tipo de explicación que se espera de la teoría de conjuntos en matemática *no es* el tipo de explicación que se espera de la postulación de una pluralidad de mundos posibles en metafísica. La “explicación” de una entidad abstracta es una prueba o una “biyección” con alguna otra estructura isomórfica a ella. La “explicación” que se espera de los hechos modales es una “explicación” del tipo que se da cuando se postula un electrón. Se trata de dar cuenta de entes reales, generalmente localizados espacio-temporalmente y capaces de entrar en interacciones causales. Se trata de dar cuenta de qué es realmente –por ejemplo– la naturaleza de cierta entidad física y que constituye, o parece constituir, la raíz de sus poderes causales. Para esto se va a necesitar que la entidad postulada o el tipo de entidades postuladas, efectúen cierto trabajo causal-explicativo, lo que no sucede aquí ni puede suceder. Si, por el contrario, Lewis protestase que el tipo de explicación que él espera de la postulación de una pluralidad de mundos posibles es sencillamente una función de clarificación conceptual sobre las nociones que ya poseemos, ¿cómo creer que del mero contenido de nuestros conceptos ha de postularse una pluralidad de entidades físicas semejantes para las que no existe ni el menor indicio *físico* como para justificar su existencia? Posiblemente Lewis ha hecho descansar la plausibilidad de su alegato en una ambigüedad consentida entre estas dos formas diferentes de explicación, para presentar algo que seguiría el modelo de las explicaciones matemáticas para lograr lo que las usuales explicaciones físicas consiguen en la ciencia natural¹⁵.

En conclusión, puede decirse que la posición metafísica de Lewis, lejos de ser capaz de resolver las dificultades planteadas por la aporía de Benacerraf, las hace todavía más agudas con su postulación de una pluralidad de entidades físicas para las que no existe ni puede

¹⁵ Sobre esta cuestión J. Divers ha sostenido en su excelente trabajo sobre metafísica de mundos posibles que las objeciones epistemológicas contra la teoría de Lewis son dudosas dado el poco desarrollo de la epistemología modal, esto es, dada la falta de claridad de en qué es lo que consiste nuestro conocimiento modal. Es efectivo que existe una considerable falta de claridad sobre las cuestiones epistemológicas implicadas, pero –tal como se ha indicado aquí– esto no impide que exista un grueso problema epistemológico en la postura de Lewis que es lo que se encuentra en la base de la “incredulidad” de la que se queja Lewis como de la principal dificultad que debe enfrentar su teoría (cfr. D. Lewis, *On the Plurality of Worlds*, 133-135; la posición de J. Divers en *Possible Worlds*, 149-165).

existir ningún tipo de evidencia de las que se utilizan para justificar la existencia de este tipo de entidades. En vano se va a pretender, por lo tanto, buscar refugio en la metafísica lewisiana para resolver las dificultades epistemológicas que afectan o pueden afectar a la metafísica modal.

3. NECESIDAD A POSTERIORI

En los influyentes trabajos de Saul Kripke se ha defendido la idea de que existen proposiciones necesarias cuya justificación es a posteriori¹⁶. Estas proposiciones necesarias a posteriori están constituidas por identificaciones como aquella por la que se ha dicho que el Lucero de la Mañana es el Lucero de la Tarde. Toda identidad, con independencia de la fuente por la que es justificada, debe ser necesaria. También son proposiciones necesarias a posteriori las identificaciones teóricas que son el resultado de la investigación empírica ordinaria y por las que, por ejemplo, se llega a decir que el calor es energía cinética molecular promedio. Kripke ha contribuido más que nadie para que en el pensamiento contemporáneo se separen las cuestiones modales de las cuestiones epistemológicas. Paradójicamente esto ha sido una causa de que surjan problemas epistemológicos respecto de la modalidad, tal como habrá ocasión de explicar. En todo caso, una idea que alguien puede sentirse inclinado a sugerir sobre la existencia de proposiciones necesarias a posteriori es si acaso no son éstas un ejemplo inmediato de enunciados modales que han sido justificados por la experiencia y que, por lo tanto, se han conseguido con la intervención de vínculos causales (los vínculos causales implicados en las percepciones sensibles ordinarias).

Las concepciones modales pre-kripkeanas identificaban la necesidad con la analiticidad y con la justificación a priori. Una proposición se decía necesaria cuando su verdad estaba garantizada sin importa cómo es que sea el mundo, y la única forma en que pudiera darse tal cosa –para esta concepción de la primera mitad del siglo XX– es cuando tal situación acaece dado el significado de los términos que se emplean para formular proposiciones sobre el mundo¹⁷. La necesidad surge, por lo tanto, sólo de las convenciones por las que decidimos instaurar

¹⁶ Cfr. S. Kripke, *Naming and Necessity*, Oxford: Blackwell, 1980; "Identity and Necessity" en J. Kim & E. Sosa (eds.), *Metaphysics. An Anthology*, Oxford: Blackwell, 1999, 72-89.

¹⁷ Cfr. por ejemplo, A. J. Ayer, *Language, Truth and Logic*, London: Penguin, 1936, 64-83.

ciertos significados para las palabras que empleamos y nada más. Por supuesto, existen aquí muchos problemas que fueron objeto de una ardua discusión por décadas, pero las discusiones no eran motivo para rechazar la identificación fundamental entre las nociones modales y las nociones epistemológicas. Como sólo la analiticidad es la fuente de que una proposición sea necesaria, no hay ninguna dificultad epistemológica especial para aceptar las proposiciones modales. Kripke, en cambio, disipa de entrada esta confusión de lo ontológico con lo epistémico. En efecto, mientras que el carácter a priori o a posteriori de cierta proposición tiene que ver con el hecho de si tal proposición ha de ser justificada con independencia de la experiencia o haciendo apelación a ella, las nociones de ‘necesidad’ y ‘posibilidad’, en cambio, tratan sobre lo que los entes son o podrían ser objetivamente con independencia de lo que alguien pueda conocer o creer. Señala Kripke:

A veces [el concepto de necesidad] se usa en un sentido epistemológico y puede entonces significar simplemente *a priori*. También, por supuesto, a veces se lo usa en un sentido físico cuando la gente distingue entre necesidad física y necesidad lógica. Pero lo que me ocupa aquí es una noción que no es una noción de la epistemología sino de la metafísica, en algún sentido no-peyorativo (espero). Preguntamos si algo podría haber sido verdadero o podría haber sido falso. Bien, si algo es falso, obviamente no es necesariamente verdadero. Si es verdadero, ¿podría haber sido de otro modo? ¿Es posible que en este respecto el mundo hubiese sido diferente de la forma en que es? Si la respuesta es ‘sí’, entonces este hecho del mundo es contingente. Esto en y por sí mismo no tiene nada que ver con el conocimiento de nadie sobre nada. Es ciertamente una tesis filosófica, y no una cuestión de equivalencia definicional obvia, si todo lo *a priori* es necesario o todo lo necesario es *a priori*. Ambos conceptos pueden ser vagos. Esto puede ser otro problema. Pero, en todo caso, tratan de dos dominios diferentes, dos áreas diferentes, la epistemológica y la metafísica.¹⁸

No hay ninguna objeción de principio para que se den casos de proposiciones necesarias cuya justificación sea a posteriori, o bien de proposiciones contingentes cuya justificación sea a priori. Kripke cree ver casos de ambos tipos de combinaciones, en especial, en el caso de

¹⁸ S. Kripke, *Naming and Necessity*, 35-36; cfr. “Identity and Necessity”, 79-80.

identidades que no pueden ser justificadas a priori y en el caso de identificaciones teóricas en la ciencia natural¹⁹. En primer lugar, toda identidad entre dos individuos cualquiera 'a' y 'b' debe ser necesaria. El razonamiento para esta conclusión es sencillo. Sea por hipótesis que:

$$(7) \quad a = b$$

Por otro lado, es obvio que la identidad de 'a' consigo misma es necesaria, pues esto no sería disputado por quien quiera identificar la necesidad con la aprioricidad.

$$(8) \quad \Box(a = a)$$

La ley de la indiscernibilidad de los idénticos sostiene que dos objetos idénticos han de poseer las mismas propiedades, esto es:

$$(9) \quad \forall x \forall y ((x = y) \rightarrow \forall P (Px \leftrightarrow Py))$$

Pero sucede que 'a' posee la siguiente propiedad:

$$(10) \quad \lambda x \Box(a = x)$$

Por lo tanto, dado (7), 'b' debe poseer también la misma propiedad. Así:

$$(11) \quad \Box(a = b)$$

Por lo tanto, si 'a' y 'b' son idénticos deben ser necesariamente idénticos. Esto choca con el hecho de que frecuentemente las identificaciones de objetos como el Lucero de la mañana con el Lucero de la tarde son descubrimientos empíricos y, por ello, son identidades

¹⁹ El caso de proposiciones contingentes a priori viene dado por casos en los que una descripción es utilizada para fijar la referencia de un término (aunque no se trate de casos en los que el significado del término es una descripción). Por ejemplo, se dice que el nombre "Perico Los Palotes" se impone a quienquiera que haya inventado la rueda. Sucede aquí que la proposición "Perico Los Palotes inventó la rueda" resulta a priori, dada la forma en que se ha fijado el referente para el nombre propio "Perico Los Palotes". Pero también sucede que la proposición "Perico Los Palotes inventó la rueda" es contingente, pues otra persona podría haber inventado la rueda o bien podría haber sucedido que nunca se hubiese inventado la rueda. Algo semejante sucede cuando se utiliza una barra de metal para fijar la magnitud física "metro". Sea esa barra S. Como se ha fijado qué es un metro por la longitud de S, sucede que sostener "S mide un metro" es a priori, pero no es un hecho necesario pues S podría haber sido más larga o más corta y no exactamente de un metro (cfr. S. Kripke, *Naming and Necessity*, 53-57).

que se justifican a posteriori. Esto, sin embargo, sólo indica que la necesidad de una proposición debe tomarse como una cuestión diferente del modo como tal proposición ha sido justificada. Los filósofos antes de Kripke habían visto aquí una suerte de “paradoja” pues parecía obvio que existen identidades contingentes como decir que, por ejemplo:

(12) (El inventor del silogismo) = (el tutor de Alejandro Magno)

Esta proposición (12) parece contingente pues el inventor del silogismo podría no haber sido el tutor de Alejandro Magno. Aristóteles podría no haber inventado el silogismo o podría no haber sido el tutor de Alejandro Magno. El que de hecho el mismo sujeto haya inventado el silogismo y haya sido el tutor de Alejandro Magno es algo contingente. Aquí es donde Kripke introduce la distinción crucial entre un designador rígido y un designador accidental. Un designador rígido refiere al mismo objeto en todos los mundos en los que ese objeto existe. Un designador accidental no refiere al mismo objeto en todos los mundos posibles (en los que la expresión refiere). La cuestión aquí, como es bien conocido, es que un nombre propio como Aristóteles es un designador rígido, mientras que una descripción como “el inventor del silogismo” es un designador accidental. Dado esto no resulta accidental la identidad del inventor del silogismo con el tutor de Alejandro Magno. Es contingente, en efecto, que un mismo individuo sea quien actualmente reciba estas dos determinaciones, pero no es contingente que el individuo denotado actualmente por la descripción “el inventor del silogismo” sea tal que sea idéntico con el individuo que es denotado actualmente por la descripción “el tutor de Alejandro Magno”, pues no es contingente que Aristóteles sea idéntico consigo mismo. Este contraste puede ser presentado como un contraste entre expresiones *de re* y *de dicto*. Así:

(13) $\exists x \exists y ((x \text{ es el inventor del silogismo}) \ \& \ (y \text{ es el tutor de Alejandro Magno}) \ \& \ (x = y))$

Pero también:

(14) $\diamond \neg \exists x \exists y ((x \text{ es el inventor del silogismo}) \ \& \ (y \text{ es el tutor de Alejandro Magno}) \ \& \ (x = y))$

Esto es, hay mundos posibles en los que la persona que inventa el silogismo no coincide con el tutor de Alejandro Magno. En cambio:

(15) $\exists x \exists y ((x \text{ es el inventor del silogismo}) \ \& \ (y \text{ es el tutor de Alejandro Magno}) \ \& \ \square(x = y))$

Esto es, en el mundo actual el individuo que inventó el silogismo y el individuo que fue el tutor de Alejandro Magno son el mismo individuo y *este* individuo es necesariamente idéntico consigo mismo.

Kripke luego desarrolla líneas de argumentación análogas para sostener que las identificaciones teóricas por las que se sostiene que, por ejemplo, el agua es H₂O, el calor es energía cinética molecular promedio o el oro es el elemento de número atómico 79 son necesarias y no contingentes²⁰. También en este caso fue necesario introducir ciertas precisiones para despejar las dudas que se presentaban de inmediato a los críticos embebidos del punto de vista tradicional. En particular, la idea de que el agua es necesariamente H₂O parecía sospechosa pues esto es un descubrimiento empírico y no se trata de una proposición que haya sido justificada a priori. La distinción kripkeana entre lo ontológico y lo epistémico apacigua esta duda, pero subsiste otra cuestión u otra familia de cuestiones. En efecto, ¿no hubiera sido posible que el agua fuese una sustancia diferente que una masa de moléculas de H₂O? ¿Cuál es el significado del término “agua”? Aquí es donde se introduce una poderosa analogía entre el modo de operar los nombres “comunes” que hacen denotación a clases naturales y el modelo semántico de los nombres propios como designadores rígidos conectados por una cadena histórica de usos a un objeto en particular al cual designan en todos los mundos posibles. El punto es que un término como “agua” designa cierta clase natural de sustancia con la que tenemos contacto en el mundo actual (contacto “causal” si se quiere) y que queda determinada como poseyendo una idéntica naturaleza. Es “agua” todo aquello que posea la misma naturaleza o esencia que las entidades con las que de hecho poseemos contacto en el mundo y al que hemos hecho denotación mediante, por ejemplo, una indicación deíctica. Es “agua” todo lo que posea la misma naturaleza que *este* líquido que llena nuestros mares, ríos y lagos y que cumple las funciones metabólicas que cumple el agua en nuestros organismos.

²⁰ Cfr. S. Kripke, *Naming and Necessity*, 115-144. Las consideraciones de Kripke aquí son muy semejantes a las de H. Putnam en “The Meaning of ‘Meaning’”, *Mind, Language and Reality. Philosophical Papers Volume 2*, Cambridge: Cambridge U.P., 1975, 215-271.

Así como en el caso de un nombre propio como Aristóteles uno se podría sentir inclinado a pensar que el significado del nombre viene dado por una descripción o por un conjunto de descripciones, en el caso de un nombre común de aquellos que designan clases naturales podríamos sentirnos inclinados a pensar que su significado pudiera venir dado por cierto conjunto de “notas” que se atribuyen usualmente a las entidades de la clase natural en cuestión. Así, uno podría sentirse inclinado a pensar que el oro tiene como significado “metal amarillo y brillante” o que el agua tiene como significado “sustancia inodora, incolora e insípida”. La cuestión es que si éste fuese el significado de “oro” o de “agua” entonces sería un hecho contingente que el oro es el elemento con número atómico 79 y sería un hecho contingente que el agua es una sustancia compuesta por moléculas de H_2O . En efecto, las mismas apariencias de color amarillo brillante podrían ser producidas en otros mundos posibles para seres racionales como nosotros por otras sustancias que no son sustancias con número atómico 79, como asimismo las mismas apariencias de un líquido inodoro, incoloro e insípido podrían ser producidas por una sustancia de composición química diferente de H_2O . El famoso experimento mental de Putnam pone precisamente esto de relieve. Nuestras intuiciones en estos casos son que sustancias con número atómico diferente de 79 no serían oro y que sustancias que no sean H_2O no serían agua²¹.

Resulta, por lo tanto, o parece resultar, que así como es necesario que el Lucero de la tarde sea idéntico al Lucero de la mañana, pues se

²¹ Estas ideas son demasiado bien conocidas como para insistir en ellas. El experimento mental propuesto por Putnam consiste en suponer que fuese descubierto un planeta en la galaxia casi completamente idéntico a la Tierra, sea la Tierra-Gemela. En la Tierra-Gemela todos los sucesos históricos acaecen exactamente igual que en la Tierra. Puede incluso suponerse que cada uno de nosotros posee un “doble” en la Tierra-Gemela que posee exactamente las mismas propiedades intrínsecas que posee cada uno de nosotros y que posee también exactamente la misma historia personal y exactamente los mismos estados intencionales que posee cada uno de nosotros. En la Tierra-Gemela se habla el mismo idioma (o los mismos idiomas) que en la Tierra, así es que ellos hablan español (quizás sea mejor hablar de español-gemelo). La expedición reporta todo esto pero luego se percató de que lo que los hablantes de la Tierra-Gemela denominan “agua” en su lenguaje y que llena los mares, ríos y lagos, etc. de la Tierra-Gemela es una sustancia que no es H_2O sino, por ejemplo, Arsénico. ¿Cuáles son las intuiciones en este caso sobre la identidad o diferencia de significado? Lo que puso de relieve Putnam es que en este caso nuestras intuiciones nos indican que el término “agua” en el idioma de la Tierra-Gemela tendría un significado *diferente* del término “agua” en nuestro lenguaje. Nótese que, además, la diferencia de significado se produciría aún cuando se hiciese un paralelo en los estados intencionales de los hablantes de la Tierra y la Tierra-Gemela en el siglo XVI, por ejemplo (se entiende que en el siglo XVI de la historia de la Tierra-Gemela que es un paralelo exacto de la historia de la Tierra) y no habría ni una sola diferencia entre tales estados intencionales en lo que se refiere a lo que unos y otros denominan “agua” (cfr. H. Putnam, “The Meaning of ‘Meaning’”, especialmente 223-227; también “Explanation and Reference” en *Mind, Language and Reality*, 196-214).

trata de nombres propios que designan al mismo objeto en todos los mundos posibles, es también necesario que el agua sea H₂O pues se trata de formas de designar entidades que poseen exactamente la misma naturaleza o esencia en todos los mundos posibles. El nombre común "agua" funciona como un designador rígido de la propiedad de ser agua y que es responsable explicativamente de las apariencias usuales del agua para nosotros, esto es, de su carácter inodoro, incoloro e insípido, así como de las funciones metabólicas que cumple el agua en los organismos de seres vivos terrestres. Es exactamente esta propiedad la que denominamos también como la que corresponde a ser una sustancia compuesta de moléculas de H₂O y en cualquier situación contrafáctica el nombre "agua" va a estar designando tal tipo de sustancia en esa situación, así como el nombre propio "Aristóteles" va a estar designando a Aristóteles en cualquier situación contrafáctica que se considere como relativa a Aristóteles. Kripke ha enfatizado que esta cuestión no tiene que ver con el modo en que podrían ser "identificados" individuos, objetos o propiedades en diferentes mundos posibles. Los mundos posibles se especifican como situaciones contrafácticas (situaciones "completas" en un sentido peculiar) en las que se considera si un cierto individuo, Aristóteles por ejemplo, podría haber poseído o no cierta propiedad o relación, por ejemplo, si es que Aristóteles podría haber poseído una formación intelectual diferente de la formación intelectual que de hecho poseía en el mundo actual. Parece que un presupuesto para la inteligibilidad de considerar semejante situación contrafáctica es que se haga referencia al mismo individuo que se designa como Aristóteles en el mundo actual y se haga referencia a la misma propiedad de poseer una formación intelectual determinada que se instancia en el mundo actual. El hecho de que en estos mundos posibles o situaciones contrafácticas completas sea difícil desde un punto de vista epistémico "reconocer" a un individuo o a una propiedad no impide de ninguna manera que la situación contrafáctica considerada sea una situación contrafáctica que envuelve lo que le hubiese acaecido a tal individuo poseyendo o no poseyendo tal propiedad.

Se ha hecho notar que la estrecha conexión que parece existir entre el esencialismo defendido por Kripke y la semántica de nombres propios y nombres comunes no implica que esta semántica obligue a aceptar tesis esencialistas fuertes. Kripke y otros autores han defendido diferentes tesis esencialistas, como la de la identidad de un objeto

conigo mismo, el origen de un objeto, la propiedad “sortal” de un objeto que lo determina como un tipo de objeto o, en el caso de un conjunto, los elementos que lo componen²². Esto, sin embargo, no debe tomarse como si se tratase de una posición filosófico-semántica compacta. La idea de que los nombres propios y los nombres comunes tienen como significado un componente indexical irreductible a una entidad o a una clase natural de entidades con la que existe una conexión ininterrumpida de usos lingüísticos desde el bautismo inicial hasta el uso presente, no tiene que tomarse como la idea de que tales entidades individuales o tales clases naturales poseen una esencia o una naturaleza con las propiedades modales usualmente contempladas por los defensores del esencialismo. Esto es, no se trata de que *deba* existir un conjunto de propiedades poseídas por todo objeto *x* al que se haga referencia en todos los mundos posibles en que ese objeto *x* exista –una esencia de *x*– para que sea inteligible hacer referencia a *x* en el mundo actual de la forma en que lo hacemos. Mucho menos es necesario que, por ejemplo, el origen del objeto *x* en cuestión sea esencial a *x*, o la propiedad sortal de *x* sea esencial a *x*. La cuestión de si existe o no una “esencia” de *x* es una cuestión independiente del modelo semántico para comprender el significado de los nombres propios que podemos emplear para hacer referencia a ese objeto *x*. La cuestión ulterior de si el origen de *x* y su propiedad sortal deban ser consideradas como parte de esa esencia son también cuestiones independientes. El núcleo de la llamada teoría de la “referencia directa” defendida por Kripke y por Putnam tiene que ver con que el significado de un nombre propio está constituido por la referencia a un objeto particular y no por las descripciones que los hablantes puedan creer que son verdaderas de ese objeto. De la misma manera, el núcleo de esta teoría en lo que respecta a nombres comunes consiste en que el significado de tales nombres se constituye por la referencia a ciertas clases naturales con independencia de las descripciones que los hablantes puedan creer que son verdaderas de las entidades que pertenecen a tales clases naturales²³.

Lo que sí se debe hacer notar a este respecto es que, aunque no se pueda derivar ninguna tesis esencialista no trivial a partir de la teoría

²² Cfr. las obras citadas de Kripke y también D. Wiggins, *Sameness and Substance Renewed*, Cambridge: Cambridge U.P., 2001, 107-138 sobre la necesidad de un sortal.

²³ Una exposición ya clásica de la independencia entre esencialismo y la teoría de la referencia directa se encuentra en N. Salmon, *Reference and Essence*, Oxford: Blackwell, 1982; “How Not to Derive Essentialism from the Theory of Reference” *The Journal of Philosophy* 76 (1979), 703-725.

de la referencia directa²⁴, las argumentaciones de Kripke y de Putnam han puesto de relieve que nuestras intuiciones ordinarias sobre cómo está constituido el mundo son, generalmente, esencialistas. Precisamente lo que permite hacer el contraste entre, por ejemplo, un designador rígido y un designador accidental, por lo tanto, aquello en lo que se apoya la diferencia intuitiva entre una descripción definida y un nombre propio, es que parece razonable para nuestras intuiciones comunes pensar que Aristóteles podría haber tenido muchas propiedades que de hecho no posee en el mundo actual o que podría haber carecido de muchas propiedades que de hecho posee en el mundo actual sin que por esto deje de ser Aristóteles, el mismo individuo. Sólo es inteligible la identidad de Aristóteles en todos los mundos posibles en los que existe y en los que debe poseer propiedades diferentes si es que hay algo que hace que Aristóteles sea Aristóteles en cada uno de tales mundos. Por supuesto, también podría alguien sostener que hablar de situaciones contrafácticas en las que pudiese verse envuelto Aristóteles, esto es, el mismo individuo que designamos como Aristóteles en el mundo actual no tiene sentido, mientras al mismo tiempo acepta la teoría de la referencia directa para la semántica del nombre propio "Aristóteles", pero precisamente la teoría de la referencia directa adquiere para nosotros verosimilitud por el contraste que provocan nuestras intuiciones modales entre el nombre propio "Aristóteles" con una descripción contingente de Aristóteles. Si no hay identidad de Aristóteles en diferentes mundos posibles este contraste entre designadores rígidos y accidentales no puede hacerse. De la misma manera, lo que hace plausible el modelo semántico propuesto para los nombres comunes es que se contempla lo que podría suceder en mundos posibles en los que una sustancia

²⁴ Existen infinitas propiedades esenciales triviales que pueden ser atribuidas a cualquier objeto. Estas propiedades consisten en propiedades extrínsecas que tienen que ver con las leyes lógicas que valen en todos los mundos posibles. Así, por ejemplo, comoquiera que sea el objeto 'a', en todos los mundos posibles en los que 'a' existe posee la propiedad de ser tal que no cae en la propiedad de caer al mismo tiempo bajo P y no caer bajo P, para una propiedad P cualquiera. También en todos los mundos posibles es tal que cae bajo la propiedad de caer bajo la propiedad P si es que cae bajo la propiedad P, para una propiedad P cualquiera, etcétera. Otra propiedad esencial trivial de 'a' es ser idéntico a 'a'. Y otras clase de propiedad esencial trivial es la que resulta de las propiedades indexadas a un mundo. Esto es, si 'a' posee la propiedad P en el mundo posible w_1 , entonces en todos los mundos posibles en los que 'a' existe posee la propiedad de ser P-en- w_1 , para todas las predicaciones verdaderas de 'a' en w_1 . Es obvio que no se está hablando aquí de estas esencias triviales (cfr. R. Stalnaker, "Anti-Essentialism" en *Ways a World Might Be. Metaphysical and Anti-Metaphysical Essays*, Oxford: Clarendon Press, 2003, 71-85, especialmente 73; la idea de propiedades indexadas a un mundo proviene de A. Plantinga, *The Nature of Necessity*, Oxford: Clarendon Press, 1974, 62-65).

diferente de H_2O posee la misma apariencia que tiene para nosotros el agua en el mundo actual. Nuestras intuiciones nos indican que eso *no es* agua aunque lo parezca externamente y, por lo tanto, tales intuiciones indican que la esencia del agua es para nosotros ser un compuesto de H_2O . Por supuesto, estas intuiciones no son sacrosantas e intocables, pueden ser gravemente incoherentes o pueden ser sencillamente falsas, pero debemos suponer, sin embargo, que el hecho de que parezcan recomendadas por nuestros juicios espontáneos hace que deban ser consideradas como gozando de cierta presunción a su favor mientras no aparezcan razones de peso en contra.

La cuestión que interesa considerar aquí, sin embargo, es qué relevancia tienen estas consideraciones kripkeanas para responder a las dificultades epistemológicas planteadas por Benacerraf para el caso matemático. En particular, el hecho de que ciertas proposiciones necesarias posean una justificación a posteriori, ¿no es una indicación de que se trata de conocimientos que se adquieren precisamente suponiendo una conexión “causal” con el estado de cosas de que se trata, tal como Benacerraf lo ha requerido? Si se llega a conocer, por ejemplo, que el Lucero de la mañana es el Lucero de la tarde es mediante investigación empírica. Si se llega a conocer que el agua es H_2O es también mediante investigación empírica que no está sujeta a ningún tipo de objeciones epistemológicas. La cuestión, sin embargo, es que el carácter empírico de la justificación de estas proposiciones no alcanza a su carácter específicamente modal, que es lo que está en discusión aquí. En un revelador pasaje señala Kripke, discutiendo la tesis de la necesidad del material de que está compuesto un objeto:

Si la visión esencialista es correcta, sólo puede serlo si es que se distingue radicalmente entre las nociones de verdad a posteriori y a priori, por una parte y verdad contingente y necesaria, por otra parte, pues aunque el enunciado que esta mesa, si es que existe, no fue hecha de hielo, es necesario, esto ciertamente no es algo que conozcamos a priori. Lo que sabemos es que, primero, las mesas no están usualmente hechas de hielo, ellas están usualmente hechas de madera. Esto parece de madera. No se siente frío y probablemente se sentiría frío si estuviese hecho de hielo. Por lo tanto, concluyo, probablemente esto no está hecho de hielo. Aquí todo mi juicio es a posteriori (...) Por lo que debemos decir que aunque no podemos conocer a priori si esta mesa fue hecha de hielo o no, dado que no está hecha de hielo, *necesariamente* no está hecha de hielo. En otras palabras, si P es el enunciado que la mesa

no está hecha de hielo, uno conoce por análisis filosófico a priori algún condicional de la forma "si P entonces necesariamente P ". Si esta mesa no está hecha de hielo, es necesario que no esté hecha de hielo. Por otro lado, entonces, conocemos por investigación empírica que P , el antecedente del condicional, es verdadero –que esta mesa no está hecha de hielo. Podemos concluir por *modus ponens*:

$$\begin{array}{l} P \supset \Box P \\ \underline{P} \\ \Box P \end{array}$$

La conclusión –" P " – es que es necesario que la mesa no esté hecha de hielo, y esta conclusión es conocida a posteriori, pues una de sus premisas en las que se basa es a posteriori.²⁵

Esta distinción es relevante para todos los casos en que se presentan o pueden presentarse proposiciones necesarias a posteriori. Existe un componente a posteriori en tales verdades, que se determina mediante percepción sensible, o mediante las investigaciones empíricas usuales que efectúa la ciencia natural para la determinación de un hecho. Luego, hay un condicional según el cual se conecta a ese hecho como antecedente una consecuencia de carácter modal. Es este condicional, donde se introduce lo específicamente modal, el que se justifica mediante "análisis filosófico a priori". Ese condicional no es justificado mediante la experiencia, no hay en su justificación conexiones causales con ningún estado de cosas del mundo actual, de manera directa o indirecta. Siguiendo el estilo del ejemplo kripkeano, se puede considerar cómo es que se sabe que el agua es necesariamente H_2O . Existe un conjunto de información sobre qué sustancia es la que llena los mares, ríos y lagos, qué sustancia es la que bebemos ordinariamente y por qué es que necesitamos beberla para sobrevivir. Esta información se determina mediante una indagación totalmente a posteriori. Se constata en condiciones experimentales tales y cuales cosas que indican la presencia de un compuesto molecular de dos átomos de hidrógeno y un átomo de oxígeno. Es una verdad a posteriori que el agua es H_2O . Luego debe postularse un condicional de esta forma:

$$(16) \text{ (el agua es } H_2O) \rightarrow \Box(\text{el agua es } H_2O)$$

²⁵ S. Kripke, "Identity and Necessity", 81.

Dado el antecedente determinado como una proposición a posteriori se sigue el consecuente modal según el cual el agua es necesariamente H_2O . Como puede notarse, la conclusión de que el agua es necesariamente H_2O depende de una fuente a posteriori, por lo que debe decirse que la proposición ha sido justificada sin independencia de la experiencia. El punto, sin embargo, es que la premisa de la que depende el carácter específicamente modal de la conclusión tiene una fuente independiente de la experiencia, a priori. Uno puede aquí preguntar, ¿cómo es que se conoce la verdad de un condicional como (16)? Es tal premisa modal la que está haciendo las familiares apelaciones a mundos posibles con los que no existe conexión causal y es este tipo de apelación lo que suscita los familiares pruritos epistemológicos.

Hay que hacer notar, además, que éste no es un problema que afecte de manera selectiva a las proposiciones necesarias. Las proposiciones contingentes poseen exactamente el mismo tipo de dificultad. Considérese, por ejemplo, esta proposición:

(17) Micifuz cazó un ratón

Si se estima que (17) es contingente, esto es que:

(18) \diamond (Micifuz cazó un ratón) & $\diamond \neg$ (Micifuz cazó un ratón)

Se sigue que hay mundos posibles en los que Micifuz caza un ratón y mundos posibles en los que Micifuz no caza un ratón. Consta en el mundo actual que Micifuz cazó un ratón por conocimiento a posteriori, por lo que podría considerarse justificado a posteriori el primer término de la conjunción, pero subsiste el problema con el segundo término de la conjunción. Nuevamente aquí debe suponerse que existe un componente estrictamente modal cuando se juzga la contingencia de la proposición (18), un componente cuya justificación específica no es a posteriori sino a priori. Se trata de cierto conocimiento de lo que podría acaecer en una situación contrafáctica (el que Micifuz no haya cazado un ratón en esos mundos) que no es dada a nuestro conocimiento empírico. Respecto de tal componente va a surgir un problema epistemológico tal como ha surgido para el caso de proposiciones necesarias que están haciendo afirmaciones sobre lo que acaece en la totalidad de los mundos posibles. Sucede, por lo tanto, que las dificultades

epistemológicas del estilo de las indicadas por Benacerraf para la fundamentación de las matemáticas no están restringidas a los juicios sobre necesidad o imposibilidad, sino que abarcan también los juicios sobre contingencia. La modalidad de una proposición no resulta mejor fundamentada cuando se considera el hecho de que las fuentes de conocimiento modal se encuentran muchas veces en juicios justificados a posteriori, tal como lo han puesto de relieve Kripke y Putnam.

Las ideas de Kripke y Putnam sobre la semántica de nombres propios y comunes y, luego, las ideas sobre proposiciones necesarias a posteriori han resultado cruciales para comprender la autonomía del dominio ontológico de lo modal. Han contribuido también poderosamente para aceptar el carácter generalizado de los presupuestos modales en nuestra comprensión más básica de la realidad. Los juicios sobre lo necesario y lo posible no son un resabio de oscurantismo escolástico, sino que se encuentran firmemente imbricados en lo más profundo de nuestra comprensión de la naturaleza por leyes, sin mencionar nuestra comprensión de nosotros mismos como seres humanos para quienes el mundo se despliega como un horizonte de posibilidades abiertas que exigen de nosotros decisiones libres. La nueva concepción de lo modal como un ámbito ontológico objetivo, sin embargo, ha atraído de inmediato las dificultades epistemológicas que se están discutiendo aquí. La separación de lo modal respecto de lo epistémico ha exigido una explicación de cómo es que existe conocimiento de lo modal, si es que se trata de un dominio ontológico objetivo, tan objetivo como el dominio de las entidades y propiedades físicas. Y es este tipo de problema aquel para el que todavía no tenemos una respuesta satisfactoria²⁶.

²⁶ Una versión preliminar de este trabajo fue presentado en el II Coloquio de Filosofía organizado por el Departamento de Artes y Humanidades de la Universidad Andrés Bello el 20 de octubre de 2005. Agradezco los comentarios y sugerencias de los participantes en ese Coloquio.